CCNA to dopiero początek.
Opanuj sieci CISCO w praktyce!

Wojownik sieci

HELION

O'REILLY®

Gary A. Donahue

okpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Tytuł oryginału: Network Warrior, 2nd edition

Tłumaczenie: Lech Lachowski (wstęp, rozdz. 1 – 21),
Marek Pałczyński (rozdz. 22 – 31),
Marcin Komarnicki (rozdz. 32 – 41)

Korekta merytoryczna: Marcin Romanowski

ISBN: 978-83-246-5201-3

© 2012 Helion S.A.

Authorized Polish translation of the English edition of *Network Warrior, 2nd Edition,*
ISBN 9781449387860 © 2011 Gary Donahue. All rights reserved.
This translation is published and sold by permission of O'Reilly Media, Inc.,
which owns or controls all rights to publish and sell the same.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from the Publisher.

Wszelkie prawa zastrzeżone. Nieautoryzowane rozpowszechnianie całości lub fragmentu niniejszej publikacji w jakiejkolwiek postaci jest zabronione. Wykonywanie kopii metodą kserograficzną, fotograficzną, a także kopiowanie książki na nośniku filmowym, magnetycznym lub innym powoduje naruszenie praw autorskich niniejszej publikacji.

Wszystkie znaki występujące w tekście są zastrzeżonymi znakami firmowymi bądź towarowymi ich właścicieli.

Autor oraz Wydawnictwo HELION dołożyli wszelkich starań, by zawarte w tej książce informacje były kompletne i rzetelne. Nie biorą jednak żadnej odpowiedzialności ani za ich wykorzystanie, ani za związane z tym ewentualne naruszenie praw patentowych lub autorskich. Autor oraz Wydawnictwo HELION nie ponoszą również żadnej odpowiedzialności za ewentualne szkody wynikłe z wykorzystania informacji zawartych w książce.

Wydawnictwo HELION
ul. Kościuszki 1c, 44-100 GLIWICE
tel. 32 231 22 19, 32 230 98 63
e-mail: *helion@helion.pl*
WWW: *http://helion.pl* (księgarnia internetowa, katalog książek)

Drogi Czytelniku!
Jeżeli chcesz ocenić tę książkę, zajrzyj pod adres
http://helion.pl/user/opinie/wojsie_ebook
Możesz tam wpisać swoje uwagi, spostrzeżenia, recenzję.

Printed in Poland.

- Poleć książkę na Facebook.com
- Kup w wersji papierowej
- Oceń książkę

- Księgarnia internetowa
- Lubię to! » Nasza społeczność

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Spis treści

Przedmowa **9**

1. Co to jest sieć? **15**

2. Huby i switche **19**
Huby 19
Switche 23

3. Autonegocjacja **33**
Co to jest autonegocjacja? 33
Jak działa autonegocjacja? 34
Kiedy autonegocjacja zawodzi 35
Autonegocjacja w praktyce 37
Konfiguracja autonegocjacji 37

4. VLAN-y **39**
Łączenie VLAN-ów 39
Konfiguracja VLAN-ów 43

5. Trunking **51**
Jak działają kanały trunkowe 51
Konfiguracja trunków 55

6. Protokół VTP **61**
VTP Pruning 64
Zagrożenia VTP 65
Konfiguracja VTP 66

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

7. Agregacja portów ... 73
EtherChannel 73
Cross-stack EtherChannel 84
Multichassis EtherChannel (MEC) 84
Virtual Port Channel 84

8. Spanning Tree ... 91
Burza broadcastowa 92
Niestabilność tablicy adresów MAC 96
Zapobieganie powstawaniu pętli przez Spanning Tree 97
Zarządzanie Spanning Tree 101
Dodatkowe funkcje Spanning Tree 104
Najczęstsze problemy ze Spanning Tree 108
Zasady projektowania umożliwiające uniknięcie problemów ze Spanning Tree 112

9. Routing i routery ... 113
Tabele routingu 114
Rodzaje tras 116
Tabela routingu IP 117
Virtual Routing and Forwarding 122

10. Protokoły routingu ... 125
Komunikacja między routerami 126
Metryki i rodzaje protokołów 128
Dystans administracyjny 130
Wybrane protokoły routingu 132

11. Redystrybucja ... 151
Redystrybucja w protokole RIP 153
Redystrybucja w protokole EIGRP 156
Redystrybucja w protokole OSPF 158
Redystrybucja wzajemna 159
Pętle redystrybucji 160
Ograniczanie redystrybucji 162

12. Tunele ... 169
Tunele GRE 170
Tunele GRE i protokoły routingu 174
GRE i listy dostępu 179

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

13. Protokoły redundancji 181
HSRP 181
Śledzenie interfejsu HSRP 184
Kiedy HSRP nie wystarcza 186
Nexus i HSRP 188
GLBP 189

14. Mapy routingu 197
Budowa mapy routingu 198
Przykład polityki routingu 200

15. Algorytmy przełączania w routerach Cisco 207
Przełączanie procesowe 209
Przełączanie kontekstu z obsługą przerwań 210
Ścieżki przełączania — konfiguracja i zarządzanie 215

16. Przełączniki wielowarstwowe 221
Konfiguracja interfejsów SVI 223
Modele przełączników wielowarstwowych 227

17. Przełączniki wielowarstwowe Cisco 6500 231
Architektura 233
CatOS kontra IOS 249
Instalacja VSS 252

18. Cisco Nexus 263
Nexus — dostępny sprzęt 263
NX-OS 271
Ikonografia Nexusa 277
Cechy architektury Nexusa 277

19. Cechy przełącznika Catalyst 3750 313
Łączenie w stos 313
Zakresy interfejsów 315
Makra 316
Linki flex 319
Kontrola burz 320
Zabezpieczenie portów 324
SPAN 327
VLAN głosowy 330
QoS 332

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

20. Nomenklatura telekomunikacyjna 335
Słownik terminów telekomunikacyjnych 336

21. T1 347
Jak działa T1 w dupleksie 347
Rodzaje T1 348
Kodowanie 349
Ramkowanie 351
Monitorowanie wydajności 354
Alarmy 355
Rozwiązywanie problemów z T1 358
Konfiguracja T1 362

22. DS3 367
Ramkowanie 368
Kodowanie liniowe 371
Konfiguracja łączy DS3 371

23. Frame Relay 377
Zamawianie usługi Frame Relay 380
Projektowanie sieci Frame Relay 381
Nadsubskrypcja 383
Lokalny interfejs zarządzania 384
Konfiguracja protokołu Frame Relay 385
Rozwiązywanie problemów z protokołem Frame Relay 393

24. MPLS 397

25. Listy dostępu 403
Definiowanie list dostępu 403
Listy ACL w przełącznikach wielowarstwowych 414
Refleksyjne listy dostępu 418

26. Uwierzytelnianie w urządzeniach Cisco 425
Uwierzytelnianie standardowe (nie AAA) 425
Uwierzytelnianie AAA 437

27. Podstawy działania zapór sieciowych 445
Sprawdzone rozwiązania 445
Strefa DMZ 447
Rozwiązania alternatywne 451

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

28. Konfiguracja zapór sieciowych ASA 455
Konteksty 456
Interfejsy i poziomy bezpieczeństwa 456
Nazwy 459
Grupy obiektów 460
Inspekcje 463
Zarządzanie kontekstami 465
Przełączanie awaryjne 475
NAT 486
Operacje uzupełniające 490
Rozwiązywanie problemów 493

29. Sieci bezprzewodowe 495
Standardy komunikacji bezprzewodowej 496
Bezpieczeństwo 497
Konfiguracja punktu dostępowego 500
Rozwiązywanie problemów 504

30. VoIP 507
Jak działają systemy VoIP 507
Przykład systemu VoIP w małym biurze 515
Rozwiązywanie problemów 549

31. Podstawy QoS 555
Rodzaje mechanizmów QoS 558
Działanie mechanizmów QoS 560
Typowe przypadki niezrozumienia mechanizmu QoS 567

32. Projektowanie QoS 571
Scenariusz LLQ 571
Konfiguracja routerów 575
Scenariusze kształtowania ruchu 579

33. Sieci przeciążone 589
Rozpoznawanie sieci przeciążonych 589
Rozwiązanie problemu 594

34. Sieci konwergentne 595
Konfiguracja 595
Monitorowanie QoS 597
Rozwiązywanie problemów w sieciach konwergentnych 599

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

35. Projektowanie sieci 605
Dokumentacja 605
Konwencja nazewnicza dla urządzeń 614
Projekty sieci 615

36. Projektowanie sieci IP 625
Publiczne a prywatne zakresy IP 625
VLSM 627
CIDR 630
Przyznawanie przestrzeni adresowej w sieciach IP 631
Przydzielanie podsieci IP 633
Łatwe podsieciowanie 637

37. IPv6 645
Adresacja 647
Prosta konfiguracja routera 652

38. NTP 661
Czym jest dokładny czas? 661
Projektowanie NTP 662
Konfiguracja NTP 664

39. Awarie 669
Błąd ludzki 669
Awaria wielu komponentów 670
Łańcuch katastrof 671
Brak testów awarii 672
Rozwiązywanie problemów 672

40. Maksymy GAD-a 677
Maksyma 1. 677
Maksyma 2. 679
Maksyma 3. 681

41. Unikanie frustracji 685
Dlaczego wszystko jest takie zagmatwane 685
Jak przedstawiać swoje pomysły kadrze zarządzającej 688
Kiedy i dlaczego aktualizować 691
Dlaczego kontrola zmian jest Twoim przyjacielem 694
Jak nie być komputerowym głupkiem 695

Skorowidz 699

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przedmowa

Przykłady praktyczne zamieszczone w tej książce zostały zaczerpnięte z mojego własnego doświadczenia oraz doświadczenia tych, z którymi i dla których miałem przyjemność pracować. Z oczywistych powodów natury prawnej oraz ze względu na uczciwość dokładne dane, a także wszelkie informacje mogące ujawnić tożsamość osób trzecich zostały zmienione.

Wszystkie przykłady, z nielicznymi wyjątkami, oparte są na protokole TCP/IP i zostały w nich wykorzystane urządzenia Cisco. Można by się spierać, czy w tego typu publikacji nie powinno się zawrzeć przykładów dotyczących również innych protokołów oraz sprzętu różnych producentów. Do pewnego stopnia jest to argument słuszny. Jednak książka opisująca szeroką gamę technologii przy jednoczesnym przedstawianiu przykładów opartych na urządzeniach większej liczby producentów obejmowałaby niepraktycznie duży zakres materiału. Faktem jest, że Cisco Systems (z pewnością ku wielkiemu rozczarowaniu konkurencji) jest wiodącym graczem w dziedzinie sieci. Podobnie ma się rzecz z TCP/IP, który jest protokołem internetowym i wykorzystywany jest przez większość urządzeń sieciowych. Czy jest on jednak najlepszym protokołem dla tych zastosowań? Być może nie, ale jest obecnie protokołem powszechnie stosowanym, więc użyłem go we wszystkich moich przykładach. Jeszcze nie tak dawno temu egzaminy Cisco CCIE obejmowały takie zagadnienia jak mostkowanie trasy źródłowej Token Ring (ang. *Token Ring Source Route Bridging*), AppleTalk czy IPX. Jednak te czasy już minęły, co pokazuje, że nawet Cisco zrozumiało, iż TCP/IP to właściwy kierunek. To wydanie zawiera również dodatkowy rozdział opisujący protokół IPv6, ponieważ wygląda na to, że wszystko zmierza w tę stronę.

Technologia WAN może obejmować wszystko, począwszy od modemów wdzwanianych (które są już na szczęście rzadkością) po T1, DS3, SONET, MPLS itd. Przyjrzymy się wielu tematom z tej dziedziny, ale nie będziemy się w nie zbytnio zagłębiać, gdyż każdy z nich to właściwie materiał na osobną książkę. Niektóre z nich z pewnością znajdziesz już we własnej kolekcji publikacji wydawnictwa Helion.

Jak już wspomniałem, źródłem wszystkich przykładów zawartych w tej książce są prawdziwe sytuacje, z którymi w większości musiałem zmierzyć się osobiście w trakcie mojej pracy na stanowiskach inżyniera sieciowego, konsultanta, menadżera oraz dyrektora. Prowadzę również własną firmę i jak dotąd miałem przyjemność pracować z jednymi z najlepszych w tej branży. Koncepcje przedstawione w poniższych rozdziałach zostały opracowane lub pozyskane przeze mnie i mój zespół podczas rozwiązywania napotykanych w praktyce problemów.

Podczas przygotowywania drugiego wydania niniejszej książki musiałem podjąć bardzo trudną decyzję. Czy opisywać komendy konfiguracji systemu CatOS, czy też pominąć je na rzecz

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

nowszych przykładów dla NX-OS? Wybór nie był łatwy nie tylko dlatego, że zamieszczenie przykładów z CatOS w pierwszym wydaniu spotkało się z pochwałami ze strony czytelników, ale również z tego powodu, iż w momencie, kiedy piszę tę przedmowę (na początku 2011 r.), przełączniki CatOS są nadal wykorzystywane w sieciach dużych przedsiębiorstw oraz sieciach e-commerce. Z tych względów zdecydowałem się pozostawić opisy konfiguracji CatOS i po prostu dodać przykłady dla NX-OS.

Odzew ze strony czytelników pierwszego wydania spowodował, że postanowiłem dodać do tej książki kilka tematów. Są to: Cisco Nexus, sieci bezprzewodowe, MPLS, IPv6 oraz VoIP. Niektóre z nich są opisane dość szczegółowo, podczas gdy inne, jak MPLS, zostały omówione pobieżnie z powodów, które podaję bezpośrednio w odpowiednich rozdziałach. Tematy takie jak Nexus oraz VoIP zostały opisane szerzej, przez co objętość tej i tak obszernej książki znacznie się zwiększyła. Usunąłem rozdziały dotyczące rozkładu obciążenia serwerów (ang. *server load balancing*), ponieważ nie byłem z nich zadowolony oraz nie udało mi się zorganizować modułu lub urządzenia ACE do aktualizacji przykładów.

Jeśli chodzi o przykłady, to zaktualizowałem je tak, aby odzwierciedlić w stosownych rozdziałach zmiany, jakie dokonały się w sferze sprzętu. Tam, gdzie w pierwszej edycji używałem przełączników 3550, teraz zastosowałem 3750, a firewalle PIX zostały zastąpione urządzeniami ASA. Dodałem również przykłady oparte na przełącznikach Cisco Nexus w tych rozdziałach, w których uznałem to za uzasadnione. Stąd też wiele rozdziałów zawiera jednocześnie przykłady z CatOS, IOS oraz NX-OS. Korzystaj z tego, gdyż w trzeciej edycji CatOS z pewnością zostanie usunięty.

Dlaczego warto przeczytać tę książkę

Niniejsza książka skierowana jest do osób posiadających wiedzę o sieciach informatycznych na pierwszym poziomie certyfikacji. Każdy, kto może poszczycić się wiedzą z zakresu certyfikatu CCNA (lub certyfikatu równoważnego tudzież certyfikatu wyższego poziomu), również skorzysta na przeczytaniu tej książki. Celem, jaki przyświecał mi podczas pisania *Wojownika sieci*, było objaśnienie złożonych zagadnień w łatwy do zrozumienia sposób. Chociaż książka ta jest zasadniczo wprowadzeniem do wielu tematów, możesz również potraktować ją jako zbiór wskazówek do wykonywania typowych zadań związanych z zawartymi w niej zagadnieniami. W duszy jestem nauczycielem, a ta publikacja pozwala mi dotrzeć ze swoimi naukami do większej liczby osób, niż mógłbym sobie wyobrazić. Mam nadzieję, że uznasz moje rozważania za pouczające i przyjemne w lekturze.

Przez lata obserwacji zauważyłem, że wiele osób z branży komputerowej, sieciowej czy telekomunikacyjnej jest często źle poinformowanych w temacie podstawowych pojęć z zakresu tych dziedzin. Wierzę, że w większości przypadków jest to rezultat niskiego poziomu szkoleń oraz korzystania z materiałów, które nie opisują złożonych zagadnień w odpowiedni sposób. Mam nadzieję udowodnić w tej książce, jak proste mogą być niektóre z nich. Oczywiście, jak zwykłem mawiać, „jest to proste, jeśli wiesz, jak to zrobić", więc bardzo się starałem, żeby każdy, kto przeczyta moją książkę, był w stanie zrozumieć koncepcje w niej zawarte.

Jeśli to czytasz, to zgaduję, że chcesz poszerzyć swoją wiedzę o sieciach. Ja również! Nauka powinna być niekończącą się przygodą i jestem zaszczycony, że pozwalasz mi być kompanem w Twojej podróży. Studiuję i uczę się na temat komputerów, sieci i telekomunikacji już od 29 lat, ale moja podróż wciąż nie dobiegła końca.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Niniejsza książka nie zawiera objaśnienia budowy modelu odniesienia ISO/OSI, ale pokrótce tłumaczy różnice pomiędzy hubami, przełącznikami oraz routerami. Powinieneś posiadać podstawowe pojęcie na temat tego, co oznacza termin warstwa druga i jakie jest jej odniesienie do modelu ISO/OSI. Poza tym książka ta stara się opisać wszystko, ale w sposób odmienny niż pozostałe publikacje.

W tej książce staram się przekazać Ci wiedzę, która będzie Ci potrzebna w prawdziwym świecie. Dlaczego wybrać przełącznik warstwy trzeciej zamiast przełącznika warstwy drugiej? Jak stwierdzić, czy Twoja sieć zachowuje się właściwie? Jak sobie radzić z burzą rozgłoszeniową? Jakie są objawy burzy rozgłoszeniowej? Po czym poznać, że w sieci powstała pętla Spanning Tree, i jak się jej pozbyć? Czym jest T1 albo DS3? Jak one działają? W tej książce znajdziesz odpowiedzi na wszystkie te pytania oraz na wiele innych. Starałem się zgromadzić w niej informacje, które bywają źle interpretowane przez inżynierów sieciowych bez ich własnej winy. *Wojownik sieci* zawiera przykłady konfiguracji inspirowane prawdziwymi sytuacjami oraz strukturami istniejących sieci, a przy tym dostarcza licznych anegdot z mojej pracy w terenie — mam nadzieję, że będziesz się przy nich dobrze bawił.

Konwencje stosowane w tej książce

W tej książce zastosowane zostały następujące konwencje typograficzne:

Czcionka pogrubiona
: Stosowana przy wprowadzaniu nowych terminów dla podkreślenia istoty informacji.

Czcionka pochylona
: Formatowanie stosowane dla linków URL.

`Czcionka o stałej szerokości znaków`
: Styl stosowany do zapisu komend, fragmentów kodu i listingów z urządzeń (przedstawionych w takiej formie, w jakiej oryginalnie pojawiają się na ekranie).

`Pochylona czcionka o stałej szerokości znaków`
: Styl stosowany w poleceniach do zaznaczenia argumentów, pod które należy samodzielnie podstawić określone wartości.

`Pogrubiona czcionka o stałej szerokości znaków`
: Styl stosowany do wyróżniania istotnych fragmentów listingów.

Podpowiedź, sugestia lub wskazówka ogólna.

Ostrzeżenie lub sugestia zachowania ostrożności.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Podziękowania

Pisanie książki to ciężka praca — o wiele trudniejsza, niż mi się wydawało. Choć spędziłem samotnie niezliczone godziny nad klawiaturą komputera, nie doprowadziłbym tego przedsięwzięcia do końca bez pomocy wielu osób.

Chcę podziękować mojej pięknej żonie Lauren za cierpliwość, miłość i wsparcie. Lauren, będąc moim domowym korektorem, stanowiła również pierwszą linię obrony w gąszczu zasad gramatyki. Wiele rozdziałów z pewnością zanudziło ją na śmierć, ale wiem, że podobało jej się przynajmniej kilka. Dziękuję, że pomogłaś mi osiągnąć ten cel w życiu.

Chcę podziękować Meghan i Colleen za to, że starały się zrozumieć, iż w trakcie pisania nie mogłem się z nimi bawić. Mam nadzieję, że kończąc tę książkę, pomogłem zaszczepić w Was cechę, jaką jest wytrwałość. Jeśli nie, możecie być pewne, że będę przywoływał to jako przykład przez resztę Waszego życia. Kocham Was obie uczuciem „większym niż Wszechświat".

Chcę podziękować mojej matce — bo jest moją mamą i za to, że nigdy we mnie nie zwątpiła, nie przestała we mnie wierzyć i zawsze mi pomagała, nawet kiedy nie powinna (pozdrawiam Cię, Mamo!).

Chcę podziękować mojemu ojcu za to, że nie pobłażał mi, kiedy nie należało, że nauczył mnie myśleć logicznie i doceniać piękno, które tkwi w szczegółach. Pamiętam, jak siadywaliśmy we dwóch przed komputerem RadioShack Model III i wpisywaliśmy proste programy z czasopism. Doszedłem tu, gdzie jestem, głównie pod Twoim wpływem, dzięki Twoim wskazówkom i naukom. Uczyniłeś ze mnie mężczyznę, którym jestem dzisiaj. Dziękuję, Tato. Tęsknię za Tobą.

Chcę podziękować mojej wiernej suce nowofunlandce o imieniu Cozy, która tragicznie musiała zostać uśpiona i odeszła w moich ramionach, aby nie cierpieć dłużej bólu powodowanego rakiem. Choroba dopadła ją, kiedy pracowałem nad pierwszym wydaniem tej książki. Gdyby nie pamięć o niej, prawdopodobnie nie publikowałbym dzisiaj. Jej śmierć wywołała we mnie wielki smutek, od którego uciekałem w pisanie. Brakuje mi Ciebie, Cozy — biegaj swobodnie bez bólu po moście z tęczy, aż spotkamy się ponownie.

Chcę podziękować Mattowi Maslowskiemu za to, że pozwolił mi korzystać ze sprzętu w swoim laboratorium, jeśli brakowało czegoś w moim własnym, oraz za pomoc przy zagadnieniach związanych z urządzeniami Cisco, kiedy sam nie byłem pewien właściwego rozwiązania. Nikomu nie zaufałbym bardziej w kwestiach sieciowych. Dzięki, stary.

Chcę podziękować Jeffowi Fry'owi, CCIE# 22061, za zapewnienie mi tymczasowego dostępu do pary nieskonfigurowanych przełączników Cisco Nexus 7000. To było dla mnie istotne i dzięki temu drugie wydanie jest bardziej kompletne.

Chcę podziękować Jeffowi Cartwrightowi, który zaoferował mi moją pierwszą ekscytującą posadę u dostawcy usług internetowych i nauczył mnie prawie wszystkiego, co wiem na temat telekomunikacji. Wciąż pamiętam, jak uczył mnie o zagęszczeniu jedynek, prowadząc samochód drogą międzystanową 80 i gryzmoląc jednocześnie na kolanie wykresy, kiedy ja starałem się nie wyglądać na mocno przestraszonego. Dzięki za korektę rozdziałów poświęconych zagadnieniom telekomunikacyjnym. Nikomu innemu nie zaufałbym bardziej w tej materii.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Chcę podziękować Mike'owi Stevensowi za pomoc w pracy nad czytelnością materiału oraz za jedne z najbardziej kolorowych wspomnień, jakie zamieściłem w tej książce. Jego pomoc z firewallami PIX była znacząca dla ukończenia pierwszej edycji. Powinieneś być wdzięczny, że nie zamieściłem żadnych zdjęć. Wciąż mam pewną fotografię z centrum danych Secaucus…

Chcę podziękować Peterowi Martinowi za pomoc w laboratorium w kilku tematach, w których nie miałem doświadczenia. Dodatkowe podziękowania za bycie jednym z technicznych recenzentów książki *Wojownik sieci* — Twoje komentarze były zawsze celne, a Twój wysiłek uczynił tę książkę jeszcze lepszą.

Chcę podziękować kolejnemu recenzentowi technicznemu, Yvesowi Eynardowi. Wyłapałeś kilka powalających błędów i doceniam czas, który na to poświęciłeś. Ta książka jest jeszcze lepsza dzięki Twoim wysiłkom.

Chcę podziękować Salowi Conde'owi oraz Edowi Homowi za udostępnienie przełącznika 6509E wraz z modułami.

Chcę podziękować Michaelowi Heubergerowi, Helge'owi Brummerowi, Andy'emu Vassaturo, Kelly Huffman, Glennowi Bradleyowi, Billowi Turnerowi oraz reszcie zespołu w Północnej Karolinie za danie mi szansy pracy z platformą Nexus 5000 oraz za wysłuchiwanie moich nieustannych nawiązań do tej książki w codziennych rozmowach. Mogę sobie wyobrazić, że nie ma nic gorszego niż przebywanie lub pracowanie ze wszystkowiedzącym autorem książek.

Chcę podziękować Christopherowi Leongowi za techniczne recenzje rozdziałów poświęconych zagadnieniom telekomunikacji oraz VoIP.

Chcę podziękować Robertowi Schafferowi za pomoc w przypomnieniu mi zagadnień, nad którymi kiedyś pracowaliśmy, a o których dawno zapomniałem.

Chcę podziękować Jennifer Frankie za pomoc w skontaktowaniu się z ludźmi, których sam nie byłbym stanie odszukać, oraz za pomoc w znalezieniu informacji, do których bym inaczej nie dotarł.

Chcę podziękować Mike'owi Loukidesowi, mojemu wydawcy, za to, że nie ograniczał mi swobody, nie zwątpił we mnie, i za to, że w ogóle dał mi szansę. Pomogłeś mi stać się lepszym pisarzem. Nie wiem, jak Ci się odwdzięczę.

Chcę podziękować Rachel Head, mojej adiustatorce, która uczyniła pierwszą edycję tej książki bardziej czytelną.

Chcę podziękować wszystkim wspaniałym ludziom z wydawnictwa O'Reilly. Pisanie tej książki było świetnym doświadczeniem, w dużej mierze dzięki ludziom, z którymi pracowałem w O'Reilly.

Chcę podziękować mojemu dobremu przyjacielowi Johnowi Tocado, który powiedział mi kiedyś: „Jeśli chcesz pisać, to pisz!". Ta książka jest dowodem na to, że można odmienić czyjeś życie jednym zdaniem. Możesz się spierać, że sam zmieniłem swoje życie, i nie będę się kłócił, ale się mylisz. Bywały chwile, kiedy byłem przytłoczony ogromem pracy, jaki pozostawał do ukończenia książki. Poważnie zastanawiałem się wtedy, czy nie zrezygnować. To Twoje słowa sprawiły, że tego nie zrobiłem. Dziękuję.

Nie wiem nawet, jak podziękować wszystkim pozostałym, którzy dodawali mi zachęty. Życie i praca z pisarzem muszą być czasami irytujące. Gdy goniły mnie terminy, z pewnością bywałem humorzasty, denerwujący i frustrujący, za co przepraszam.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przez ostatnie kilka lat moim celem było ukończenie tej książki. Wszystkie inne obowiązki, z wyjątkiem dbania o zdrowie i rodzinę, odłożyłem na bok. Zdaję sobie sprawę, że publikacja tej książki jest spełnieniem mojego marzenia. Jeśli też posiadasz własne marzenia, mogę dać Ci jedną radę: musisz dążyć do ich spełnienia, a tak się stanie. Naprawdę wystarczy tylko tyle.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 1.

Co to jest sieć?

Zanim zaczniemy, chciałbym zdefiniować niektóre pojęcia i ustanowić kilka podstawowych zasad. Dla celów tej książki (mam nadzieję, że przyda Ci się to również w życiu zawodowym) **sieć komputerową** możemy określić jako „co najmniej dwa komputery połączone ze sobą w taki sposób, że istnieje możliwość współdzielenia informacji między nimi". Nie szukaj tej definicji w publikacjach RFC, gdyż stworzyłem ją w tej chwili. Dla naszych potrzeb nadaje się jednak doskonale.

Istnieje wiele rodzajów sieci komputerowych: sieć lokalna — LAN (ang. *local area network*), sieć rozległa — WAN (ang. *wide area network*), sieć miejska — MAN (ang. *metropolitan area network*), sieć uczelniana — CAN (ang. *campus area network*), sieć Ethernet, sieć Token Ring, sieć światłowodowa FDDI (ang. *fiber distributed data interface*), sieć oparta na szerokopasmowej technologii ATM (ang. *asynchronous transfer mode*), sieć pakietowa Frame Relay, sieć w standardzie T1, sieć DS3[1], sieć mostkowana (ang. *bridged network*), sieć rutowalna (ang. *routed network*), sieć typu punkt-punkt (ang. *point-to-point network*) i wiele innych. Starsi czytelnicy mogą pamiętać jeszcze program Laplink, który umożliwiał kopiowanie plików pomiędzy dwoma komputerami połączonymi ze sobą specjalnym kablem równoległym. Taką strukturę również mogliśmy nazwać siecią, nie była ona jednak zbyt skalowalna (tylko dwa komputery) ani dość szybka. Mimo to umożliwiała transmisję danych z jednego komputera do drugiego dzięki istnieniu połączenia.

Połączenie jest w tym przypadku istotnym czynnikiem. Jego obecność odróżnia prawdziwą sieć od czegoś, co nazywany ***sneakernet***, czyli przenoszenia plików za pomocą fizycznych nośników wymiennych. Jeśli podłączasz napęd USB (czy ktoś jeszcze korzysta z dyskietek?) do komputera, nie istnieje żadne wskazanie mówiące, że dane, które kopiujesz, pochodzą z innego komputera. Nie istnieje żadne połączenie między komputerami. Fizyczne połączenie wymaga zastosowania pewnego rodzaju adresacji lub identyfikacji węzłów w sieci (nawet jeśli jest to identyfikacja oparta na stopniu podrzędności typu „master/slave" lub „primary/secondary").

Urządzenia w sieci są zazwyczaj połączone fizycznie za pomocą kabli. Trochę inaczej jest oczywiście w przypadku sieci bezprzewodowej, gdzie kabel zastąpiony jest łączem radiowym. Każdy węzeł w sieci bezprzewodowej posiada swój adres, a transportowane ramki mają określone źródło i przeznaczenie, jak w innych typach sieci.

[1] T1 i DS3 to standardy amerykańskie, ich europejskimi odpowiednikami są E1 i E3 — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Sieci często rozróżnia się według zasięgu, jaki obejmują, czego najprostszym przykładem są LAN, MAN, WAN i CAN. Sieci LAN najczęściej obejmują jeden budynek lub choćby piętro budynku. WAN pokrywa dużo większe obszary i z reguły łączy ze sobą poszczególne sieci lokalne. W praktyce sieć WAN obejmuje całą kulę ziemską, a nie ma przeciwwskazań, by posunąć się nawet krok dalej. Sieć MAN możemy spotkać w miejscach, gdzie możliwe jest zastosowanie punktów wymiany ruchu — IXP (ang. *Internet eXchange Point*) — łączących ze sobą sieci lokalne na obszarze miasta lub regionu. Sieć CAN jest zbliżona do MAN, z tym ograniczeniem, że obejmuje jedynie obszar danego **kampusu** (kampus jest z reguły definiowany jako grupa budynków pod zarządem jednego podmiotu, takiego jak uczelnia czy przedsiębiorstwo).

Można by się spierać, czy określenia MAN i CAN nie powinny być stosowane wymiennie, co w niektórych przypadkach wydaje się słusznym podejściem. Są też jednak tacy, którzy nie zgodzą się z tym twierdzeniem, gdyż uważają, że muszą zaistnieć szczególne okoliczności umożliwiające nazwanie danej sieci mianem sieci uczelnianej, a stosowanie w takich sytuacjach odmiennego terminu byłoby nieprawidłowe. Różnica podejścia może wynikać stąd, że w przypadku sieci uczelnianej jest większe prawdopodobieństwo, iż wszystkie budynki uda się połączyć z wykorzystaniem własnej sieci. Prowadzenie przez dany podmiot własnej sieci w całym mieście jest praktycznie niewykonalne. W środowiskach miejskich to zazwyczaj dostawcy usług telekomunikacyjnych zapewniają technologię umożliwiającą połączenie za pomocą ich struktur sieciowych.

Sieci MAN i CAN właściwie mogą być również sieciami WAN. Różnica tkwi często w semantyce. Jeśli dwa budynki uczelni są połączone za pomocą sieci Frame Relay, to są częścią sieci WAN czy CAN? A co jeśli Frame Relay jest elementem infrastruktury kampusowej, a nie jest zapewniane przez dostawcę usług telekomunikacyjnych? Czy jest to jakaś różnica? A jeśli uczelnia znajduje się w obrębie aglomeracji miejskiej, to czy sieć uczelniana może być nazywana siecią MAN?

Zazwyczaj projektanci sieci nadają jej nazwę, która przykleja się do niej na stałe. Jeśli grupa konsultantów zaprojektuje sieć WAN, ale nazwie ją w dokumentacji siecią MAN, to podmiot, dla którego została ona stworzona, będzie właśnie w ten sposób ją nazywał przez cały czas jej funkcjonowania.

Wyobraź sobie sytuację, w której sieci lokalne są połączone z siecią uczelnianą, a sieci uczelniane połączone z siecią rozległą. Sam widzisz, jak trudne do sklasyfikowania mogą być takie struktury, szczególnie dla niewtajemniczonych.

W tej branży problemem jest częste szafowanie pojęciami oraz błędne stosowanie definicji. Dodatkowo, jak w tym przypadku, definicje mogą być nieco niejasne, co prowadzi do nieporozumień.

Powinieneś zawsze być ostrożny w kwestii terminologii, którą stosujesz. Zdarza się, że dyrektor działu informatyki określa daną sieć mianem WAN, a inżynierowie nazywają ją CAN. Musisz wtedy wyprowadzić z błędu tego, kto się myli, lub stosować w komunikacji z każdą ze stron jej własny język i nazewnictwo. Takie sytuacje są niestety częstsze, niż mógłbyś przypuszczać. Unikaj arbitralnych rozstrzygnięć, jeśli chodzi o wyodrębnianie struktur jednej z tych sieci (MAN, WAN, CAN) względem drugiej. W przypadkach pozostałych sieci definicje są już bardziej dokładne.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Dla naszych potrzeb zdefiniujemy poniżej następujące rodzaje sieci:

LAN

Sieć lokalna, ograniczająca się do zamkniętej przestrzeni, jaką stanowi budynek lub piętro budynku. Stosuje się tutaj technologie o krótkim zasięgu, takie jak Ethernet, Token Ring i tym podobne. LAN pozostaje pod kontrolą firmy lub podmiotu, który z niej korzysta.

WAN

Sieć rozległa, która służy do łączenia ze sobą sieci lokalnych przez zewnętrznego dostawcę usług internetowych. Przykładem może być chmura Frame Relay (dostarczona przez operatora telekomunikacyjnego) łącząca biura korporacji w Nowym Jorku, Bostonie, Los Angeles i San Antonio.

CAN

Sieć uczelniana łącząca ze sobą sieci lokalne i/lub budynki leżące na wydzielonym obszarze należącym do jednego podmiotu i pozostającym pod jego kontrolą. Taka sytuacja umożliwia położenie podziemnego okablowania światłowodowego, które łączy poszczególne budynki. Przykładem mogą być kampusy uczelniane lub parki przemysłowe.

MAN

Sieć miejska, która łączy ze sobą sieci lokalne i/lub budynki leżące na obszarze większym niż kampus uczelniany. Dla przykładu MAN może łączyć różne biura jednej firmy w obszarze aglomeracji miejskiej poprzez usługi operatora telekomunikacyjnego. Również w tym przypadku trzeba być ostrożnym, jeśli chodzi o definicje. Wiele firm na Manhattanie posiada budynki lub centra danych znajdujące się na przeciwległym brzegu rzeki, w New Jersey. Są one traktowane jako część aglomeracji Nowego Jorku, więc należą do tej samej sieci miejskiej, mimo że znajdują się w innym stanie.

Terminologia i język, jakimi się posługujesz, są podobne do każdego protokołu: uważaj na terminy, jakie stosujesz na co dzień, ale nie bądź zbyt pedantyczny. Bezustanne upominanie innych jest denerwujące. Słuchaj ludzi dookoła i pomagaj ich edukować. Chęć do dzielenia się wiedzą to czynnik, który pozwala odróżnić dobrego pracownika działu IT od przeciętnego.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 2.

Huby i switche

Huby

W początkach Ethernetu stosowano okablowanie w standardzie 10Base-5, zwanym **grubym Ethernetem** (ang. *thick-net*). Posługiwano się kablem koncentrycznym o dużej średnicy, co mogło powodować pewne problemy w trakcie instalacji. Jego następca, 10Base-2, wykorzystywał już cieńszy kabel koncentryczny podobny do tych używanych w sieciach telewizji kablowej. Z uwagi na różnicę w grubości stosowanego kabla określano go z kolei jako **cienki Ethernet** (ang. *thin-net*). Technologie te wymagały zastosowania dużych metalowych konektorów do łączenia kabli. W przypadku 10Base-5 były to konektory o złączach N, a w przypadku 10Base-2 konektory o złączach BNC. W obydwu standardach konieczna była również instalacja specjalnych terminatorów na końcach tras kablowych. Usunięcie konektorów lub terminatorów powodowało przerwanie działania całej sieci. Kable te stanowiły fizyczny szkielet sieci Ethernet.

Wraz z rozwojem Ethernetu zaczęto stosować nieekranowaną skrętkę (UTP) zakończoną złączami RJ-45. W tym okresie huby stały się nowymi szkieletami sieci w większości instalacji. Wiele firm zaczęło dodatkowo instalować te urządzenia w sieciach 10Base-2, aby zwiększyć ich elastyczność. W założeniu huby miały łączyć struktury UTP i 10Base-2 BNC. Z uwagi jednak na zalety skrętki UTP, które okazały się szczególnie istotne w procesie instalacji okablowania, to ona stała się w niedługim czasie standardem.

Hub, zwany też koncentratorem, jest prostym rozgałęźnikiem umożliwiającym przekazywanie sygnału do wszystkich podłączonych do niego kabli ethernetowych. Z tego powodu może być również zwany **regeneratorem** (ang. *repeater*). Ważne jednak, by mieć świadomość, że o ile każdy hub jest regeneratorem, o tyle nie każdy regenerator musi być hubem.

Regenerator służy do replikowania sygnału. Umożliwia on zwiększenie zasięgu transmisji w przypadku przekroczenia dopuszczalnej odległości od zdalnego hosta lub pomiędzy grupą użytkowników sieci w standardzie 10Base-T. Innymi słowy, jeśli planowane połączenie przekracza dopuszczalną odległość, to możemy ją zwiększyć, umieszczając na trasie sygnału regenerator.

Segmenty sieci rozdzielone są regeneratorami lub hubami. Rysunek 2.1 przedstawia regenerator umożliwiający zwiększenie odległości pomiędzy serwerem a komputerem osobistym.

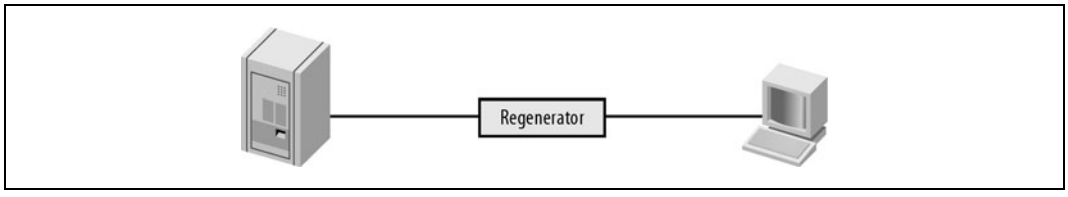
Hub działa jak regenerator, z tą różnicą, że regenerator posiada tylko dwa złącza, natomiast koncentrator może mieć ich więcej. Oznacza to, że hub regeneruje sygnał na wiele kabli, a nie tylko na jeden. Rysunek 2.2 pokazuje hub podłączony do kilku komputerów w sieci.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

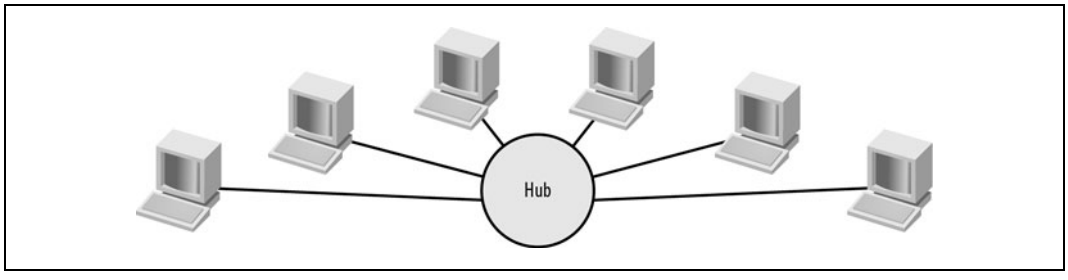
Byłem zaskoczony tym, że nie ma wyraźnego ograniczenia zasięgu sieci w specyfikacji 10Base-T, podczas gdy standardy 10Base-5 oraz 10Base-2 posiadają ograniczenia odległości odpowiednio do 500 m i 200 m. 10Base-T opisuje jedynie charakterystykę, jaką powinien posiadać kabel transmisyjny.

Specyfikacja kabli kategorii 5e (TIA/EIA-568-B.2-2001) określa wartości, bazując na kablu o długości 100 m. Mówiąc ściślej, oznacza to, że kabel musi spełniać te warunki do długości 100 m. Jedną rzeczą jest powiedzenie, że „krzywa opóźnienia propagacji nie może być większa niż 45 ns na 100 m", a czym innym, że „kabel nie może przekraczać długości 100 m".

Odkładając jednak na bok semantykę, należy stwierdzić, że zawsze warto stosować zasadę, według której długość kabla Cat-5e nie powinna przekraczać 100 m.



Rysunek 2.1. Regenerator przedłużający dystans pojedynczego łącza w standardzie 10Base-T



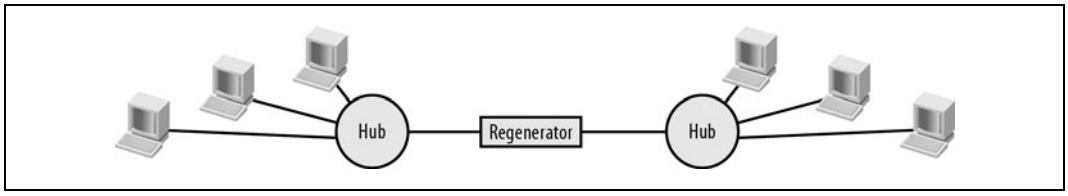
Rysunek 2.2. Hub łączący kilka stacji roboczych z siecią

W strukturach sieci Ethernet regeneratory i koncentratory traktowane są w ten sam sposób. Reguła standardu, zwana zasadą 5-4-3, stwierdza, że pomiędzy dwoma dowolnymi hostami sieci Ethernet może być maksymalnie 5 segmentów. Segmenty mogą być połączone za pomocą 4 regeneratorów, ale tylko 3 z nich mogą być obsadzone stacjami końcowymi. Choć reguła ta może wydawać się nieco archaiczna w kontekście dzisiejszych technologii sieciowych, stanowiła niegdyś źródło sporych problemów dla tych, którzy jej nie rozumieli.

Od kiedy huby stały się tańsze, dodatkowe węzły były często stosowane jako regeneratory w bardziej złożonych sieciach. Rysunek 2.3 przedstawia przykład tego, jak dwie odległe grupy użytkowników mogą być ze sobą połączone za pomocą dwóch hubów na końcach i regeneratora w środku.

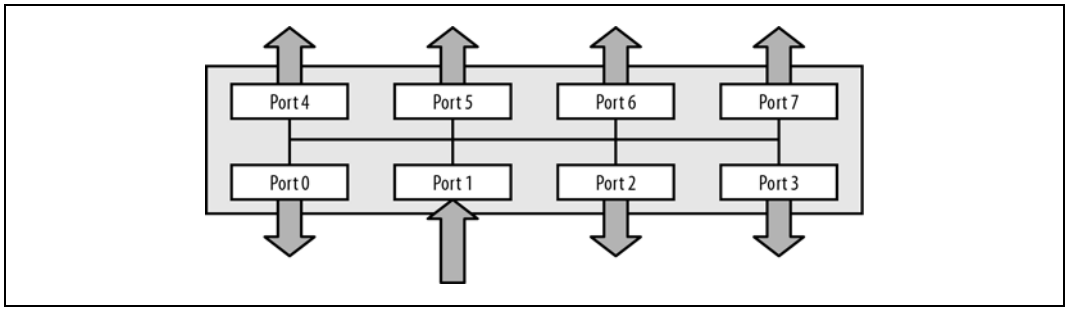
Huby są bardzo prostymi urządzeniami. Każdy sygnał otrzymany na dowolnym porcie regenerowany jest na wszystkie pozostałe porty. Są to najprostsze elementy sieci Ethernet i jako takie są dla niej „przezroczyste" (ich obecność w sieci może być istotna jedynie dla celów zarządzania). Huby nie dokonują procesu analizy czy zmiany ramek, a ich działanie nie jest w żaden sposób oparte na informacjach transportowanych przez ramki.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 2.3. Regenerator łączący huby

Rysunek 2.4 ilustruje działanie koncentratora. Łatwo się domyślić, że ten model operacyjny może stać się problematyczny w rozleglejszych sieciach. Natężenie ruchu może być chwilami tak duże, że sieć będzie przepełniona. Przykładowo kiedy jedna osoba drukuje duży plik, reszta użytkowników sieci będzie doświadczać problemów z transferem danych, w czasie gdy plik ten jest przesyłany do drukarki.



Rysunek 2.4. Hub przekazuje sygnały na wszystkie porty, bez względu na to, gdzie zostały skierowane

Jeśli jedno z urządzeń wykonuje transmisję po danym kablu, to kolejne urządzenie będzie musiało chwilę poczekać, a potem spróbować ponowić transmisję. Jeśli dwie stacje transmitują pakiety w tym samym momencie, to wystąpi **kolizja** (ang. *collision*). Obie stacje odnotowują tę kolizję, wycofują się i próbują ponowić transmisję. W sieciach, w których występuje duże natężenie ruchu pakietów, z pewnością wystąpi wiele tego typu kolizji.

Dzięki zastosowaniu huba więcej stacji jest w stanie korzystać z tej samej sieci w danym momencie. Jeżeli wszystkie stacje będą aktywne, to sieć będzie znacząco zwalniać z powodu występowania nadmiernej liczby kolizji.

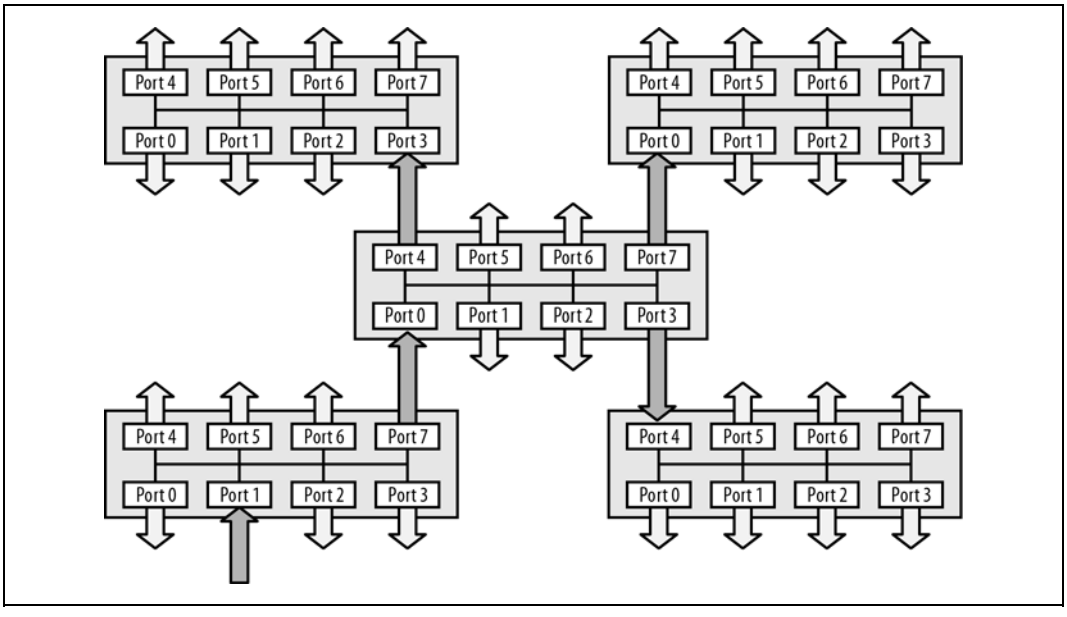
Kolizje ograniczone są do segmentów sieci. Segment sieci Ethernet to część sieci, w której urządzenia mogą komunikować się, używając adresów MAC warstwy drugiej modelu ISO/OSI. Aby umożliwić komunikację z urządzeniami poza segmentem sieci Ethernet, niezbędne jest dodatkowe urządzenie, takie jak router. Kolizje są również ograniczone do **domen kolizyjnych** (ang. *collision domains*). Domena kolizyjna to obszar sieci Ethernet, w którym może dojść do wystąpienia kolizji. Jeśli jedna ze stacji może uniemożliwić innej stacji wysłanie pakietu, gdyż zajmuje w tym momencie sieć, to obie te stacje znajdują się w tej samej domenie kolizyjnej.

Domena rozgłoszeniowa (ang. *broadcast domain*) to segment sieci Ethernet, w którym pakiety mogą być transmitowane bezpośrednio pomiędzy stacjami. Pakiety rozgłoszeniowe pozostają w obrębie warstwy trzeciej sieci (chyba że zostaną przekazane dalej), która jest zazwyczaj ograniczona urządzeniem warstwy trzeciej, takim jak router. Pakiety te mogą być przekazywane przez switche (urządzenia warstwy drugiej), ale zatrzymują się na routerach.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Wiele osób błędnie uważa, że pakiety broadcast nie rozchodzą się poza struktury ograniczane przez przełączniki lub wirtualne sieci lokalne (VLAN). Oczywiście tak być powinno w prawidłowo zaprojektowanych i skonfigurowanych sieciach, nie jest to jednak zasadą. Przykładowo łączymy ze sobą dwa switche kablem skrosowanym, zakładając, że jeden z nich ma skonfigurowany na wszystkich portach VLAN 10, a drugi na wszystkich portach VLAN 20. Stacje podłączone do jednego switcha będą mogły komunikować się ze stacjami podłączonymi do drugiego, o ile są w tej samej sieci IP. Broadcasty i sieci IP nie są ograniczone do VLAN-ów, choć na pozór mogłoby się tak wydawać.

Rysunek 2.5 pokazuje sieć hubów połączonych za pośrednictwem koncentratora centralnego. Kiedy ramka jest transmitowana na port 1 huba znajdującego się na dole po lewej stronie, jest przekazywana na wszystkie pozostałe porty tego huba, również na ten, który połączony jest z koncentratorem centralnym. Następnie hub centralny przekazuje ramkę na wszystkie swoje porty, propagując ją do pozostałych koncentratorów w sieci. Schemat ten przedstawia ideę szkieletu sieci, w której wszystkie urządzenia otrzymują każdą ramkę przesłaną w tej sieci.

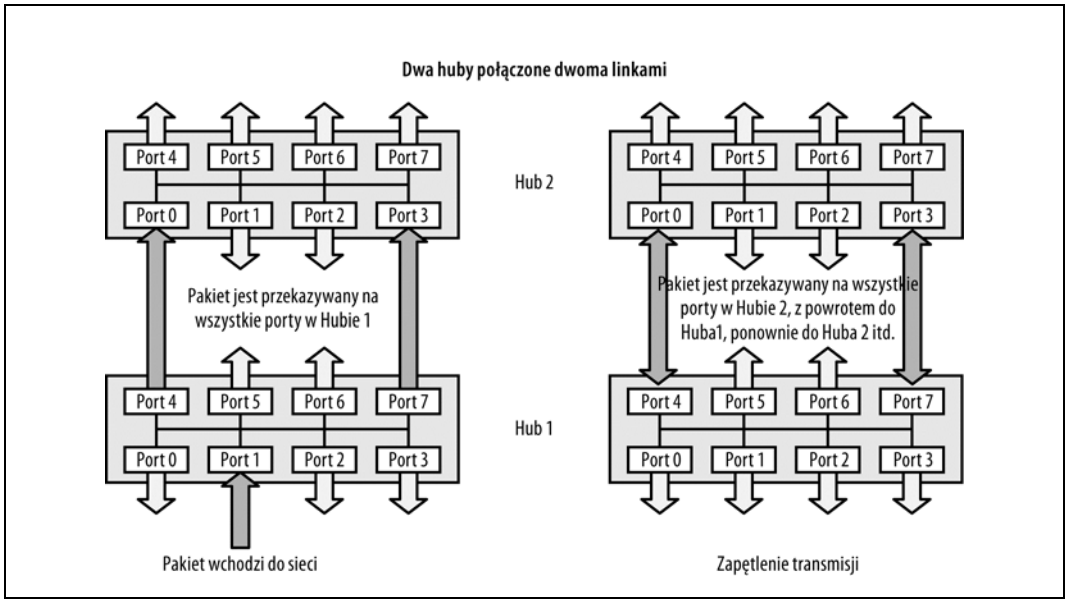


Rysunek 2.5. Sieć oparta na hubach

W większych sieciach tego typu mogą pojawić się jeszcze inne problemy. Będą to tak zwane **późne kolizje** (ang. *late collisions*). Późna kolizja występuje, kiedy dwie stacje wykonają w tym samym czasie pozytywnie zakończony test na brak zajętości sieci i rozpoczną transmisję, która kończy się jednak kolizją. Taka kolizja może wystąpić, jeśli sieć jest na tyle rozległa, że czas propagacji transmitowanej ramki z jednego końca sieci na drugi jest dłuższy niż czas potrzebny na wykonanie testu na brak zajętości sieci.

Kolejnym niepożądanym zjawiskiem przy korzystaniu z koncentratorów jest **burza rozgłoszeniowa** (ang. *broadcast storm*). Rysunek 2.6 przedstawia dwa huby, które zostały ze sobą połączone dwoma linkami. Kiedy ramka transmitowana jest do huba 1, jest przekazywana na każdy jego

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 2.6. Burza broadcastowa

port, w tym oba porty, które połączone są z hubem 2. Hub 2 powiela ramkę na wszystkie swoje porty, przekazując ją w ten sposób z powrotem w kierunku huba 1. Ten zaś ponownie zwraca sygnał w kierunku huba 2. W efekcie następuje zapętlenie transmisji.

Każdy, kto choć raz doświadczył na własnej skórze burzy broadcastowej, wie, jaka to „frajda", oczywiście pod warunkiem, że uznamy wydzierającego się na nas szefa za element zabawy. Robi się jeszcze ciekawej, kiedy dołączy do Was szef szefa. Objawem burzy broadcastowej jest brak możliwości wysyłania danych przez jakiekolwiek urządzenie z uwagi na permanentny ruch odbywający się w sieci. Diody koncentratora świecą jednostajnie, zamiast migać, a kierownictwo grozi Ci uszkodzeniem ciała.

Jedynym sposobem na wybrnięcie z tej trudnej sytuacji jest przerwanie pętli transmisji. W tym celu powinno wystarczyć wyłączenie i ponowne uruchomienie urządzeń sieciowych. Z uwagi na to, że koncentratory nie są z reguły urządzeniami zarządzalnymi, pewnym wyzwaniem może okazać się odnalezienie w warstwie drugiej urządzeń odpowiedzialnych za powstanie pętli.

Koncentratory mają, jak widać, wiele wad i są obecnie rzadko stosowane w nowoczesnych sieciach. Zostały już dość dawno temu zastąpione przez **switche**, które oferują większą szybkość, automatyczne wykrywanie pętli oraz wiele innych dodatkowych funkcji.

Switche

Kolejnym etapem w rozwoju sieci ethernetowych były switche, zwane też przełącznikami. Switche różnią się od hubów tym, że odgrywają aktywną rolę w transmisji ramek. Koncentrator po prostu powiela każdy sygnał otrzymany na jednym z portów na pozostałe porty. Switch natomiast zapamiętuje, do których portów podpięte są określone urządzenia, i przesyła ramki tylko do tych portów, do których zostały one skierowane.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

To, co nazywamy **pakietem** (ang. *packet*) w protokole TCP/IP, jest określane jako **ramka** (ang. *frame*), kiedy mówimy o koncentratorach, mostach i przełącznikach. Z technicznego punktu widzenia są to jednak dwie różne rzeczy. Dopiero kiedy pakiety TCP zostaną enkapsulowane z informacją warstwy drugiej, tworzą ramkę. Potocznie oba te terminy są stosowane wymiennie, do czego zresztą sam się przyczyniłem. Poprawne jest używanie pojęcia ramka, w przypadku gdy odnosimy się do przełączników lub koncentratorów.

W czasach kiedy większość firm rozwijała technologię przełączników, Cisco skoncentrowało swoją energię na routerach, przez co nie posiadało produktu, dzięki któremu mogłoby konkurować na rynku. Cisco zrobiło więc najmądrzejszą rzecz, jaką wtedy mogło zrobić — kupiło jedną z najlepszych firm zajmujących się produkcją switchy, Kalpana, i dołączyło jej urządzenia do swojej oferty. W rezultacie przełączniki Cisco nie posiadały tego samego systemu operacyjnego co routery Cisco. Podczas gdy w routerach zastosowano Internetwork Operating System (IOS), w switchach zaimplementowany został **CatOS**, którego nazwa pochodzi od nazwy rodziny przełączników — **Catalyst**.

Jeśli chodzi o terminologię, należy wspomnieć, że słowa „przełączanie" i „przełącznik" mają wiele znaczeń nawet w świecie technologii sieciowych. Można wyróżnić przełączniki Ethernet, przełączniki Frame Relay, przełączniki warstwy trzeciej, przełączniki wielowarstwowe itd. Oto kilka z ważniejszych pojęć:

Przełącznik
: Termin stosowany dla wszystkich urządzeń przełączających bez względu na dziedzinę i przedmiot przełączania. W technologii sieciowej przełącznik to zasadniczo switch ethernetowy. W telekomunikacji przełącznik służy do wielu innych zadań, których nie będziemy omawiać w tym rozdziale.

Przełącznik ethernetowy
: Przełącznikiem ethernetowym jest każde urządzenie, które transmituje w sieci Ethernet ramki, opierając się na adresach MAC warstwy drugiej. Podczas gdy hub powiela każdą ramkę na każdy port, switch przekazuje ramki tylko do tych portów, dla których zostały przeznaczone. Switch dzieli sieć na mniejsze domeny kolizyjne na każdym porcie, natomiast koncentrator tworzy domenę kolizyjną obejmującą wszystkie porty.

Przełącznik warstwy trzeciej
: Jest to switch z funkcją rutowania. Ogólnie rzecz biorąc, w przełączniku warstwy trzeciej sieci VLAN mogą być skonfigurowane jako wirtualne interfejsy. Te przełączniki są już dzisiaj rzadkością, gdyż większość switchy jest obecnie wielowarstwowa.

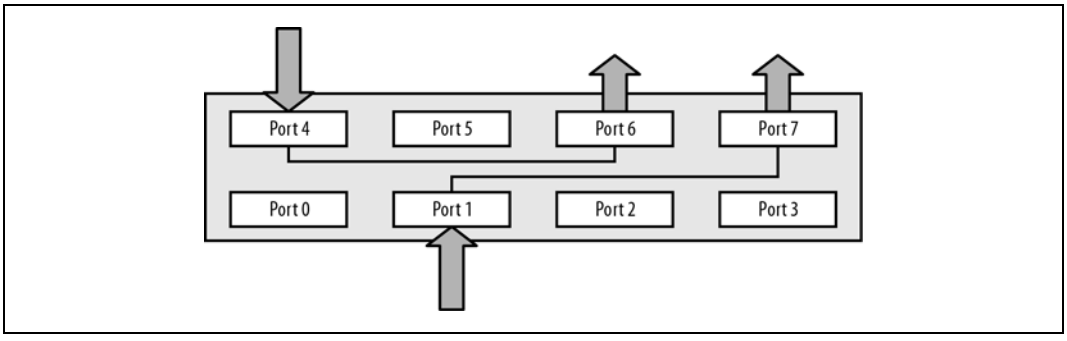
Przełącznik wielowarstwowy
: Podobny do przełącznika warstwy trzeciej. Umożliwia jednak kontrolę opartą na wyższych warstwach modelu ISO/OSI. Oznacza to wykorzystanie w tym celu protokołów TCP, UDP czy szczegółów zawartych w polu danych pakietu.

Przełączanie
: W sieci Ethernet termin ten oznacza przekazywanie ramek na podstawie ich docelowych adresów MAC. W telekomunikacji przełączanie jest czynnością polegającą na nawiązaniu połączenia między dwoma stronami. Z kolei w routingu jest to proces przekazywania pakietów z jednego interfejsu do drugiego w obrębie routera.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Switche różnią się od koncentratorów jedną zasadniczą rzeczą: sygnał, który wchodzi na jeden z portów przełącznika, nie jest przekazywany na wszystkie pozostałe porty (chyba że takie jest przeznaczenie sygnału), tak jak dzieje się to zawsze w koncentratorze. Nowoczesne przełączniki oferują wiele innych zaawansowanych funkcji, jednak to właśnie sposób przekazywania ramek decyduje o tym, że dane urządzenie możemy nazwać switchem.

Rysunek 2.7 pokazuje konkretny przypadek, w którym wykorzystano dwie ścieżki transmisji ramek: z portu 4 do 6 oraz z portu 1 do 7. Dodatkową zaletą działania switcha jest to, że obie te transmisje mogą odbywać się równolegle. Zwiększa to znacznie postrzeganą prędkość sieci. Dedykowana ścieżka pomiędzy wybranymi portami tworzona jest każdorazowo na potrzeby konkretnej transmisji. Pozostałe porty w danym momencie nie uczestniczą w tym procesie.



Rysunek 2.7. Przełącznik transmituje ramki tylko do tych portów, do których zostały skierowane

W jaki więc sposób switch decyduje, gdzie przesyłać konkretne ramki transmitowane przez różne stacje w sieci? Każda ramka Ethernet zawiera w sobie informacje o źródłowym oraz docelowym adresie MAC. Przełącznik otwiera ramkę, odczytuje źródłowy adres MAC i zapisuje go w specjalnej tablicy (chyba że jest on już tam obecny). Tablica ta zwana jest **tablicą CAM** (ang. *content-addressable memory table*) w przypadku systemów operacyjnych CatOS lub **tablicą adresów MAC** (ang. *MAC address table*) w systemach IOS. Zawiera ona mapę adresów MAC rozpoznanych na konkretnych portach urządzenia. Następną czynnością jest odczytanie przez przełącznik docelowych adresów MAC dla tych ramek i porównanie ich z tablicą. Jeśli istnieje odpowiednia para adresów, tworzone jest połączenie między portem źródłowym a docelowym. Jeśli switch nie znajdzie pary, przesyła ramkę na wszystkie porty.

Jeśli stacja korzystająca z adresu IP chce przesłać pakiet do urządzenia o innym określonym adresie IP w tej samej sieci, musi najpierw określić adres MAC urządzenia używającego tego docelowego adresu IP. W tym celu wysyłane jest żądanie typu ARP request (ang. *Address Resolution Protocol*). Jest to pakiet rozgłoszeniowy przekazywany na wszystkie porty przełącznika. Pakiet ARP po enkapsulacji do postaci ramki zawiera adres MAC hosta wysyłającego żądanie. Switch może więc określić, który port oznaczyć jako port źródłowy. Kiedy stacja z żądanym adresem IP odpowie na zapytanie, przełącznik może odczytać, na którym porcie zlokalizowany jest docelowy adres MAC (ramka z odpowiedzią będzie zawierać ten adres MAC).

W przełącznikach z systemem operacyjnym IOS możemy skorzystać z komendy `show mac-address-table`, aby wyświetlić tablicę adresów MAC i odpowiadających im portów. Wiele adresów MAC na jednym porcie oznacza zazwyczaj, że do tego portu podłączony jest kolejny przełącznik lub urządzenie sieciowe:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Switch1-IOS>sho mac-address-table
Legend: * - primary entry //pierwotne wejście
        age - seconds since last seen //czas w sekundach od ostatniego pokazania się
        n/a - not available //niedostępny
  vlan   mac address     type    learn     age     ports
------+----------------+--------+-----+----------+-------------------------
*   24  0013.bace.e5f8   dynamic  Yes        165   Gi3/4
*   24  0013.baed.4881   dynamic  Yes         25   Gi3/4
*   24  0013.baee.8f29   dynamic  Yes         75   Gi3/4
*    4  0013.baeb.ff3b   dynamic  Yes          0   Gi2/41
*   24  0013.baee.8e89   dynamic  Yes        108   Gi3/4

*   18  0013.baeb.01e0   dynamic  Yes          0   Gi4/29
*   24  0013.2019.3477   dynamic  Yes        118   Gi3/4
*   18  0013.bab3.a49f   dynamic  Yes         18   Gi2/39
*   18  0013.baea.7ea0   dynamic  Yes          0   Gi7/8
*   18  0013.bada.61ca   dynamic  Yes          0   Gi4/19
*   18  0013.bada.61a2   dynamic  Yes          0   Gi4/19
*    4  0013.baeb.3993   dynamic  Yes          0   Gi3/33
```

Z powyższej tabeli możesz odczytać, że jeśli urządzenie z adresem `MAC 0013.baeb.01e0` spróbuje komunikować się z urządzeniem o adresie `MAC 0013.baea.7ea0`, to przełącznik będzie ustanawiał połączenie między portami Gi4/29 i Gi7/8.

Z pewnością zauważyłeś, że podałem w opisie pierwszy człon komendy jako „show", natomiast podczas wpisywania komendy użyłem wersji skróconej „sho". Urządzenia Cisco umożliwiają stosowanie takich skrótów poleceń, pod warunkiem, że brzmienie skrótu nie będzie takie samo dla większej liczby poleceń.

Ta informacja jest również przydatna, jeśli potrzebujesz określić, w którym miejscu dane urządzenie podłączone jest do przełącznika. Najpierw musisz poznać adres MAC urządzenia, którego szukasz. Oto przykład z systemu Solaris:

```
[root@unix /]$ifconfig -a
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232 index 1
        inet 127.0.0.1 netmask ff000000
dmfe0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
        inet 172.16.1.9 netmask ffff0000 broadcast 172.16.255.255
        ether 0:13:ba:da:d1:ca
```

Teraz możesz użyć tego adresu MAC (wyświetlony w ostatniej linijce powyżej) w poleceniu `show mac-address-table | include` *`adres-mac`* w interfejsie systemu operacyjnego IOS przełącznika:

```
Switch1-IOS>sho mac-address-table | include 0013.bada.d1ca
*   18  0013.bada.d1ca   dynamic  Yes          0   Gi3/22
```

Zwróć uwagę, że format wyświetlania adresów MAC może się różnić w zależności od systemu operacyjnego. Musisz więc przekonwertować go na odpowiedni dla CatOS lub IOS format. IOS wyświetla każdą grupę dwubajtowych par oddzieloną kropką. Solaris, jak i większość pozostałych systemów operacyjnych, wyświetla każdy oktet oddzielony dwukropkiem lub myślnikiem (CatOS używa myślnika jako ogranicznika podczas wyświetlania adresów MAC w systemie szesnastkowym). Jedne systemy używają do zapisu systemów dziesiętnych, inne zaś szesnastkowych.

Dane wynikowe zadanego polecenia informują, że nasz serwer podłączony jest do portu Gi3/22.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W systemach operacyjnych NX-OS polecenia są takie same jak w przypadku IOS. Jedyna różnica jest taka, że w interfejsie przełącznika mamy informację o sprzęcie, z którego korzystamy (w tym przypadku z urządzenia Nexus 5020 z 2148T skonfigurowanym jako FEX100):

```
NX-5K-1(config-if)# sho mac-address-table
VLAN      MAC Address       Type    Age       Port
---------+-----------------+-------+---------+-----------------------------
100       0005.9b74.b811    dynamic 0         Po100
100       0013.bada.d1ca    dynamic 40        Eth100/1/2
Total MAC Addresses: 2
```

W przypadku przełączników z systemem CatOS działa to trochę inaczej, ponieważ komenda `show cam` zawiera opcję wyświetlenia konkretnego adresu MAC.

```
Switch1-CatOS: (enable)sho cam 00-00-13-ba-da-d1-ca
* = Static Entry. + = Permanent Entry. # = System Entry. R = Router Entry.
X = Port Security Entry $ = Dot1x Security Entry
VLAN  Dest MAC/Route Des    [CoS]  Destination Ports or VCs / [Protocol Type]
----  ------------------    -----  ------------------------------------------
20    00-13-ba-da-d1-ca             3/48 [ALL]
Total Matching CAM Entries Displayed  =1
```

Rodzaje przełączników

Switche Cisco możemy podzielić na dwa rodzaje: switche o stałej konfiguracji oraz modularne. Przełączniki o stałej konfiguracji są mniejsze, zazwyczaj o wysokości 1 U. Zawierają na ogół jedynie porty Ethernet. Przeznaczone są do instalacji, w których nie ma potrzeby stosowania większych przełączników.

Przykładami przełączników Cisco o stałej konfiguracji są switche z serii 2950, 3550 oraz 3750. Przełączniki 3750 charakteryzują się możliwością **łączenia w stos**, czyli stakowania (ang. *stack*). Stakowanie polega na łączeniu kilku przełączników w jedną logiczną całość. Może to być użyteczne, gdy potrzebujesz większej liczby portów niż ta, którą oferuje pojedynczy switch o stałej konfiguracji (maksymalnie 48 portów). Wadą tej metody łączenia switchy jest ograniczenie prędkości magistrali całego stosu do 32 lub 64 Gb/s. Dla porównania niektóre z przełączników modularnych typu *chassis-based* oferują prędkość magistrali nawet do 720 Gb/s. Takie duże modułowe przełączniki są jednak z reguły droższe niż stos przełączników o stałej konfiguracji.

Korzyści, jakie oferują przełączniki o stałej konfiguracji, to:

Cena
: Są one z reguły dużo tańsze niż ich modularni kuzyni.

Rozmiar
: Przełączniki o stałej konfiguracji nie przekraczają najczęściej rozmiaru 1 U, przez co mogą być montowane w niewielkich szafach czy pomieszczeniach, w których nie zmieściłyby się switche typu *chassis-based*. Nawet dwa switche połączone ze sobą w stos są wciąż mniejsze niż najmniejszy przełącznik *chassis-based*.

Ciężar
: Switche te są dużo lżejsze od switchy *chassis-based*, do montażu których potrzeba z reguły minimum dwóch osób.

Zasilanie
: Przełączniki o stałej konfiguracji mogą być zasilane bezpośrednio ze standardowej sieci energetycznej, a więc można je montować praktycznie wszędzie. Większe przełączniki

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

chassis-based, gdy są w pełni wyposażone we wszystkie moduły, wymagają specjalnych zasilaczy oraz gniazd zasilania prądu zmiennego. Wiele przełączników jest również dostępnych w opcji zasilania prądem stałym (DC).

Z drugiej strony duże przełączniki modularne Cisco typu *chassis-based* mają pewną przewagę nad swoimi mniejszymi konkurentami. Oto ich podstawowe zalety:

Możliwość rozbudowy
Przełączniki te mogą obsługiwać setki portów Ethernet, a architektura *chassis-based* umożliwia łatwą rozbudowę modułów przetwarzania. Dla serii 6500 dostępne są moduły zapewniające prędkość magistrali 720 Gb/s. W przypadku połączenia w stos 7 switchy z serii 3750 tak, aby uzyskać porównywalną liczbę portów, musimy pogodzić się z ograniczeniem prędkości magistrali do 32 Gb/s.

Elastyczność
Przełączniki Cisco z serii 6500 umożliwiają rozbudowę o moduły oferujące funkcjonalność, której nie posiadają standardowe switche. Do takich modułów możemy zaliczyć:

- moduły Firewall (ang. *Firewall Services Module* — FWSM);
- moduły detekcji włamań (ang. *Intrusion Detection System Module* — IDSM);
- moduły kontroli w warstwie aplikacji (ang. *Application Control Engines* — ACE);
- moduły analizy sieci — (ang. *Network Analysis Module* — NAM);
- moduły WAN (FlexWAN).

Redundancja
Niektóre modele przełączników o stałej konfiguracji są w stanie obsługiwać jednostki dystrybucji zasilania, które mogą zapewnić pewien stopień redundancji zasilania. Oczywiście wiąże się to z dodatkowymi kosztami. Switche typu *chassis-based* z zasady są przystosowane do obsługi wielu źródeł zasilania (starsze modele z serii 4000 wymagały 3 źródeł zasilania dla systemu redundancji i nawet jeszcze więcej dla obsługi VoIP). Większość switchy *chassis-based* obsługuje też system podwójnych modułów zarządzania.

Prędkość
Przełączniki Cisco z serii 6500 posiadają moduły zarządzania oferujące prędkości magistrali do 720 Gb/s. Najszybszy przełącznik o stałej konfiguracji — model 4948 — zapewnia zaledwie 48 Gb/s. 4948 jest przeznaczony do zamontowania jako najwyższy element szafy w celu wspierania pozostałych urządzeń montowanych na niższych poziomach. Z uwagi na swoją specjalizację nie posiada on możliwości stakowania, więc jest ograniczony do 48 portów.

Urządzenia *chassis-based* mają oczywiście swoje wady, do których należą głównie duży ciężar i rozmiary oraz pobieranie dużej ilości energii. Jeśli jednak potrzebujesz mocy i elastyczności oferowanej przez te przełączniki, musisz potraktować wspomniane niedogodności jako element kosztów wynikających z prowadzenia działalności.

Podstawowymi dwoma produktami Cisco w technologii *chassis-based* są przełączniki Catalyst z serii 4500 i 6500. Jest jeszcze seria 8500, jednak przełączniki te są rzadko spotykane w środowisku korporacyjnym.

Switche z linii Nexus mogą być, w zależności od modelu, zakwalifikowane do jednej lub do drugiej kategorii. Nexus 7000 dostępny jest w obudowach posiadających 10 lub 18 slotów (gniazd na moduły). Model ten posiada wszelkie zalety i wady przełączników *chassis-based*,

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

jednak według specyfikacji nie obsługuje na razie wszystkich typów modułów. Linia ta nadal się rozwija, więc w przyszłości może to ulec zmianie.

Przełączniki Nexus z serii 5000 posiadają, w zależności od modelu, jeden (5010) lub dwa (5020) sloty na moduły rozszerzające. Dodatkowo można do nich podłączyć urządzenia zwane FEX (ang. *Fabric Extender*), które fizycznie przypominają przełączniki, w rzeczywistości działają jednak bardziej jak moduły podłączone do urządzeń z serii 5000. Przełączniki z serii Nexus 2000 działają wszystkie jako FEX. Z kolei modele 5000 (pojedynczo lub w parze) dają możliwość podłączenia do dwunastu (24 dla pary urządzeń) 48-portowych urządzeń FEX. Działają przez to bardziej jako 12-slotowy przełącznik *chassis-based* niż przełącznik o stałej konfiguracji. Więcej szczegółów na temat rozbudowy switchy Nexus zamieściłem w oddzielnym rozdziale.

Planowanie instalacji przełączników typu chassis-based

Instalacja tego typu przełączników wymaga więcej planowania niż instalacje wykorzystujące mniejsze i nie tak skomplikowane urządzenia. Przy konfiguracji switchy *chassis-based* należy wziąć pod uwagę wiele elementów. Decyzja, z jakich modułów będziesz korzystał, wpłynie na wielkość zasilacza, który będziesz musiał zastosować. Powinieneś też zdecydować, czy skorzystać z zasilania prądem zmiennym, czy też stałym, i jakiego natężenia prądu wymaga instalacja. Z uwagi na duży ciężar i rozmiary urządzeń trzeba przeznaczyć na nie odpowiednią ilość miejsca w szafie rackowej. Poniżej znajdziesz kilka podstawowych zasad stosowanych podczas planowania instalacji przełączników *chassis-based*.

Miejsce w szafie rackowej

Przełączniki tego typu mogą być dość duże. Przykładowo model 6513 wymaga 19 U przestrzeni w szafie rackowej. Wersja NEBS modelu 6509 to 21 U wysokości. Nexus 7010 z 10 slotami również zapełni 21 U, a model 7018 z 18 slotami wymaga 25 U. Szafa telekomunikacyjna o wysokości około 213 cm to w przeliczeniu 40 U, takie switche zajmują więc znaczącą część dostępnej przestrzeni.

Bądź ostrożny w przypadku planowania instalacji przełączników Nexus. Nexus z serii 5000 charakteryzują się podobną głębokością co serwery rackowe i zajmują całą głębokość szafy. Z tego powodu nie mogą być montowane w szafach z dwoma profilami.

Większe przełączniki są bardzo ciężkie i powinny być instalowane tak nisko w szafie, jak to możliwe. Lżejsze modele (takie jak 4506, który ma wysokość 10 U) można instalować nieco wyżej.

Do instalacji przełączników *chassis-based* potrzeba zawsze przynajmniej dwóch osób. Często przydaje się też osoba trzecia, która pomoże naprowadzić urządzenie na odpowiednie miejsce w szafie rackowej. Pamiętaj, żeby podnosić takie przełączniki dopiero po wymontowaniu wszystkich modułów i zasilaczy.

Zasilanie

Każdy moduł pobiera określoną ilość energii mierzonej w watach. Jeśli już wybierzesz, z których modułów będziesz korzystał, będziesz mógł obliczyć wymaganą moc zasilacza. W przypadku systemów zasilania redundantnego każdy z zasilaczy w parze musi mieć moc wystarczającą

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

do zasilenia całego urządzenia ze wszystkimi modułami. Tak więc jeśli moduły wymagają 3200 W, będziesz potrzebował dwóch zasilaczy po 4000 W dla systemu zasilania redundantnego. Jeśli użyjesz dwóch zasilaczy o mocy 3000 W każdy, to oba z nich będą musiały pracować równolegle. W przypadku awarii jednego z nich część modułów zostanie odłączona.

W zależności od miejsca instalacji przełącznika możesz zostać zmuszony do wyboru typu zasilania: AC lub DC. W przypadku zasilaczy DC pamiętaj, aby wyznaczyć **kanały A i B**. Przykładowo jeśli potrzebujesz 40 amperów zasilania DC, to powinieneś mieć **40 amperów DC na kanale A i B**. Oznacza to dwa 40-amperowe obwody dla celów zachowania zasilania w razie awarii. Szczegóły dotyczące sposobu instalacji znajdziesz w specyfikacjach technicznych określonych produktów.

W przypadku zasilaczy AC powinieneś znać wymagane napięcie, natężenie oraz rodzaj gniazda dla każdego kanału zasilania. Typowy zasilacz wymaga jednego kanału zasilającego, jednak niektóre mogą potrzebować ich więcej. Przyda się tutaj znajomość terminologii dotyczącej wtyczek i gniazd. Wszystkie te informacje powinny być zawarte w specyfikacji technicznej dostępnej na stronie internetowej producenta.

Podłączając kable zasilające do gniazd zasilacza, należy zawsze korzystać z zacisków umieszczonych zazwyczaj obok gniazda na panelu zasilacza. Zaciski te przytrzymują kable, zapobiegając ich wypięciu w razie przypadkowego szarpnięcia kablem.

Chłodzenie

Wiele przełączników typu *chassis-based* jest chłodzonych od strony ścian bocznych. Powietrze zasysane jest z jednego boku i wydmuchiwane z drugiego. Zazwyczaj montaż switchy w szafie rackowej zapewnia odpowiedni przepływ powietrza. Zwróć jednak uwagę, czy kable, które biegną z reguły po bokach szafy, nie utrudniają prawidłowej wentylacji.

W przełączniku 6509 mamy do czynienia z pionową cyrkulacją powietrza. Pionowo są też umieszczone moduły w obudowie. Otwory wentylacyjne znajdują się na przednim panelu obudowy — pamiętaj, żeby utrzymywać je w czystości. We wszystkich przełącznikach Nexus obieg powietrza odbywa się od strony panelu przedniego w kierunku ściany tylnej. O ile w większości przypadków umieszcza się porty Ethernet na przednim panelu szafy, to dla przełączników Nexus nie jest to najlepszy pomysł. Zakłócałoby to obieg powietrza w szafie.

Pracowałem kiedyś nad projektem wykorzystującym 6 przełączników 6506. Wyciągnęliśmy je z klatek i umieściliśmy obok siebie na kilku paletach. Nie zauważyliśmy, że ułożyliśmy je tak, iż ciepłe powietrze wydmuchiwane z otworów wentylacyjnych jednego przełącznika trafiało wprost do otworów wentylacyjnych, którymi było zasysane przez kolejny przełącznik. W rezultacie ostatni z przełączników w szeregu wyłączył się z przegrzania. Zawsze upewnij się, że zostawiłeś odpowiednią ilość miejsca pomiędzy switchami w trakcie instalacji, tak aby zapewnić prawidłową wentylację.

Instalowanie i usuwanie modułów

Moduły przełączników *chassis-based* są wsuwane wzdłuż szyn umieszczonych po obu stronach gniazda (slotu). Przy instalowaniu modułu należy być bardzo ostrożnym, aby nie zablokować go przez niewłaściwe ułożenie podczas wsuwania do gniazda. Wiele modułów jest wyposażonych w dużą liczbę komponentów, a ich cena może wynosić nawet kilkadziesiąt tysięcy

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

złotych. Dotyczy to chociażby modułów FWSM, IDSM, CSM czy ACE. Widziałem inżynierów, którym udało się zniszczyć takie moduły poprzez niewłaściwe naprowadzenie na szyny montażowe gniazda. Pamiętaj też, aby podczas instalacji użyć opaski uziemiającej na rękę.

Podczas instalowania przełączników lub modułów zawsze powinieneś używać opaski uziemiającej. Taka opaska jest łatwa w użyciu i często dostarczana ze sprzętem jako element zestawu. Zakładanie jej może wydawać się głupie, lecz pomyśl, jak się poczujesz, kiedy zniszczysz sprzęt za kilkaset tysięcy złotych?

Prowadzenie kabli

Prowadząc kable do modułów, miej na uwadze, że być może w przyszłości zajdzie konieczność usunięcia modułu. Podciągnięcie 48 kabli Ethernet do danego elementu może być dość trudnym zadaniem. Pamiętaj, żeby zostawić odpowiednio dużo luzu na kablach, tak aby móc je łatwo odsunąć na bok w przypadku wyjmowania modułu. Wykonanie takiej operacji może się okazać konieczne w razie awarii elementu, kiedy trzeba będzie wymienić go na nowy. Im więcej rzeczy zaplanujesz z wyprzedzeniem, tym łatwiejsze będziesz miał zadanie.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 3.

Autonegocjacja

Kiedy jestem wzywany do klienta w celu zdiagnozowania problemów związanych ze spowalnianiem działania sieci lub „wolno" działającym urządzeniem, patrzę najpierw na statystyki błędów oraz ustawienia autonegocjacji przełączników, a także urządzeń do nich podłączonych. Jeśli miałbym sporządzić listę najbardziej powszechnych problemów, z którymi zetknąłem się w ciągu wielu lat praktyki, to kłopoty z autonegocjacją znalazłyby się w pierwszej piątce, jeśli nie na pierwszym miejscu.

Czemu kwestie związane z autonegocjacją są tak często występującym problemem? Prawdę mówiąc, wiele osób nie do końca rozumie, za co jest ona odpowiedzialna i jak działa. Przyjmują więc błędne założenia, które prowadzą do powstawania problemów.

Co to jest autonegocjacja?

Autonegocjacja to funkcja, która umożliwia wybranie właściwego trybu dupleks oraz prędkości połączenia podczas komunikacji pomiędzy wybranym portem przełącznika, routera, serwera lub innego urządzenia a urządzeniem znajdującym się po drugiej stronie łącza. Parametry transmisji są dynamicznie dostosowywane do wartości, jaką oferują interfejsy łącza. Przyjrzyjmy się konfiguracji dostępnych parametrów:

Prędkość

Dla każdego interfejsu dostępna jest konkretna prędkość łącza określana zazwyczaj w megabitach na sekundę (Mb/s). Standardowe prędkości dla Ethernetu to 10 Mb/s, 100 Mb/s oraz 1000 Mb/s. Ta ostatnia nazywana jest **Gigabitowym Ethernetem**. Wiele przełączników obsługuje już prędkość 10 Gb/s, a w niedalekiej przyszłości może to być nawet 100 Gb/s.

Dupleks

Dupleks określa sposób przepływu danych dla konkretnego interfejsu. W trybie półdupleks (ang. *half duplex*) dane mogą być w danym momencie tylko wysyłane albo tylko odbierane, nigdy jednocześnie. Rozmowa przez radio z reguły obywa się w trybie półdupleks — jeśli chcesz mówić, naciskasz odpowiedni przycisk i w tym czasie nie możesz słyszeć drugiej osoby. W pełnym dupleksie (ang. *full-duplex*) można jednocześnie wysyłać i odbierać dane, tak jak w przypadku rozmowy przez telefon.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jak działa autonegocjacja?

Przede wszystkim powiedzmy, czego funkcja autonegocjacji **nie robi**. Jeśli autonegocjacja jest włączona na porcie urządzenia znajdującego się po jednej stronie łącza, nie determinuje to automatycznie konfiguracji na porcie urządzenia po drugiej stronie Ethernetu — nie zachodzi automatyczne dopasowanie parametrów. Dość powszechne jest jednak przekonanie, że tak właśnie ta funkcja powinna działać. Może to prowadzić do wielu problemów.

Autonegocjacja jest protokołem i jak każdy protokół działa tylko wtedy, kiedy jest uruchomiona po obu stronach łącza. Innymi słowy, jeśli autonegocjacja jest włączona tylko po jednej stronie, to **nie może** ona dopasować parametru prędkości oraz trybu dupleks po drugiej stronie. Jeśli jednak również na drugiej końcówce uruchomiony został ten protokół, to oba urządzenia **wspólnie** decydują o wyborze najodpowiedniejszej prędkości i właściwego dupleksu. Każde urządzenie przedstawia, w jakich prędkościach i dupleksach może pracować, a następnie wybierana jest optymalna zgodność (preferowane są wysokie prędkości oraz tryb pełnego dupleksu).

Pierwotnym powodem zamieszania jest to, że autonegocjacja wydaje się działać praktycznie w każdym przypadku. Dzieje się tak z powodu funkcji zwanej **wykrywaniem równoległym** (ang. *parallel detection*). Jest ona uruchamiana, gdy protokół autonegocjacji nie zostanie wykryty po drugiej stronie łącza. Detekcja równoległa polega na tym, że do lokalnych sterowników 10Base-T, 100Base-TX oraz 100Base-T4 wysyłany jest sygnał. Jeśli któryś ze sterowników wykryje sygnał, interfejs jest ustawiany na daną prędkość.

Detekcja równoległa pozwala ustalić jedynie prędkość transmisji, nie obsługuje jednak trybów dupleksu. To bardzo ważne rozróżnienie, ponieważ powszechnie stosowane standardy Ethernetu mogą wspierać różne poziomy konfiguracji dupleksu:

10Base-T
: 10Base-T został zaprojektowany bez wsparcia dla pełnego dupleksu. Niektóre implementacje tego standardu mogą mimo to obsługiwać pełny dupleks, jednak większość z nich nie daje takiej możliwości.

100Base-T
: 100Base-T od dawna wspiera tryb pełnego dupleksu, który jest preferowany dla połączeń o prędkości 100 Mb/s od momentu powstania standardu. Domyślnym ustawieniem dla 100Base-T jest jednak półdupleks, a pełny dupleks musi być ustawiony ręcznie.

1000Base-T
: Gigabitowy Ethernet charakteryzuje się wydajniejszym mechanizmem działania protokołu autonegocjacji niż Ethernet 10 Mb i 100 Mb. W większości przypadków interfejsy gigabitowe powinny być skonfigurowane do automatycznej negocjacji.

10 Gigabit Ethernet
: Połączenia 10-gigabitowe (10G) generalnie determinowane są przez transceivery światłowodowe oraz specjalne końcówki kabla miedzianego różniące się od złączy RJ-45 stosowanych w pozostałych typach Ethernetu. To, jakich złączy 10G użyjemy, zależeć będzie od zastosowanego sprzętu. W urządzeniach Nexus 6500 interfejsy 10G są z reguły realizowane przez moduły XENPAK, które nie negocjują do niższych prędkości. W przełącznikach Nexus 5000 niektóre porty obsługują prędkości 1G/10G, które ustawia się za pomocą odpowiedniej komendy.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

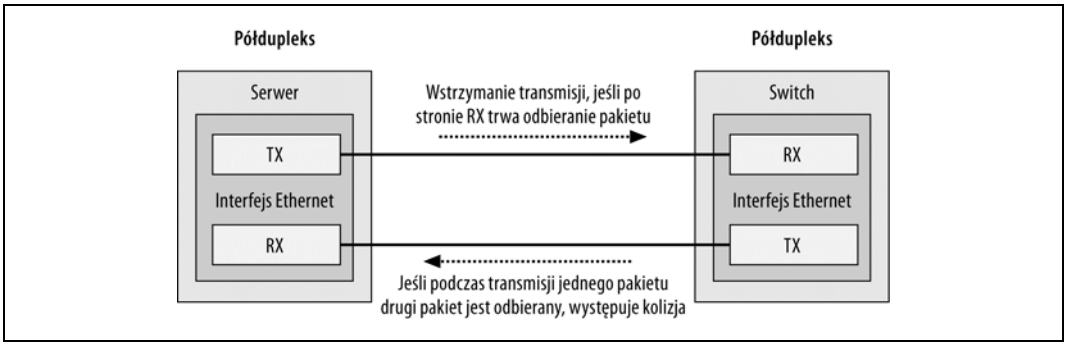
Jak już wspomniałem, standard 10Base-T nie wspiera generalnie pełnego dupleksu, a domyślnym ustawieniem dla 100Base-T jest półdupleks. Z tych względów, kiedy zawodzi protokół autonegocjacji i uruchamia się detekcja równoległa (obsługująca jedynie tryby prędkości), najbezpieczniejszym rozwiązaniem jest to, aby sterownik zszedł do półdupleksu.

Z uwagi na stałą ewolucję technologii sieciowych oraz rozwój sprzętu łącza o wyższych prędkościach i ze stabilniej działającym protokołem autonegocjacji prawdopodobnie wyeliminują w przyszłości problemy związane z tym tematem. Jednak trzeba mieć na uwadze, że w praktyce cały czas spotykamy działające w strukturach sieciowych routery sprzed 20 lat. Wiedza na temat protokołu autonegocjacji może być więc użyteczna jeszcze przez długi czas.

Kiedy autonegocjacja zawodzi

Jeśli na łączach 10/100 autonegocjacja nie działa, to prawdopodobnie po jednej stronie ustawiono parametry 100 Mb/s oraz pełny dupleks, a po drugiej włączono protokół autonegocjacji. Powoduje to, że po jednej stronie mamy 100 Mb/s oraz pełny dupleks, a po drugiej 100 Mb/s oraz półdupleks.

Rysunek 3.1 przedstawia połączenie w trybie półdupleks. W środowisku półdupleks linia odbiorcza (RX) jest monitorowana. Jeśli jakakolwiek ramka zajmuje linię RX, to kolejna ramka nie zostanie wysłana, dopóki linia nie będzie wolna. Jeśli w danym momencie linia RX odbiera ramkę, a linia transmisyjna (TX) wysyła ramkę, występuje kolizja. Taka sytuacja jest odnotowywana przez licznik błędów kolizji, a ramka, która nie została wysłana, jest retransmitowana po losowym opóźnieniu wynikającym ze wstrzymania pierwotnej wysyłki. Może się to wydawać sprzeczne z intuicyjnym podejściem w przypadku nowoczesnych struktur, w których stosuje się przełączniki. Powinieneś jednak pamiętać, że Ethernet został zaprojektowany z przeznaczeniem do działania na pojedynczych przewodach, a switche i skrętka zostały wprowadzone dużo później.

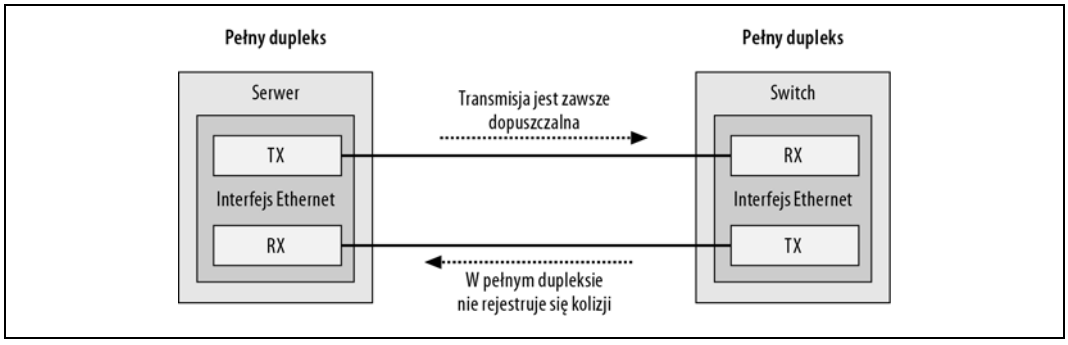


Rysunek 3.1. Półdupleks

Rysunek 3.2 przedstawia łącze w pełnym dupleksie. W takim trybie linia RX nie jest monitorowana, a linia TX jest zawsze traktowana jako dostępna. W pełnym dupleksie nie występują kolizje, gdyż linie RX i TX działają niezależnie.

Jeśli po jednej stronie łącza mamy do czynienia z pełnym dupleksem, a po drugiej z trybem półdupleks, doprowadzi to do wystąpienia znacznej liczby kolizji po stronie półdupleksu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

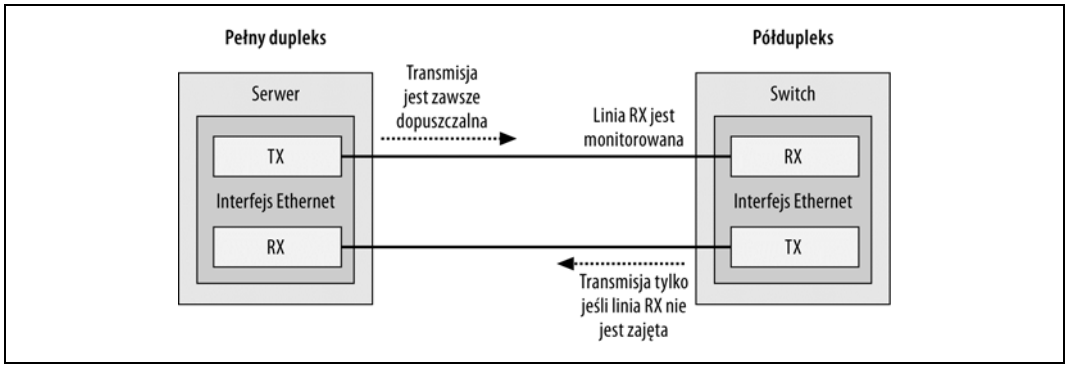


Rysunek 3.2. Pełny dupleks

Po stronie pełnego dupleksu ramki wysyłane są bez sprawdzania linii RX, więc w przypadku dużego natężenia ruchu sieciowego przechodzącego przez to urządzenie jest duże prawdopodobieństwo, że ramki będą wysyłane nieprzerwanie.

Końcówka działająca w półdupleksie będzie monitorować linię RX i nie wyśle żadnej ramki, dopóki linia się nie zwolni. W takim przypadku upłynie sporo czasu, zanim pojawi się możliwość transmisji ramki. Wystąpi też duża liczba kolizji, a samo urządzenie będzie sprawiało wrażenie powolnego działania. Kwestia ta może nie wydawać się tak oczywista, ponieważ standardowo interfejs w półdupleksie rejestruje kolizje, a interfejs w pełnym dupleksie tego nie robi. Skoro w trybie pełnego dupleksu nie jest przeprowadzany test na zajętość linii przed wysłaniem ramki, nie istnieje też możliwość rejestracji występowania błędów. Problem będzie się więc objawiał nadmiernym występowaniem kolizji, ale tylko po stronie półdupleksu.

Rysunek 3.3 pokazuje sytuację, w której nie zadziałała funkcja autonegocjacji.



Rysunek 3.3. Typowy scenariusz problemów z autonegocjacją

W praktyce, jeśli stwierdzisz, że interfejs skonfigurowany pod autonegocjację ustawił się na 100 Mb/s w trybie półdupleks, to prawdopodobnie na drugiej końcówce skonfigurowano ręcznie parametry 100 Mb/s w pełnym dupleksie. Interfejsy 100 Mb/s nieobsługujące pełnego dupleksu są bardzo rzadko spotykane, więc w przypadku prawidłowej konfiguracji protokołu autonegocjacji po obu stronach łącza urządzenia w zasadzie nigdy nie powinny przejść do pracy w półdupleksie.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Autonegocjacja w praktyce

Korzystanie z zalet protokołu autonegocjacji jest dość łatwe, pod warunkiem stosowania się do jednej zasady:

> Upewnij się, że po obu stronach łącza urządzenia skonfigurowane są w ten sam sposób.

Jeśli ustawiasz autonegocjację po jednej stronie łącza, musisz tak samo skonfigurować drugą końcówkę. Jeśli wybierasz tryb 100 Mb/s w pełnym dupleksie, również skonfiguruj tak samo oba urządzenia.

Bądź ostrożny, stosując tryb 10 Mb/s w pełnym dupleksie, ponieważ żadne urządzenia standardu Ethernet 10Base-T nie obsługują pełnego dupleksu.

W gigabitowym Ethernecie wykorzystuje się bardziej wydajny mechanizm autonegocjacji niż ten opisany w niniejszym rozdziale. Urządzenia w standardzie Gigabit Ethernet powinny być zawsze ustawione w trybie autonegocjacji, chyba że istnieje istotny powód, by postąpić inaczej (np. interfejs, który nie znegocjuje się prawidłowo). Powinieneś jednak traktować taką sytuację jako rozwiązanie tymczasowe, które będzie funkcjonować tylko do momentu wymiany niewłaściwie działającego elementu.

Konfiguracja autonegocjacji

W przełącznikach Cisco autonegocjacja jest ustawiona domyślnie. W interfejsie systemu IOS możesz konfigurować prędkość transmisji oraz tryb dupleksu poleceniami `speed` i `duplex`.

W pierwszej kolejności należy wybrać prędkość, a dopiero potem skonfigurować dupleks. W przeciwnym przypadku system zgłosi błąd, wyświetlając poniższy komunikat:

```
2950(config-if)#duplex half
Duplex can not be set until speed is set to non-auto value
        //Tryb dupleksu nie może być zdefiniowany, dopóki nie jest skonfigurowana inna wartość prędkości niż auto
```

W przypadku próby ustawienia trybu półdupleks w przełączniku gigabitowym na portach gigabitowych również otrzymasz ostrzeżenie:

```
SW-3750(config-if)#duplex half
% Duplex cannot be set to half when speed autonegotiation subset contains 1Gbps
        //Nie można ustawić trybu półdupleks, jeśli zakres wybranego protokołu autonegocjacji dopuszcza prędkość 1 Gb/s
```

Aby skonfigurować prędkość transmisji interfejsu, stosuje się komendę `speed`. Jeśli ustawienia interfejsu były wcześniej zmieniane, możesz powrócić do domyślnego trybu autonegocjacji, używając polecania `auto`:

```
SW-3750(config-if)#speed ?
  10    Force 10 Mbps operation
  100   Force 100 Mbps operation
  1000  Force 1000 Mbps operation
  auto  Enable AUTO speed configuration
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Po ustawieniu prędkości możesz wybrać tryb dupleksu — `auto`, `full` lub `half`:

```
SW-3750(config-if)#duplex ?
  auto  Enable AUTO duplex configuration
  full  Force full duplex operation
  half  Force half-duplex operations
```

W przypadku systemów operacyjnych NX-OS polecenia są takie same jak dla IOS:

```
NX-5K(config)#int e1/5
NX-5K(config-if)#speed ?
  10    Force 10 Mbps operation
  100   Force 100 Mbps operation
  1000  Force 1000 Mbps operation
  auto  Enable AUTO speed configuration
NX-5K(config-if)#
NX-5K(config-if)#speed 1000
NX-5K(config-if)#duplex full
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 4.

VLAN-y

Wirtualne sieci lokalne, czyli VLAN-y (ang. *Virtual LAN*), to wirtualny podział w obrębie przełącznika pozwalający wyodrębnić logiczne sieci LAN, z których każda funkcjonuje tak, jakby była skonfigurowana na osobnym fizycznie switchu. Przed wprowadzeniem sieci wirtualnych jeden przełącznik mógł obsługiwać tylko jedną sieć LAN. VLAN-y umożliwiły obsługę wielu sieci lokalnych na pojedynczym switchu. Zakładając brak luk w systemie operacyjnym przełącznika, nie ma możliwości, by ramka wygenerowana w obrębie jednego VLAN-u przedostała się do innego.

Kiedy pisałem pierwsze wydanie książki *Wojownik sieci*, nie słyszałem jeszcze o istnieniu exploitów potrafiących łamać bariery VLAN-ów (ang. *VLAN-hopping*). Istnieją jednak sposoby na obejście tych barier. Ataki typu „spoofing" lub „double tagging" (podwójne tagowanie) to przykłady uzyskiwania dostępu z jednej sieci VLAN do drugiej. Choć odczyt danych z innej sieci VLAN nie jest łatwy, to atak typu „denial-of-service" może być skuteczny pomiędzy VLAN-ami, zwłaszcza jeśli przełączniki połączone są kanałami trunkowymi. Nie polegaj więc na VLAN-ach jako głównym elemencie zabezpieczeń, szczególnie w środowiskach wysokiego ryzyka.

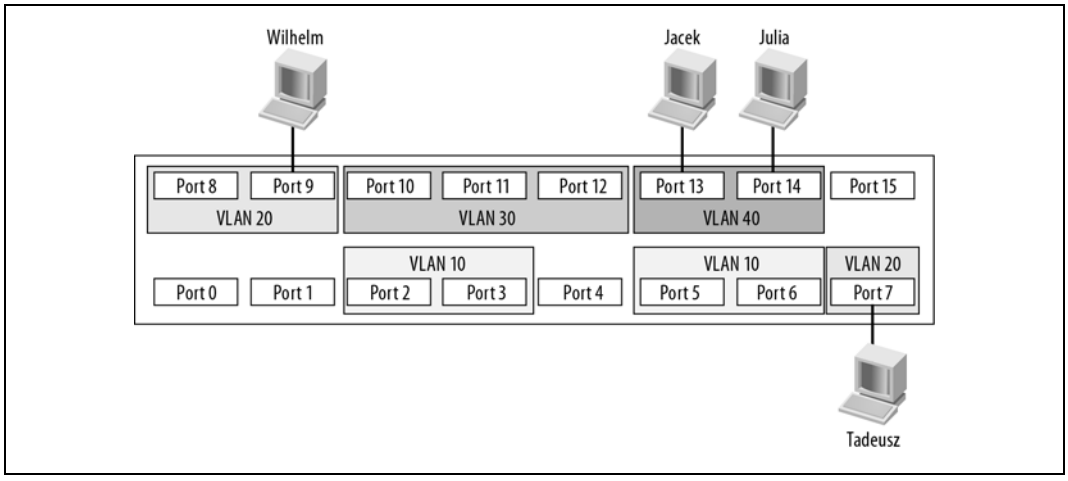
Łączenie VLAN-ów

Rysunek 4.1 pokazuje przełącznik, na którym skonfigurowano kilka wirtualnych sieci lokalnych. Tym VLAN-om nadano numery 10, 20, 30 oraz 40. Generalnie VLAN-y mogą być numerowane lub nazywane. We wdrożeniach Cisco do rozróżniania VLAN-ów domyślnie stosuje się numerację. Domyślny VLAN oznaczony jest numerem 1. Jeśli podepniesz kilka urządzeń do danego przełącznika, na którym nie konfigurowano VLAN-ów, to wszystkie będą się znajdowały w domyślnym VLAN-ie 1.

Ramki nie mogą opuszczać VLAN-ów, w obrębie których zostały wygenerowane. Oznacza to, że w naszej przykładowej konfiguracji Jacek może komunikować się z Julią, a Wilhelm może komunikować się z Tadeuszem. Wilhelm i Tadeusz nie mogą jednak w żaden sposób komunikować się z Jackiem i Julią.

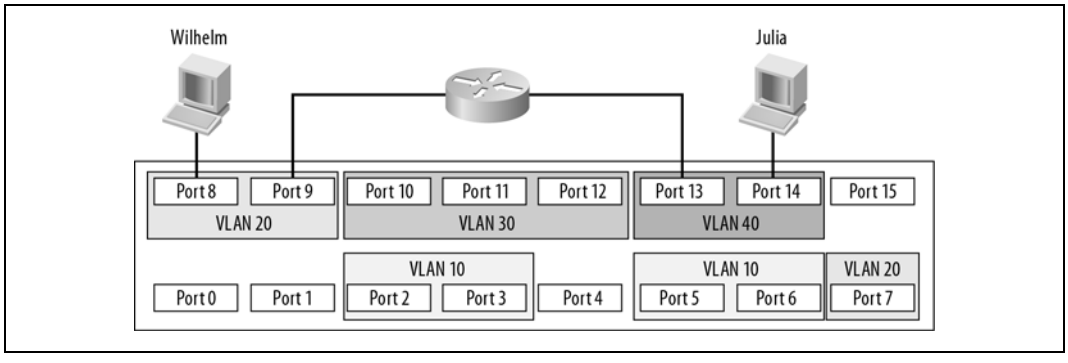
W przełączniku warstwy drugiej, aby pakiet mógł przejść z jednego VLAN-u do drugiego, muszą być one połączone zewnętrznym routerem. Umożliwi to routing pomiędzy VLAN-ami.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 4.1. VLAN-y na switchu

Rysunek 4.2 przedstawia router łączący VLAN 20 z VLAN-em 40. Zakładając prawidłową konfigurację routera, Wilhelm będzie teraz w stanie komunikować się z Julią. Nic nie będzie jednak nadal wskazywało na to, że obie stacje robocze są fizycznie podłączone do tego samego przełącznika.

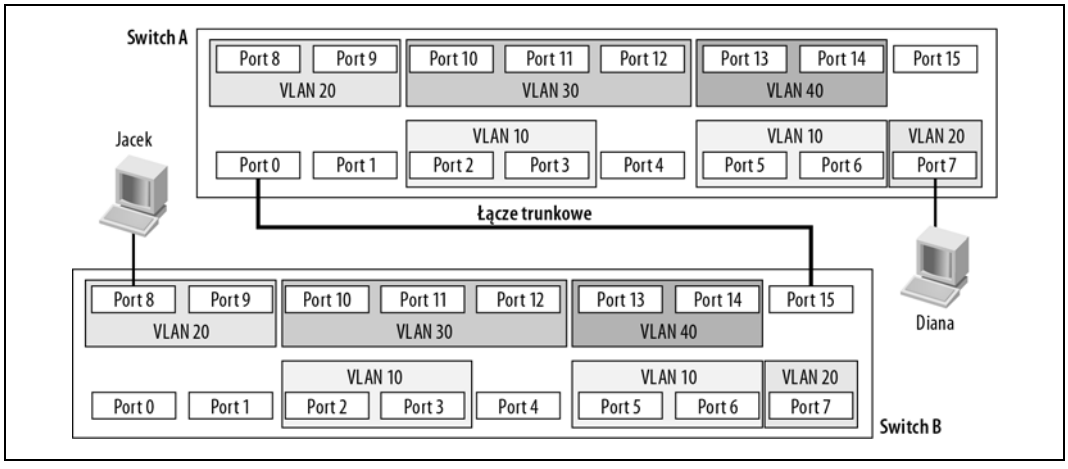


Rysunek 4.2. Zewnętrzny routing między VLAN-ami

Podczas rozbudowy sieci wykorzystującej VLAN-y również napotkasz te same ograniczenia. Jeśli podłączysz kolejny przełącznik do portu skonfigurowanego w VLAN-ie 20, to będzie on mógł wymieniać ramki jedynie z tym właśnie VLAN-em. Łącząc ze sobą dwa przełączniki, z których każdy posiada wyodrębnione po 4 VLAN-y, będziesz potrzebował 4 fizycznych połączeń — po jednym na każdy VLAN. Rozwiązaniem tego problemu może być zastosowanie **trunków** pomiędzy switchami. Trunki to kanały umożliwiające transport ramek z więcej niż jednej wirtualnej sieci lokalnej.

Rysunek 4.3 przedstawia dwa przełączniki połączone kanałem trunkowym. Jacek znajduje się w VLAN-ie 20 na przełączniku B, a Diana w VLAN-ie 20 na przełączniku A. Pomiędzy przełącznikami istnieje łącze trunkowe. Zakładając, że umożliwia ono transport ramek pomiędzy wszystkimi skonfigurowanymi VLAN-ami, Jacek będzie w stanie komunikować się z Dianą.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

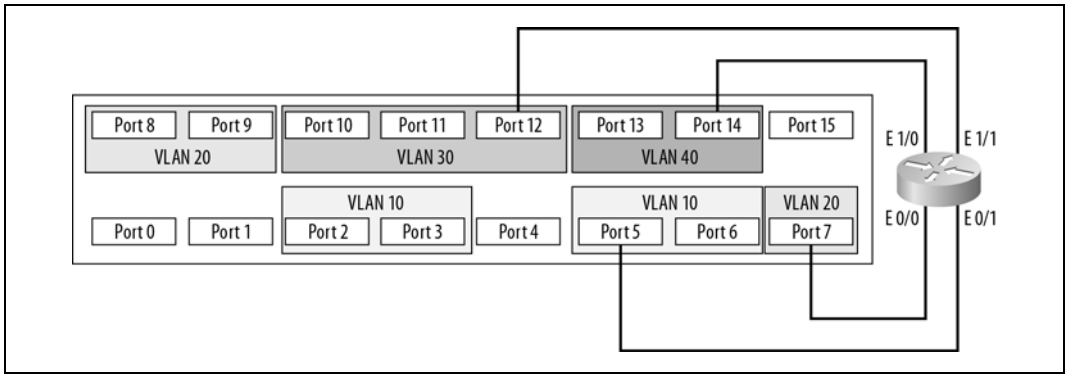


Rysunek 4.3. Dwa przełączniki połączone kanałem trunkowym

Zwróć uwagę na to, że porty wykorzystane do łącza trunkowego nie przynależą do żadnego VLAN-u. Są to **porty trunkowe** i jako takie nie są przypisane do żadnej konkretnej wirtualnej sieci lokalnej.

Trunking oferuje jeszcze inne, dodatkowe korzyści przy konfiguracji przełączników. Rysunek 4.2 przedstawia sposób, w jaki dwa VLAN-y mogą być połączone za pomocą routera, tak jakby były dwoma odseparowanymi fizycznie sieciami. Wyobraź sobie, że chcesz połączyć ze sobą **wszystkie** VLAN-y skonfigurowane na danym przełączniku. Jak tego dokonać, opierając się na powyższym schemacie? Tradycyjnie rozwiązaniem byłoby zapewnienie pojedynczego łącza między routerem a każdą z sieci. W tym przypadku każda z występujących sieci jest siecią wirtualną, należy więc stworzyć fizyczne połączenie między routerem a każdym VLAN-em.

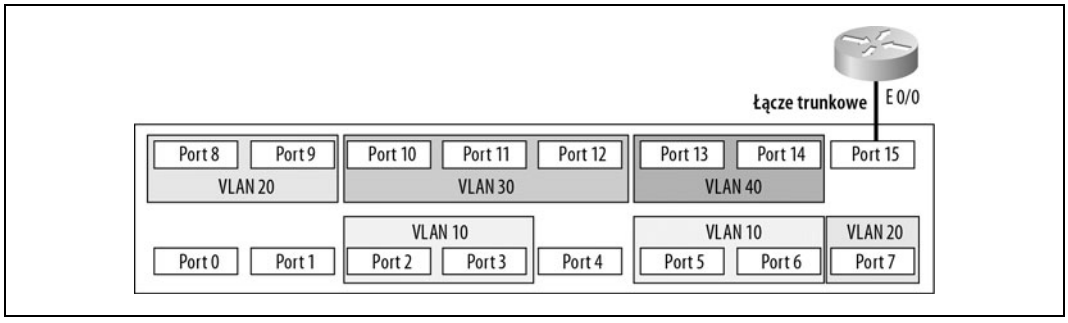
Jak widać na rysunku 4.4, metoda ta wymaga wykorzystania czterech portów zarówno po stronie routera, jak i po stronie przełącznika. Mniejsze routery nie zawsze muszą posiadać 4 interfejsy Ethernet, a każdy dodatkowy interfejs to rosnące koszty. Ponadto przy zakupie przełącznika, decydując się na urządzenie z określoną liczbą portów, masz już z reguły zaplanowane wykorzystanie wszystkich interfejsów. Przy tym sposobie konfiguracji jedna czwarta wszystkich portów przełącznika została wykorzystana do ustanowienia routingu między VLAN-ami.



Rysunek 4.4. Routing pomiędzy wieloma VLAN-ami

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Inną metodą ustanowienia routingu pomiędzy VLAN-ami jest konfiguracja zwana **routerem na patyku** (ang. *router-on-a-stick*). W takim przypadku zamiast tworzenia połączeń między routerem a każdym VLAN-em, wystarczy pojedynczy kanał trunkowy między routerem a przełącznikiem. Transmisje ze wszystkich VLAN-ów będą przepuszczane przez jedno łącze, tak jak pokazano to na rysunku 4.5.



Rysunek 4.5. Router na patyku

Takie rozwiązanie pozwala zmniejszyć liczbę portów wykorzystanych zarówno po stronie przełącznika, jak i routera. Ma ono jednak pewną wadę — mamy do czynienia z pojedynczym kanałem transmisyjnym o prędkości 10 Mb/s. Inaczej jest, kiedy każdy VLAN posiada własne łącze, co sprawia, że ma 10 Mb/s tylko dla siebie. Nie zapominaj też, że to router obsługuje ruch między VLAN-ami, są więc duże szanse, iż każda ramka będzie przechodzić przez kanał trunkowy dwukrotnie: raz w kierunku routera, a następnie z powrotem w kierunku sieci wirtualnej, do której została wysłana.

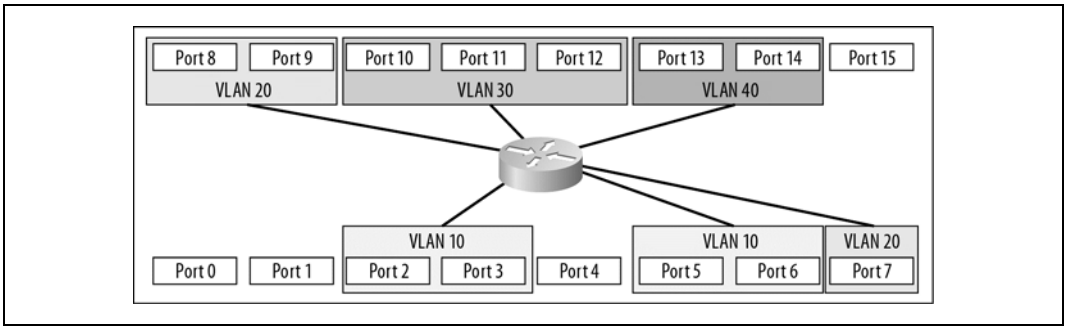
Kiedy edytowałem ten rozdział na potrzeby drugiego wydania, rozważałem aktualizację wszystkich schematów, aby dopasować je do obecnie powszechnie stosowanych prędkości transmisyjnych interfejsów. Zdecydowałem się jednak tego nie robić, gdyż ostatni raz widziałem osobę stosującą metodę routera na patyku, gdy najszybsze przełączniki oferowały interfejsy o prędkości 10 Mb/s. Moja decyzja nie wynikała więc z tego, że okazałem się zbyt leniwy, aby poprawić schematy.

Będąc dokładnym do bólu, muszę jednak wspomnieć, że podłączenie kanału trunkowego do zapory sieciowej (ang. firewall), która ma pełnić funkcję bramy domyślnej dla wielu VLAN-ów, opiera się na tej samej zasadzie. Mógłbyś więc stwierdzić, że mimo wszystko wykazałem się lenistwem. Skontrowałbym wtedy, przytaczając fakt, że najnowsza technologia przełączania wprowadzona przez Cisco w linii produktowej Nexus zniosła nazwy interfejsów opisujące ich prędkości. W zamian zastosowano dla wszystkich interfejsów Ethernet nazwę **e slot/port**. Tak więc nie jestem leniwy, lecz przewidujący.

W dalszej kolejności zacząłem pisać rozdział o protokole VoIP, w którym musiałem przedstawić konfigurację VLAN-u dla kanału VoIP, podłączonego trunkiem do routera, za pomocą metody routera na patyku. Dobrze się stało, że nie wyrzuciłem tej sekcji z książki. Potwierdziło się, że jestem przewidujący.

Rysunek 4.6 pokazuje koncepcyjnie, jak ten sam projekt może być zrealizowany za pomocą przełącznika warstwy trzeciej. Ponieważ przełącznik ten realizuje również funkcje routera, żadne zewnętrzne połączenia nie będą wymagane. Wszystkie porty przełącznika warstwy trzeciej mogą być więc przeznaczone do podpięcia urządzeń, innych przełączników lub do utworzenia kanałów trunkowych.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 4.6. Przełącznik warstwy trzeciej

Konfiguracja VLAN-ów

W systemach CatOS oraz IOS VLAN-y są standardowo konfigurowane za pomocą interfejsu wiersza poleceń CLI (ang. *command-line interpreter*), tak jak pozostałe funkcje. Jednak niektóre modele przełączników z systemami IOS, takie jak 2950 czy 3550, mogą posiadać konfigurowalną **bazę danych sieci VLAN** (ang. *VLAN database*) charakteryzującą się indywidualnymi trybami konfiguracji oraz poleceniami. Może to stanowić wyzwanie dla niewtajemniczonych użytkowników, zwłaszcza że konfiguracja tych baz danych jest całkowicie oddzielna od konfiguracji reszty funkcji przełącznika. Nawet sekwencja poleceń `write erase` oraz `reload` nie wyczyści bazy danych VLAN-ów w takich przełącznikach. Konfiguracja za pomocą bazy danych VLAN-ów jest powrotem do rozwiązań stosowanych w starszych urządzeniach, które nie oferowały innego trybu zarządzania lokalnymi sieciami wirtualnymi. Na szczęście wszystkie nowsze switche (również te posiadające bazę danych VLAN-ów) oferują opcję konfiguracji VLAN-ów za pomocą standardowego dla IOS interfejsu CLI. Switche takie jak 6500 podczas pracy w trybie natywnym IOS obsługują przy konfiguracji tylko polecenia IOS. Linia przełączników Nexus nie posiada w ogóle konfigurowalnej bazy danych VLAN-ów.

Cisco zaleca, aby konfigurację wirtualnych sieci lokalnych rozpocząć od skonfigurowania protokołu VTP (ang. *VLAN Trunking Protocol*). Ma to istotne znaczenie, gdyż kanały trunkowe nie będą negocjować bez domeny VTP. Ustanowienie domeny VTP nie jest jednak konieczne dla funkcjonowania VLAN-ów na pojedynczym przełączniku. Konfigurację VTP opisałem w kolejnych rozdziałach (rozdział 5. „Trunking" oraz rozdział 6. „Protokół VTP").

CatOS

W przypadku systemów operacyjnych CatOS VLAN-y tworzy się poleceniem `set vlan`:

```
Switch1-CatOS# (enable)set vlan 10 name Lab-VLAN
VTP advertisements transmitting temporarily stopped,
and will resume after the command finishes.
Vlan 10 configuration successful
```

Przy konfiguracji wirtualnych sieci lokalnych istnieje wiele opcji, opisałem tu jednak podstawowe minimum. Znajomość tego polecenia powinna wystarczyć na razie dla Twoich potrzeb. Aby sprawdzić status istniejących VLAN-ów, użyj polecenia `show vlan`:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Switch1-CatOS# (enable)sho vlan
VLAN Name                             Status    IfIndex Mod/Ports, Vlans
---- -------------------------------- --------- ------- ------------------------
1    default                          active    7       1/1-2
                                                        2/1-2
                                                        3/5-48
                                                        6/1-48
10   Lab-VLAN                         active    112
20   VLAN0020                         active    210     3/1-4
1002 fddi-default                     active    8
1003 token-ring-default               active    11
1004 fddinet-default                  active    9
1005 trnet-default                    active    10
1006 Online Diagnostic Vlan1          active    0       internal
1007 Online Diagnostic Vlan2          active    0       internal
1008 Online Diagnostic Vlan3          active    0       internal
1009 Voice Internal Vlan              active    0       internal
1010 Dtp Vlan                         active    0       internal
1011 Private Vlan Reserved Vlan       suspend   0       internal
1016 Online SP-RP Ping Vlan           active    0       internal
```

Zauważ, że VLAN 10 posiada własną nazwę, którą nadałeś, tworząc ten VLAN. VLAN 20, któremu nie przypisałeś żadnej nazwy, otrzymał nazwę domyślną VLAN0020. Z przedstawionej informacji można odczytać, które porty przypisane są do VLAN-u 20 oraz że większość portów nadal pozostaje w VLAN-ie 1 (jest to VLAN domyślny, dlatego znajdują się tam wszystkie porty nieprzypisane ręcznie do innych VLAN-ów).

W VLAN-ie 10 nie ma na razie żadnych portów, dodaj je więc, stosując polecenie `set vlan`:

```
Switch1-CatOS# (enable)set vlan 10 6/1,6/3-4
VLAN 10 modified.VLAN 1 modified.
VLAN  Mod/Ports
---- ------------------------
10  6/1,6/3-4
```

W ten sposób dodałeś do VLAN-u 10 porty 6/1, 6/3 oraz 6/4. Komenda `show vlan` pokaże, jakie zmiany zaszły w konfiguracji:

```
Switch1-CatOS# (enable)sho vlan
VLAN Name                             Status    IfIndex Mod/Ports, Vlans
---- -------------------------------- --------- ------- ------------------------
1    default                          active    7       1/1-2
                                                        2/1-2
                                                        3/5-48
                                                        6/2,6/5-48
10   Lab-VLAN                         active    112     6/1,6/3-4
20   VLAN0020                         active    210     3/1-4
1002 fddi-default                     active    8
1003 token-ring-default               active    11
1004 fddinet-default                  active    9
1005 trnet-default                    active    10
1006 Online Diagnostic Vlan1          active    0       internal
1007 Online Diagnostic Vlan2          active    0       internal
1008 Online Diagnostic Vlan3          active    0       internal
1009 Voice Internal Vlan              active    0       internal
1010 Dtp Vlan                         active    0       internal
1011 Private Vlan Reserved Vlan       suspend   0       internal
1016 Online SP-RP Ping Vlan           active    0       internal
```

Jak widać, zmodyfikowany został również VLAN 1, ponieważ część portów z VLAN 1 została przeniesiona do VLAN 10.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

IOS z wykorzystaniem bazy danych VLAN-ów

Ta metoda konfiguracji została zastosowana jako opcja z uwagi na jej kompletność. Starsze przełączniki umożliwiające jedynie taki tryb konfiguracji są bez wątpienia nadal stosowane. Przełączniki z systemem IOS obsługujące bazę danych VLAN-ów (np. 3750) wyświetlają następujący komunikat po wejściu w tryb konfiguracji bazy danych VLAN:

```
3750-IOS#vlan database
% Warning: It is recommended to configure VLAN from config mode,
  as VLAN database mode is being deprecated. Please consult user
  documentation for configuring VTP/VLAN in config mode. // Uwaga: Zaleca się konfigurację VLAN-ów
  // z poziomu trybu config mode, ponieważ tryb bazy danych VLAN-ów jest trybem przestarzałym. Informacje na temat
  // konfiguracji VTP/VLAN w config mode znajdziesz w dokumentacji użytkownika.
```

Jeśli masz przełącznik IOS z aktywnymi VLAN-ami, ale nie ma o tym informacji w bieżącej konfiguracji, możliwe, że zostały one utworzone w trybie bazy danych VLAN-ów. Inną możliwością jest, że powstały one poprzez zastosowanie VTP (omówię to w rozdziale 6.). W przypadku przełączników z serii 3750 ustawionych w trybie serwera VTP, nawet przy tworzeniu VLAN-ów za pomocą interfejsu CLI, nie pojawią się one w bieżącej konfiguracji.

Ponieważ bardziej prawdopodobne jest, że spotkasz się z bazą danych VLAN-ów w starszych modelach przełączników, będę kontynuował omawianie tego trybu konfiguracji na przykładzie urządzenia 2950. Oczywiście wszystkie takie przełączniki zachowują się podobnie. Jeśli napotkasz kiedyś switch skonfigurowany za pomocą bazy danych VLAN-ów, doradzałbym przekonfigurowanie go z poziomu IOS.

Aby skonfigurować wirtualne sieci lokalne w bazie danych VLAN-ów, musisz wejść do trybu konfiguracyjnego bazy danych VLAN-ów, stosując polecenie `vlan database`. Następnie możesz użyć polecenia żądania pomocy (?), aby wylistować dostępne w tym trybie komendy:

```
2950-IOS#vlan database
2950-IOS(vlan)# ?
VLAN database editing buffer manipulation commands:
  abort  Exit mode without applying the changes              //Wyjdź z danego trybu bez
                                                             // zapisywania zmian
  apply  Apply current changes and bump revision number      // Zapisz zmiany i podaj numer wersji
  exit   Apply changes, bump revision number, and exit mode  // Zapisz zmiany, podaj numer wersji
                                                             // i wyjdź z danego trybu
  no     Negate a command or set its defaults                // Neguje polecenie, aby wrócić
                                                                   // do wartości domyślnych
  reset  Abandon current changes and reread current database         // Zrezygnuj z wprowadzonych zmian
                                                                     // i przywróć bazę danych
  show   Show database information                                   // Pokaż informacje o bazie danych
  vlan   Add, delete, or modify values associated with a single VLAN // Dodaje, usuwa lub modyfikuje
                                                                     // wartości związane z pojedynczym
                                                                     // VLAN-em
  vtp    Perform VTP administrative functions                        // Funkcje administracyjne VTP
```

Aby utworzyć VLAN, zastosuj polecenie `vlan`, określając przy tym numer i nazwę VLAN-u:

```
2950-IOS(vlan)#vlan 10 name Lab-VLAN
VLAN 10 added:
    Name: Lab-VLAN
```

Polecenie `show` pozwala wyświetlić VLAN-y skonfigurowane w trybie bazy danych. Opcje tego polecenia służą do wyświetlenia informacji o bieżącej bazie danych (`show current`), różnicach między bieżącą a proponowaną bazą danych (`show changes`) lub o proponowanej bazie danych

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

w kształcie, jaki otrzyma po zapisaniu zmian poleceniem `apply` albo po wyjściu z trybu konfiguracji bazy danych VLAN-ów (komenda `exit`). Domyślnie użycie polecenia `show` traktowane jest jako `show proposed`:

```
2950-IOS(vlan)#show
  VLAN ISL Id: 1
    Name: default
    Media Type: Ethernet
    VLAN 802.10 Id: 100001
    State: Operational
    MTU: 1500
    Backup CRF Mode: Disabled
    Remote SPAN VLAN: No

  VLAN ISL Id: 10
    Name: Lab-VLAN
    Media Type: Ethernet
    VLAN 802.10 Id: 100010
    State: Operational
    MTU: 1500
    Backup CRF Mode: Disabled
    Remote SPAN VLAN: No
```

Do utworzenia podstawowego VLAN-u nie potrzeba niczego więcej. Zmiany w bazie danych zostaną zapisane po użyciu `exit`:

```
2950-IOS(vlan)#exit
APPLY completed.
Exiting....
```

Teraz, jeśli wpiszesz polecenie `show vlan` w interfejsie IOS, zobaczysz VLAN-y, które utworzyłeś:

```
2950-IOS#sho vlan
VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
1    default                          active    Fa0/1, Fa0/2, Fa0/3, Fa0/4
                                                Fa0/5, Fa0/6, Fa0/7, Fa0/8
                                                Fa0/9, Fa0/10, Fa0/11, Fa0/12
                                                Fa0/13, Fa0/14, Fa0/15, Fa0/16
                                                Fa0/17, Fa0/18, Fa0/19, Fa0/20
                                                Fa0/21, Fa0/22, Fa0/23, Fa0/24
                                                Gi0/1, Gi0/2
10   Lab-VLAN                         active
1002 fddi-default                     active
1003 token-ring-default               active
1004 fddinet-default                  active
1005 trnet-default                    active
```

Dodawanie portów do VLAN-ów odbywa się w trybie konfiguracji interfejsu IOS i jest opisane w następnej sekcji.

Standardowe polecenia IOS

Dodawanie sieci VLAN za pomocą interfejsu IOS jest stosunkowo łatwe, zakładając, że wszystkie domyślne ustawienia są dopuszczalne, co ma zazwyczaj miejsce. Powrócę tutaj do modelu 3750, ponieważ ta metoda spotykana jest z reguły w nowoczesnych przełącznikach.

Pierwszym krokiem jest wejście w tryb konfiguracji (ang. *configuration mode*). Teraz powinieneś użyć polecenia `vlan` wraz z identyfikatorem VLAN-u, który zamierzasz dodać lub zmieniać. Następnie możesz nadać VLAN-owi nazwę, stosując podkomendę `name` (podobnie jak

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

w CatOS, jeśli nie określisz nazwy, zostanie automatycznie przypisana nazwa domyślna w postaci `VLAN`*xxxx*):

```
3750-IOS#conf t
Enter configuration commands, one per line. End with CNTL/Z. // Wprowadź polecenia konfiguracyjne, jedna
                                                              // komenda w linijce. Zakończ CNTL/Z.
3750-IOS(config)# vlan 10
3750-IOS(config-vlan)# name Lab-VLAN
```

Wyjdź z trybu konfiguracyjnego (*Ctrl*+*Z*) i wprowadź polecenie `show vlan`. Zobaczysz, jakie VLAN-y są w tej chwili skonfigurowane na przełączniku:

```
3750-IOS#sho vlan
VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
1    default                          active    Gi1/0/1, Gi1/0/2, Gi1/0/3
                                                Gi1/0/4, Gi1/0/5, Gi1/0/6
                                                Gi1/0/7, Gi1/0/8, Gi1/0/9
                                                Gi1/0/10, Gi1/0/11, Gi1/0/12
                                                Gi1/0/13, Gi1/0/14, Gi1/0/15
                                                Gi1/0/16, Gi1/0/17, Gi1/0/18
                                                Gi1/0/21, Gi1/0/22, Gi1/0/23
                                                Gi1/0/24, Gi1/0/25, Gi1/0/26
                                                Gi1/0/27, Gi1/0/28, Gi1/0/29
                                                Gi1/0/30, Gi1/0/31, Gi1/0/32
                                                Gi1/0/33, Gi1/0/34, Gi1/0/35
                                                Gi1/0/36, Gi1/0/37, Gi1/0/38
                                                Gi1/0/39, Gi1/0/40, Gi1/0/41
                                                Gi1/0/42, Gi1/0/43, Gi1/0/44
                                                Gi1/0/46, Gi1/0/49, Gi1/0/50
                                                Gi1/0/51, Gi1/0/52
10   Lab-VLAN                         active
100  VLAN0100                         active
200  VLAN0200                         active
300  VLAN0300                         active
VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
1002 fddi-default                     act/unsup
1003 token-ring-default               act/unsup
1004 fddinet-default                  act/unsup
1005 trnet-default                    act/unsup
VLAN Type  SAID       MTU   Parent RingNo BridgeNo Stp  BrdgMode Trans1 Trans2
---- ----- ---------- ----- ------ ------ -------- ---- -------- ------ ------
1    enet  100001     1500  -      -      -        -    -        0      0
10   enet  100010     1500  -      -      -        -    -        0      0
100  enet  100100     1500  -      -      -        -    -        0      0
200  enet  100200     1500  -      -      -        -    -        0      0
300  enet  100300     1500  -      -      -        -    -        0      0
1002 fddi  101002     1500  -      -      -        -    -        0      0
1003 tr    101003     1500  -      -      -        -    -        0      0
1004 fdnet 101004     1500  -      -      -        ieee -        0      0
1005 trnet 101005     1500  -      -      -        ibm  -        0      0
Remote SPAN VLANs
------------------------------------------------------------------------------
Primary Secondary Type              Ports
------- --------- ----------------- ------------------------------------------
```

W IOS konkretne porty są przypisywane do VLAN-ów w trybie konfiguracji interfejsów. Każdy interfejs musisz konfigurować indywidualnie, używając polecenia `switchport access` (w CatOS wygląda to nieco inaczej: wszystkie wybrane porty można dodać za jednym razem za pomocą polecenia `set vlan`):

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
3750-IOS(config)#int g1/0/1
3750-IOS(config-if)#switchport access vlan 10
3750-IOS(config-if)#int g1/0/2
3750-IOS(config-if)#switchport access vlan 10
```

Nowsze wersje IOS umożliwiają zastosowanie określonych poleceń do wielu interfejsów jednocześnie za pomocą komendy `interface range`. Pozwala to osiągnąć ten sam rezultat, zmniejszając jednocześnie liczbę wprowadzanych komend:

```
3750-IOS(config)#interface range g1/0/1 - 2
3750-IOS(config-if-range)#switchport access vlan 10
```

Wprowadzając teraz polecenie `show vlan`, zobaczysz, że wybrane porty zostały przypisane do właściwych VLAN-ów:

```
3750-IOS#sho vlan
VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
1    default                          active    Gi1/0/3, Gi1/0/4, Gi1/0/5
                                                Gi1/0/6, Gi1/0/7, Gi1/0/8
                                                Gi1/0/9, Gi1/0/10, Gi1/0/11
                                                Gi1/0/12, Gi1/0/13, Gi1/0/14
                                                Gi1/0/15, Gi1/0/16, Gi1/0/17
                                                Gi1/0/18, Gi1/0/21, Gi1/0/22
                                                Gi1/0/23, Gi1/0/24, Gi1/0/25
                                                Gi1/0/26, Gi1/0/27, Gi1/0/28
                                                Gi1/0/29, Gi1/0/30, Gi1/0/31
                                                Gi1/0/32, Gi1/0/33, Gi1/0/34
                                                Gi1/0/35, Gi1/0/36, Gi1/0/37
                                                Gi1/0/38, Gi1/0/39, Gi1/0/40
                                                Gi1/0/41, Gi1/0/42, Gi1/0/43
                                                Gi1/0/44, Gi1/0/46, Gi1/0/49
                                                Gi1/0/50, Gi1/0/51, Gi1/0/52
10   Lab-VLAN                         active    Gi1/0/1, Gi1/0/2
100  VLAN0100                         active
200  VLAN0200                         active
300  VLAN0300                         active
1002 fddi-default                     act/unsup
```

Nexus i NX-OS

NX-OS posiada zbliżony interfejs poleceń do tego, który znasz już z IOS. Zachowuje się on jednak nieco inaczej, zwłaszcza w odniesieniu do konfiguracji interfejsów. Metoda konfiguracji VLAN-ów jest podobna jak w IOS. Na początek utwórz VLAN poleceniem `vlan` *`nr_vlan`*:

```
NX-7K-1-Cozy(config)# vlan 10
```

Kiedy VLAN jest już utworzony, wejdź w tryb konfiguracji VLAN-ów i nazwij go poleceniem `name` *`nazwa-vlan`*:

```
NX-7K-1-Cozy(config-vlan)# name Lab-VLAN
```

Jedną z ciekawszych funkcji NX-OS jest możliwość egzekwowania komend typu „show" bezpośrednio w trybie konfiguracji, bez stosowania dodatkowych poleceń. Jest to zachowanie zbliżone do wdrożonego w trybie konfiguracji urządzeń PIX oraz ASA i stanowi to bardzo pożądaną zmianę. Poniżej przykład, jak wykonać polecenie `show vlan` w trybie konfiguracji VLAN-ów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-7K-1-Cozy(config-if)# sho vlan
VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
1    default                          active    Po1, Po10
10   Lab-VLAN                         active    Po1, Eth3/2
VLAN Type
---- -----
1    enet
10   enet
Remote SPAN VLANs
-------------------------------------------------------------------------------
Primary  Secondary  Type             Ports
-------  ---------  ---------------  ------------------------------------------
```

Kolejną nowością w NX-OS jest możliwość konfigurowania wielu interfejsów bez konieczności stosowania polecenia `interface-range`. Po prostu wpisz zakres interfejsów, które chcesz konfigurować, tak jakbyś stosował komendę `interface-range` w IOS:

```
NX-7K-1-Cozy(config-vlan)# int e3/1 – 2
```

Automatycznie przechodzisz wtedy do trybu konfiguracji zakresu interfejsów. Teraz możesz przydzielać porty do VLAN-ów tak jak w IOS, stosując komendę `switchport access vlan` *`nr_vlan`*:

```
NX-7K-1-Cozy(config-if-range)# switchport access vlan 10
Warning: command rejected, Eth3/1 not a switching port
Warning: command rejected, Eth3/2 not a switching port
```

W tym momencie pojawiło się ostrzeżenie, którego standardowo nie zobaczysz w przypadku przełączników Catalyst. Przełącznik Nexus 7000 zachowuje się jednak inaczej niż switche Catalyst. Domyślnie wszystkie porty switcha są portami routera! Aby wykonać polecenia dotyczące zmiany portów w przypadku portów Nexusa, musisz najpierw zastosować dla nich tryb przełączania portów (ang. *switchport mode*), wpisując komendę `switchport`:

```
NX-7K-1-Cozy(config-if-range)# int e3/1 - 2
NX-7K-1-Cozy(config-if-range)# switchport
NX-7K-1-Cozy(config-if-range)# no shut
```

Kiedy już włączyłeś dla wybranych portów tryb switchport, możesz bez przeszkód przypisywać je do konkretnych VLAN-ów.

```
NX-7K-1-Cozy(config-if-range)# switchport access vlan 10
NX-7K-1-Cozy(config-if-range)#
```

Polecenie `show vlan` pokaże teraz Twoje porty przypisane do VLAN-u 10. Zwróć jeszcze raz uwagę, że komendę `show` wprowadzasz w trybie konfiguracji. Kocham tę funkcję!

```
NX-7K-1-Cozy(config-if-range)# sho vlan
VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
1    default                          active
10   Lab-VLAN                         active    Eth3/1, Eth3/2
VLAN Type
---- -----
1    enet
10   enet
Remote SPAN VLANs
-------------------------------------------------------------------------------
Primary  Secondary  Type             Ports
-------  ---------  ---------------  ------------------------------------------
```

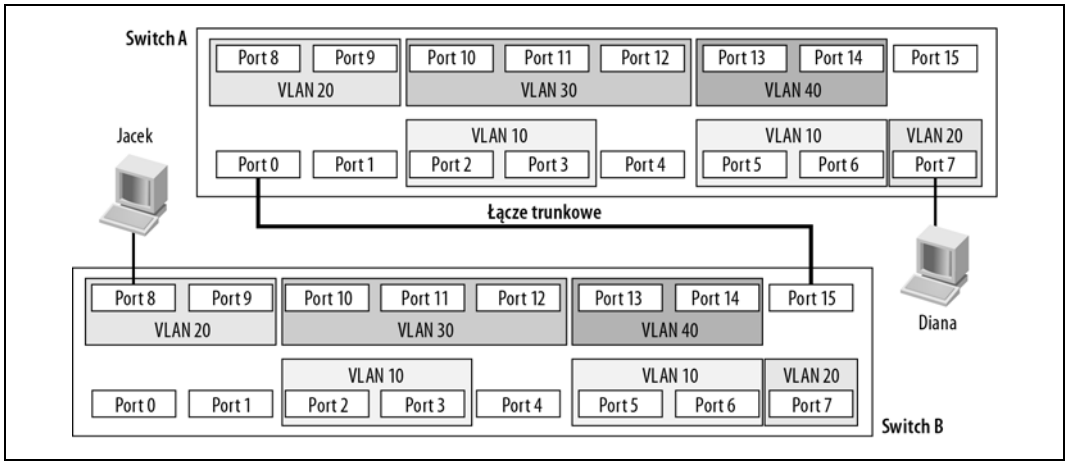

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 5.

Trunking

W terminologii Cisco **kanał trunkowy** (ang. *trunk*) to interfejs lub łącze, które umożliwia transmisję ramek dla wielu VLAN-ów jednocześnie. Jak dowiedziałeś się z poprzedniego rozdziału, możesz połączyć trunkiem dwa przełączniki, tak aby urządzenia znajdujące się w VLAN-ie na jednym z nich mogły się komunikować z urządzeniami w tym samym VLAN-ie na drugim przełączniku. Jeśli posiadamy wiele VLAN-ów, to łączymy switche w warstwie drugiej poprzez kanał trunkowy. Rysunek 5.1 przedstawia dwa switche połączone trunkiem.



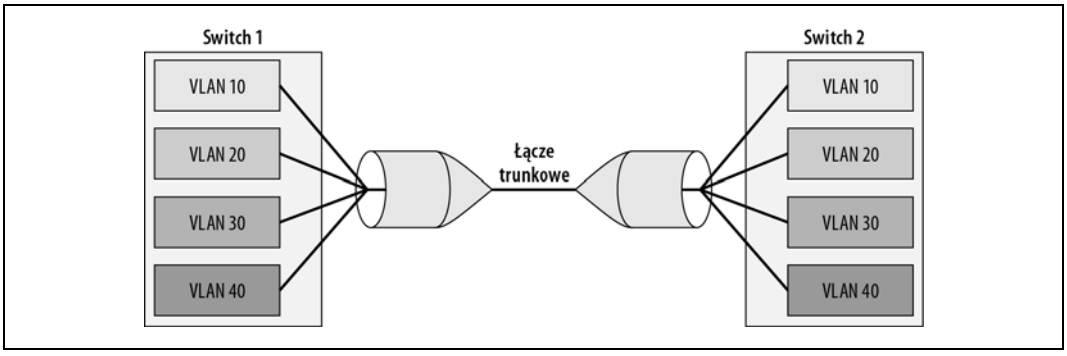
Rysunek 5.1. Kanał trunkowy łączący dwa przełączniki

Zasadniczo pojęcie trunkingu odnosi się do przełączników, jednak i inne urządzenia, jak chociażby routery czy firewalle, mogą być łączone trunkami. Schemat routera na patyku opisany w rozdziale 4. przedstawia router połączony z portem trunkowym przełącznika.

Jak działają kanały trunkowe

Rysunek 5.2 ilustruje koncepcję łącza trunkowego. VLAN-y 10, 20, 30 i 40 znajdują się po obu stronach kanału. Każda transmisja z VLAN-u 10 na przełączniku 1 skierowana do VLAN-u 10 na przełączniku 2 musi przechodzić przez trunk (oczywiście dzieje się tak samo również w drugim kierunku).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 5.2. Graficzne przedstawienie kanału trunkowego

Aby zdalny przełącznik mógł określić, w jaki sposób przesłać daną ramkę, ramka ta musi zawierać odniesienie do VLAN-u, do którego przynależy. Pakiety IP nie zawierają żadnej informacji dotyczącej VLAN-ów, podobnie jak pakiety protokołów TCP, UDP, ICMP czy jakichkolwiek innych protokołów warstwy wyższej niż druga. Pamiętaj, że koncepcja sieci wirtualnych umieszcza je w warstwie drugiej łącza danych, więc jakiekolwiek odniesienia do VLAN-ów powinny się zawierać w tej warstwie. Ethernet został wynaleziony, zanim wymyślono VLAN-y, tak więc także w protokołach Ethernet nie ma do nich odniesienia.

W celach znakowania czy **tagowania** (ang. *tag*) ramki, która ma zostać przesłana kanałem trunkowym, obie strony muszą uzgodnić stosowany protokół. Cisco wykorzystuje obecnie dla celów trunkingowych protokoły ISL (ang. *Inter-Switch Link*) oraz IEEE 802.1Q. Nie wszystkie przełączniki Cisco obsługują jednak obydwa protokoły równolegle. Przykładowo modele z serii 2950 oraz 4000 wspierają jedynie 802.1Q. Aby sprawdzić, który z protokołów trunkingowych jest wspierany przez dany przełącznik, możesz użyć polecenia `show interface capabi ↪lities` (dla IOS oraz NX-OS) lub `show port capabilities` (dla CatOS):

```
NX-7K-1-Cozy(config)# sho interface capabilities | inc Trunk
Ethernet3/1
  Trunk encap. type:     802.1Q
```

Jak widać z powyższego przykładu, przełączniki Nexus obsługują jedynie otwarty standard 802.1Q. Poniżej switch 3750 obsługujący zarówno 802.1Q, jak i ISL:

```
3750-IOS#sho int g1/0/45 capabilities | inc Trunk
  Trunk encap. type:     802.1Q,ISL
  Trunk mode:            on,off,desirable,nonegotiate
```

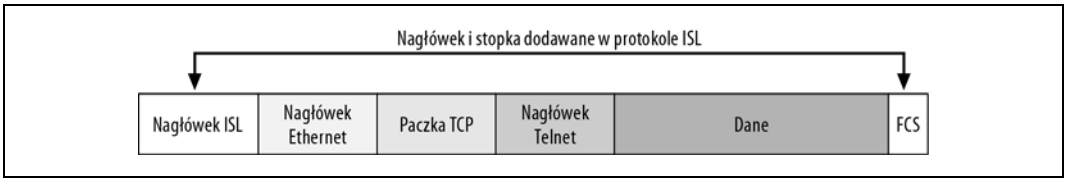
ISL różni się od 802.1Q pod kilkoma względami. Przede wszystkim ISL to protokół własny, opracowany przez firmę Cisco, natomiast 802.1Q jest standardem IEEE. Po drugie, ISL enkapsuluje ramki Ethernet w ramki ISL, podczas gdy 802.1Q dodaje do ramek znaczniki VLAN. Ponadto ISL jest w stanie obsłużyć jedynie 1000 VLAN-ów, a 802.1Q aż 4096.

W przełącznikach obsługujących zarówno ISL, jak i 802.1Q obydwa protokoły mogą być wykorzystywane. Protokół jest konfigurowany dla każdego kanału trunkowego indywidualnie. Możesz konfigurować równolegle trunki ISL oraz 802.1Q na tym samym przełączniku i w tej samej sieci, pod warunkiem, że każdy kanał z osobna wykorzystuje po obu stronach tylko jeden wybrany protokół.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ISL

Aby do ramki wysyłanej kanałem trunkowym dodać informację o VLAN-ie, ISL enkapsuluje całą ramkę w nową ramkę. Dołączany jest do ramki dodatkowy nagłówek oraz niewielki przyrostek na końcu. Nagłówek zawiera głównie informację dotyczącą numeru VLAN-u, podczas gdy stopka niesie wiadomość o sumie kontrolnej ramki. Rysunek 5.3 ilustruje dokładny schemat ramki ISL.



Rysunek 5.3. Ramka ISL po enkapsulacji

Stopka sekwencji kontrolnej ramki FCS (ang. *frame check sequence*) jest dodatkowa w stosunku do pola FCS, które oryginalnie zawiera ramka Ethernet. Ramka ISL ze stopką FCS oblicza sumę kontrolną z uwzględnieniem nagłówka ISL. Suma kontrolna FCS ramki Ethernet nie zawiera tego nagłówka.

Dołączanie dodatkowej informacji do ramki Ethernet może być problematyczne. Jeśli mamy do czynienia z ramką o maksymalnym rozmiarze 1518 bajtów, to protokół ISL, dodając kolejne 30 bajtów, powiększy wielkość ramki do 1548 bajtów. Takie ramki mogą być postrzegane jako zbyt duże, co prowadzi do występowania błędów ramek typu „giant". Mimo to przełączniki Cisco nie mają z nimi problemów.

802.1Q

Standard 802.1Q reprezentuje odmienne podejście do tagowania VLAN-ów. Zamiast dodawać kolejne nagłówki do ramki, 802.1Q wstawia informację do istniejących nagłówków. Dodatkowe czterobajtowe pole znacznika wstawiane jest pomiędzy pola „adres źródłowy" a „typ/długość". Ponieważ ramka jest modyfikowana, FCS również odzwierciedli tę zmianę.

Skoro dodane zostaną jedynie 4 bajty, to maksymalny rozmiar ramki 802.1Q może wynieść 1522 bajty. Może to skutkować występowaniem błędów ramek typu „baby giant". Ramki takie są jednak nadal obsługiwane przez urządzenia Cisco.

Który protokół wybrać

Właściwie czemu istnieją dwa protokoły? Cisco wprowadziło protokół ISL, zanim w ogóle powstał 802.1Q. Starsze przełączniki Cisco obsługują więc tylko ISL. Co dziwne, inne przełączniki, jak Nexus, wspierają tylko 802.1Q. Niekiedy, jak w przypadku systemu przełączników kasetowych (ang. *blade*), to magistrala obudowy może być decydującym czynnikiem. Przykładowo 10-gigabitowy moduł *blade* do przełącznika Catalyst 6509 obsługuje jedynie protokół 802.1Q, podczas gdy sam przełącznik wspiera oba protokoły. W takim układzie musisz zastosować 802.1Q.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W większości instalacji możesz użyć dowolnie jednego z tych dwóch protokołów, a Twój wybór nie będzie miał znaczenia. Jeśli łączysz trunkiem dwa switche Cisco, żaden z protokołów nie ma przewagi nad drugim, z tym tylko wyjątkiem, że 802.1Q obsługuje 4096 VLAN-ów, podczas gdy ISL jedynie 1000. Jedni opowiedzą się zawsze za protokołem ISL, ponieważ nie modyfikuje on oryginalnej ramki, inni preferują 802.1Q, gdyż ramki są mniejszych rozmiarów i nie stosuje się enkapsulowania. W efekcie każdy instalator wybiera taki protokół, do jakiego się przyzwyczaił.

Cisco opublikowało rekomendację dotyczącą konfiguracji kanałów trunkowych pomiędzy przełącznikami Cisco. Dokument ten (dokument Cisco ID 24067) nosi tytuł *System Requirements to Implement Trunking* („Wymagania systemowe dla implementacji trunkingu").

Jeśli łączysz przełączniki Cisco z urządzeniami innej firmy, wyborem jest zawsze 802.1Q. Pamiętaj, że nie ma restrykcji dotyczących wyboru protokołu w przypadku przełączników obsługujących obydwa protokoły. Gdy zajdzie potrzeba podłączenia przełącznika 3Com do sieci opartej w całości na urządzeniach Cisco, można zastosować trunk 802.1Q, nawet jeśli pozostałe trunki w sieci oparte są na protokole ISL. Protokół trunkingowy działa lokalnie dla każdego indywidualnego kanału. VLAN-y na switchu 3Com stosującym 802.1Q będą wciąż dostępne na pozostałych przełącznikach, podłączonych w innych miejscach trunkiem ISL.

Linia przełączników Cisco Nexus, w której zastosowano więcej otwartych standardów niż w starszych urządzeniach, nie obsługuje protokołu ISL.

Negocjacja trunków

Niektóre przełączniki Cisco obsługują protokół DTP (ang. *Dynamic Trunking Protocol*). Odpowiada on za określenie, jakie protokoły są wspierane po obu stronach łącza, i ustanowienie, jeśli to możliwe, kanału trunkowego.

W procesie negocjacji trunków wykorzystywana jest domena protokołu VTP (ang. *VLAN Trunking Protocol*). Aby negocjacja DTP zakończyła się powodzeniem, oba przełączniki muszą mieć taką samą (wspólną) nazwę domeny VTP. Więcej szczegółów na temat konfiguracji domen VTP znajdziesz w rozdziale 6.

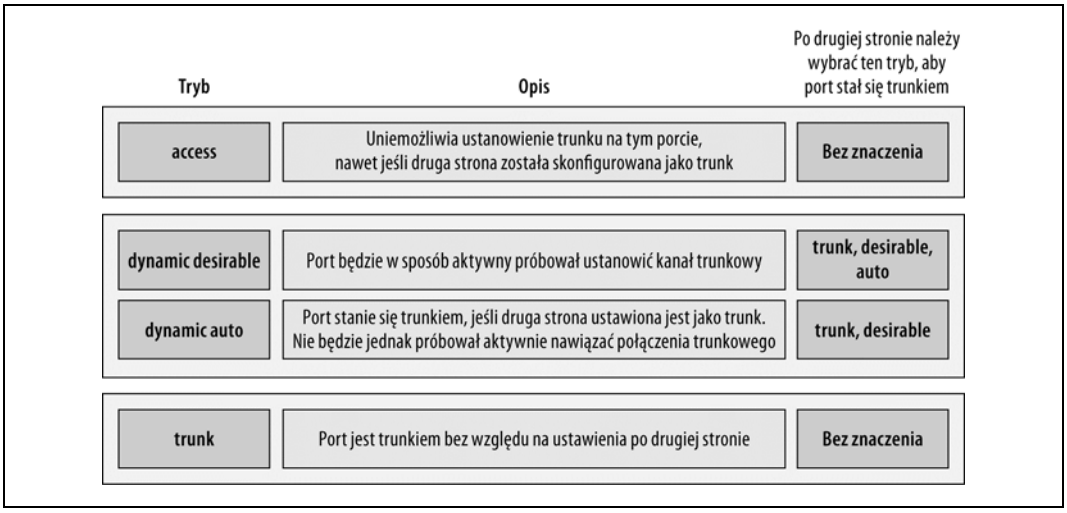
Interfejs, na którym włączono DTP, wysyła co 30 sekund ramki w celu wynegocjowania kanału trunkowego. Jeśli jednak na porcie ręcznie ustawiono opcję *trunk* (kanał trunkowy) lub *prevent trunking* (nie stosuj kanału trunkowego), protokół DTP jest niepotrzebny i można go wyłączyć. Dla IOS poleceniem wyłączającym DTP jest `switchport nonegotiate`:

```
3750-IOS(config-if)#switchport mode trunk
3750-IOS(config-if)#switchport nonegotiate
```

Wielu inżynierów stosuje ze względów bezpieczeństwa polecenia `switchport none ↪gotiate` dla wszystkich portów przełączników. Ma to swoje uzasadnienie, gdyż nieupoważniona osoba mogłaby podłączyć się przełącznikiem do jednego z portów i uzyskać połączenie trunkowe. Miałaby wtedy potencjalny dostęp do wszystkich VLAN-ów na switchu. Ja preferuję metodę domyślnego wyłączenia wszystkich portów, tak aby podłączenie dodatkowego urządzenia do sieci wymagało interwencji administratora. Oba podejścia mają swoje zalety i mogą być stosowane równolegle.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Rysunek 5.4 przedstawia możliwe tryby polecenia `switchport`. Pamiętaj, że nie każdy przełącznik obsługuje wszystkie tryby. Dany port może być ustawiony w trybie `access`, co oznacza, że nigdy nie będzie trunkiem; `dynamic`, co znaczy, że port może stać się trunkiem; lub `trunk`, co oznacza, iż port będzie trunkiem bez względu na wszelkie inne ustawienia.



Rysunek 5.4. Możliwe tryby switchport związane z funkcją trunkingu

Dwa warianty trybu `dynamic`, czyli `desirable` i `auto`, odnoszą się do sytuacji, w której na porcie włączony jest protokół DTP. W przypadku `desirable` port będzie inicjował proces negocjacji i próbował ustanowić kanał trunkowy na łączu. Przy ustawieniu `auto` port będzie nasłuchiwał ramek DTP, ale nie będzie próbował w sposób czynny utworzyć trunku.

Domyślnym trybem dla większości przełączników Cisco jest `dynamic auto`. Przy tej konfiguracji port automatycznie stanie się trunkiem, jeśli podłączony do niego port na drugim przełączniku jest ustawiony na „sztywno" jako trunk lub w trybie `dynamic desirable`.

Konfiguracja trunków

Konfigurowanie kanałów trunkowych wiąże się z koniecznością wybrania portu, który będzie trunkiem, protokołu, w jakim trunk będzie działał, oraz tego, czy i w jaki sposób port będzie negocjował. Opcjonalnie możesz również zdefiniować, które VLAN-y będą dopuszczane do transmisji na danym łączu trunkowym.

IOS

Doskonałym przykładem przełącznika z systemem operacyjnym IOS jest Cisco 3750. W tej sekcji przyjrzymy się konfiguracji gigabitowych portów jako trunków w standardzie 802.1Q na przykładzie tego właśnie switcha.

Mógłbyś pomyśleć, że pierwszym krokiem jest zdefiniowanie określonego portu jako trunku. Jednak w tym przypadku jest inaczej:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
3750-IOS(config-if)#switchport mode trunk
Command rejected: An interface whose trunk encapsulation is "Auto" can not be configured to "trunk" mode.
// Komenda odrzucona: interfejs z ustawionym enkapsulowaniem „Auto" nie może zostać skonfigurowany w trybie „trunk".
```

Dla IOS, który obsługuje zarówno ISL, jak i 802.1Q, musisz najpierw określić sposób **enkapsulowania trunku**, zanim będziesz mógł skonfigurować dany port jako trunk. Stosuje się w tym celu polecenie `trunk encapsulation`. Nie jest to do końca zręczny wybór nazwy polecenia, gdyż jak wiesz 802.1Q nie enkapsuluje ramek w sposób, w jaki robi to ISL. Trzeba jednak trzymać się składni Cisco. Kiedy już wybierzesz protokół trunkingowy, możesz zadeklarować, który port ma być trunkiem.

```
3750-IOS(config)#int g1/0/45
3750-IOS(config-if)#switchport trunk encapsulation dot1q
3750-IOS(config-if)#switchport mode trunk
```

Jeśli chciałbyś usunąć utworzony trunk z interfejsu, możesz użyć komendy `switchport mode access`.

W domyślnej konfiguracji wszystkie VLAN-y są dopuszczone do przesyłania danych przez istniejące kanały trunkowe. Możesz mieć jednak 40 VLAN-ów, a chcieć przesyłać trunkiem jedynie 3 z nich. Ponieważ wszystkie dopuszczone VLAN-y będą transmitować pakiety przez każde łącze trunkowe, wykluczenie z tego procesu niepotrzebnych VLAN-ów może zaoszczędzić sporo pasma na tych łączach. Aby zdefiniować, które VLAN-y mogą przechodzić przez trunki, użyj polecenia `switchport trunk allowed`. Poniżej dostępne opcje tego polecenia:

```
3750-IOS(config-if)#switchport trunk allowed vlan ?
  WORD    VLAN IDs of the allowed VLANs when this port is in trunking mode
                                        // ID VLAN dla dozwolonego VLAN-u, jeśli port jest w trybie trunking
  add     add VLANs to the current list  // dodaje VLAN-y do bieżącej listy
  all     all VLANs // wszystkie VLAN-y są dozwolone
  except  all VLANs except the following // dopuszczone są wszystkie VLAN-y oprócz zdefiniowanego
  none    no VLANs // żaden VLAN nie jest dozwolony
  remove  remove VLANs from the current list // usuwa VLAN-y z bieżącej listy
```

Aby dopuścić do transmisji w danym trunku tylko jeden VLAN (w tym przypadku VLAN 100), zastosuj poniższe polecenie:

```
3750-IOS(config-if)#switchport trunk allowed vlan 100
```

Teraz po wpisaniu komendy `show interface trunk` zobaczysz, że tylko VLAN 100 jest akceptowany. IOS usunął z trunku wszystkie pozostałe:

```
3750-IOS#sho int trunk
Port        Mode             Encapsulation  Status        Native vlan
Gi1/0/45    on               802.1q         trunking      1
Port        Vlans allowed on trunk
Gi1/0/45    100
Port        Vlans allowed and active in management domain
Gi1/0/45    100
Port        Vlans in spanning tree forwarding state and not pruned
Gi1/0/45    100
```

Jeśli jednak chciałbyś zezwolić na transmisję ze wszystkich VLAN-ów z wyjątkiem VLAN-u 100, skorzystaj z poniższego przykładu:

```
3750-IOS(config-if)#switchport trunk allowed vlan except 100
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ta komenda nadpisze poprzednią konfigurację, która ustanawiała VLAN 100 jako jedyny uprawniony do transmisji. Teraz wszystkie VLAN-y, **z wyjątkiem** VLAN-u 100, będą uprawnione. Poprzedni stan mógłbyś przywrócić, stosując polecenie `switchport trunk allowed vlan 100`. Obecny status odczytasz, wpisując `show interface trunk`:

```
3750-IOS#sho int trunk
Port        Mode             Encapsulation  Status        Native vlan
Gi1/0/45    on               802.1q         trunking      1
Port        Vlans allowed on trunk
Gi1/0/45    1-99,101-4094
Port        Vlans allowed and active in management domain
Gi1/0/45    1
Port        Vlans in spanning tree forwarding state and not pruned
Gi1/0/45    none
```

VLAN-y 1 – 99 i 101 – 4096 są obecnie dopuszczone na trunku. Powiedzmy, że chciałbyś teraz z tej listy usunąć VLAN-y 200 i 300. W tym celu użyj opcji `remove`:

```
3750-IOS(config-if)#switchport trunk allowed vlan remove 200
3750-IOS(config-if)#switchport trunk allowed vlan remove 300
```

Informacja wynikowa polecenia `show interface trunk` pokazuje, że w tej chwili na trunku dopuszczone są wszystkie VLAN-y z wyjątkiem 100, 200 i 300:

```
3750-IOS#sho int trunk
Port        Mode             Encapsulation  Status        Native vlan
Gi1/0/45    on               802.1q         trunking      1
Port        Vlans allowed on trunk
Gi1/0/45    1-99,101-199,201-299,301-4094
Port        Vlans allowed and active in management domain
Gi1/0/45    1
Port        Vlans in spanning tree forwarding state and not pruned
Gi1/0/45    1
```

CatOS

W CatOS do konfiguracji trunku używa się komendy `set trunk`. Dostępne są następujące opcje:

```
Switch1-CatOS# (enable)set trunk 3/1 ?
  none                      No vlans
  <mode>                    Trunk mode (on,off,desirable,auto,nonegotiate)
  <type>                    Trunk type (isl,dot1q,dot10,lane,negotiate)
  <vlan>                    VLAN number
```

W trybie `on` port jest skonfigurowany jako trunk, natomiast tryb `off` oznacza, że dany port nigdy nie będzie wykorzystany do kanału trunkowego. Opcje `desirable` oraz `auto` odnoszą się do sposobu działania protokołu DTP na danym porcie. W przypadku `desirable` port będzie inicjował negocjację, starając się ustanowić kanał trunkowy. W trybie `auto` interfejs będzie nasłuchiwał ramek DTP, ale nie będzie starał się w sposób czynny utworzyć trunku. Tryb `nonego ↪tiate` wyłącza protokół DTP. Możesz go zastosować, gdy na przeciwległym porcie drugiego przełącznika skonfigurowana jest opcja `on` lub `off`.

Typy trunku `isl` oraz `dotq1` dotyczą odpowiednio wyboru protokołu ISL i 802.1Q. Opcja `nego ↪tiate` włącza DTP w celu negocjacji protokołu. Parametry `dot10` i `lane` dotyczą technologii takich jak ATM, których nie będę tu omawiał.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jedną z przyjemniejszych cech CatOS jest to, że umożliwia on określanie kilku parametrów jednocześnie przy wpisywaniu danej komendy. Poniższe polecenie konfiguruje tryb portu desi ↪rable i ustala zarazem protokół trunkingowy 802.1Q:

```
Switch1-CatOS# (enable)set trunk 3/5 desirable dot1q
Port(s)  3/1 trunk mode set to desirable.
Port(s)  3/1 trunk type set to dot1q.
Switch1-CatOS# (enable)
2006 May 23 11:29:31 %ETHC-5-PORTFROMSTP:Port 3/5 left bridge port 3/5
2006 May 23 11:29:34 %DTP-5-TRUNKPORTON:Port 3/5 has become dot1q trunk
```

Druga strona łącza nie została skonfigurowana, ale kanał trunkowy stał się aktywny. Jest to możliwe, ponieważ domyślne ustawienia przeciwległego portu to tryb `auto`.

Jeśli zechcesz podejrzeć status trunków w CatOS, wprowadź polecenie `show port trunk`:

```
Switch1-CatOS#sho port trunk
* - indicates vtp domain mismatch
# - indicates dot1q-all-tagged enabled on the port
$ - indicates non-default dot1q-ethertype value
Port      Mode         Encapsulation  Status        Native vlan
--------  -----------  -------------  ------------  -----------
 3/5      desirable    dot1q          trunking      1
15/1      nonegotiate  isl            trunking      1
16/1      nonegotiate  isl            trunking      1
Port      Vlans allowed on trunk
--------  ---------------------------------------------------------------------
 3/5      1-4094
15/1      1-4094
16/1      1-4094
Port      Vlans allowed and active in management domain
--------  ---------------------------------------------------------------------
 3/5      1,10,20
15/1
16/1
Port      Vlans in spanning tree forwarding state and not pruned
--------  ---------------------------------------------------------------------
 3/5      1,10,20
15/1
16/1
```

Trunki 15/1 i 16/1 pokazane na listingu to trunki wewnętrzne. W przełącznikach 6500 z systemem CatOS istnieją kanały trunkowe pomiędzy modułami zarządzającymi a kartami MSFC (ang. *Multilayer Switch Feature Card*). Gdy zainstalowane są dwa moduły zarządzające, karty MSFC rozpoznawane są jako sloty 15 i 16.

Aby określić, które VLAN-y mogą transmitować przez trunki, użyj tej samej komendy `set trunk`, dołączając identyfikatory VLAN-ów otrzymujących tym samym uprawnienia. W przypadku CatOS, w przeciwieństwie do IOS, użycie tego polecenia nie usunie z listy uprawnionych tych VLAN-ów, które były dodane wcześniej:

```
Switch-2# (enable)set trunk 3/5 100
Vlan(s) 100 already allowed on the trunk
Please use the 'clear trunk' command to remove vlans from allowed list.
```

Pamiętaj, że domyślnie wszystkie VLAN-y mają uprawnienia do transmisji przez trunki. Usunięcie wybranego VLAN-u odbywa się za pomocą polecenia `clear trunk`:

```
Switch-2# (enable)clear trunk 3/5 100
Removing Vlan(s) 100 from allowed list.
Port  3/5 allowed vlans modified to 1-99,101-4094.
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli chcesz podejrzeć zaktualizowaną listę dopuszczonych VLAN-ów, nie musisz stosować polecenia `show trunk`. Komenda `clear trunk` wyświetla również nowy status interfejsu.

Aby skonfigurować CatOS tak, żeby przepuszczał tylko jeden wybrany VLAN, musisz odebrać uprawnienia wszystkim pozostałym. Tak jak usuwałeś pojedynczy VLAN, użyj komendy `clear trunk` do usunięcia wszystkich VLAN-ów **z wyjątkiem tego**, któremu chcesz pozostawić uprawnienia:

```
Switch-2# (enable)clear trunk 3/5 1-99,101-4094
Removing Vlan(s) 1-99,101-4094 from allowed list.
Port  3/5 allowed vlans modified to 100.
```

Na koniec polecenie `show trunk` wyświetli aktualny status trunków. Jak widzisz, obecnie tylko VLAN 100 jest dozwolony na trunku 3/5:

```
Switch-2# (enable)sho trunk
* - indicates vtp domain mismatch
# - indicates dot1q-all-tagged enabled on the port
$ - indicates non-default dot1q-ethertype value
Port      Mode         Encapsulation  Status        Native vlan
--------  -----------  -------------  ------------  -----------
 3/5      auto         dot1q          trunking      1
15/1      nonegotiate  isl            trunking      1
16/1      nonegotiate  isl            trunking      1
Port      Vlans allowed on trunk
--------  ---------------------------------------------------------------------
 3/5      100
15/1      1-4094
16/1      1-4094
Port      Vlans allowed and active in management domain
--------  ---------------------------------------------------------------------
 3/5
15/1
16/1
Port      Vlans in spanning tree forwarding state and not pruned
--------  ---------------------------------------------------------------------
 3/5
15/1
16/1
```

Nexus i NX-OS

Pamiętając, że w przypadku Nexusa 7000 NX-OS ustawia domyślnie wszystkie interfejsy w trybie rutowania (ang. *routed mode*), musisz najpierw skorzystać z komendy `switchport interface`:

```
NX-7K-1-Cozy(config-if)# int e3/2
NX-7K-1-Cozy(config-if)# switchport
```

Teraz ustawiasz port w trybie `trunk` poleceniem `switchport mode trunk` dla danego interfejsu:

```
NX-7K-1-Cozy(config-if)# switchport mode ?
  access  Port mode access
  trunk   Port mode trunk
NX-7K-1-Cozy(config-if)# int e3/2
NX-7K-1-Cozy(config-if)# switchport mode trunk
```

Listę VLAN-ów uprawnionych do transmisji na danym trunku konfigurujesz poleceniem interfejsu `switchport trunk allowed vlan`:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-7K-1-Cozy(config-if)# switchport trunk allowed vlan ?
  <1-3967,4048-4093>  VLAN IDs of the allowed VLANs when this port in trunking mode
                                  // ID VLAN dla dozwolonego VLAN-u, jeśli port jest w trybie trunking
  add                 Add VLANs to the current list // dodaje VLAN-y do bieżącej listy
  all                 All VLANs // wszystkie VLAN-y są dozwolone
  except              All VLANs except the following
                                  // dopuszczone są wszystkie VLAN-y oprócz zdefiniowanego
  none                No VLANs  // żaden VLAN nie jest dozwolony
  remove              Remove VLANs from the current list // usuwa VLAN-y z bieżącej listy
```

Działa to podobnie jak w przypadku wersji komendy w IOS. Jeśli ustawisz uprawnienia dla VLAN-u 100, wszystkie pozostałe VLAN-y zostaną zablokowane na trunku.

```
NX-7K-1-Cozy(config-if)# switchport trunk allowed vlan 100
```

Sprawdzisz, które VLAN-y zostały dopuszczone na trunku, poleceniem `show interface trunk`. Listing różni się nieco od tego z IOS, ale zawiera wszelkie niezbędne informacje.

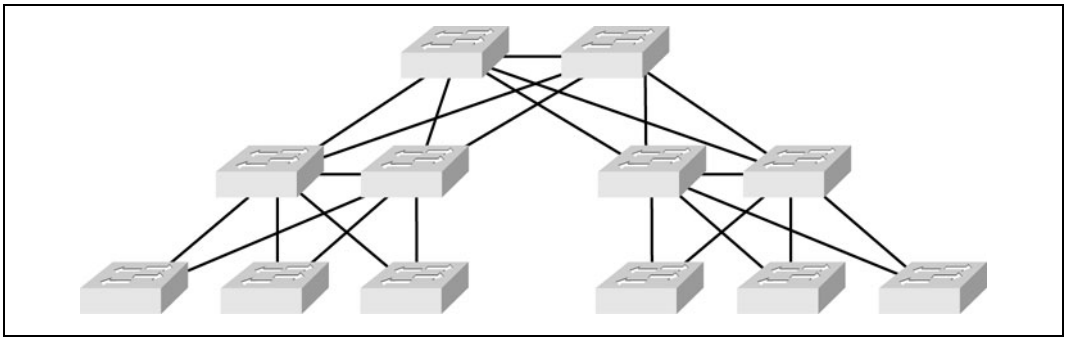
```
NX-7K-1-Cozy(config-if)# sho int trunk
--------------------------------------------------------------------------------
Port       Native  Status        Port
           Vlan                  Channel
--------------------------------------------------------------------------------
Eth3/2     1       trunking      --
Eth8/1     1       trnk-bndl     Po1
Eth8/2     1       trnk-bndl     Po1
Po1        1       trunking      --
--------------------------------------------------------------------------------
Port       Vlans Allowed on Trunk
--------------------------------------------------------------------------------
Eth3/2     100
Eth8/1     1-3967,4048-4093
Eth8/2     1-3967,4048-4093
Po1        1-3967,4048-4093
--------------------------------------------------------------------------------
Port       Vlans Err-disabled on Trunk
--------------------------------------------------------------------------------
Eth3/2     none
Eth8/1     none
Eth8/2     none
Po1        none
--------------------------------------------------------------------------------
Port       STP Forwarding
--------------------------------------------------------------------------------
Eth3/2     none
Eth8/1     none
Eth8/2     none
Po1        1,10
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 6.

Protokół VTP

W złożonych sieciach zarządzanie VLAN-ami może być czasochłonne i wiązać się z występowaniem wielu błędów. Protokół VTP (ang. *VLAN Trunking Protocol*) jest narzędziem umożliwiającym zarządzanie nazwami i numerami VLAN-ów na poziomie urządzeń centralnych. Pozwala to na automatyczną dystrybucję konfiguracji do pozostałych urządzeń. Weź na przykład pod uwagę sieć, której schemat przedstawia rysunek 6.1. Jest to typowy schemat sieci trójwarstwowej składającej się wyłącznie z przełączników warstwy drugiej. Jest tu w sumie 12 przełączników: dwa w rdzeniu, cztery w warstwie dystrybucyjnej i sześć w warstwie dostępowej (w rzeczywistej sieci opartej na tym schemacie mielibyśmy prawdopodobnie do czynienia z setkami przełączników).



Rysunek 6.1. Trójwarstwowa sieć przełączników

Załóżmy, że nasza sieć posiada w sumie 10 VLAN-ów. Nie jest to aż tak dużo, prawda? W taki sposób mogłaby wyglądać konfiguracja 10 VLAN-ów na przykładzie switcha 2950:

```
vlan 10
 name IT
!
vlan 20
 name Personnel
!
vlan 30
 name Accounting
!
vlan 40
name Warehouse1
!
vlan 50
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
 name Warehouse2
!
vlan 60
 name Shipping
!
vlan 70
 name MainOffice
!
vlan 80
 name Receiving
!
vlan 90
 name Lab
!
vlan 100
 name Production
```

Teraz wyobraź sobie, że każdy przełącznik w tym schemacie musi posiadać informacje o wszystkich VLAN-ach. Aby to osiągnąć, musiałbyś wprowadzić wszystkie powyższe komendy w każdym przełączniku z osobna. Oczywiście możesz skopiować kompletną sekwencję poleceń do pliku tekstowego i następnie wkleić go do konfiguracji każdego switcha, ale wciąż będzie to dość uciążliwe. Przyjrzyj się nazwom VLAN-ów. Masz tu między innymi: dwa magazyny (ang. *warehouse*), laboratorium (ang. *lab*) oraz główne biuro (ang. *main office*) — to dopiero powierzchnia! Żeby to wszystko skonfigurować, musiałbyś ciągnąć ze sobą laptopa i kabel do połączenia z konsolą do każdego przełącznika, co z pewnością zajęłoby sporo czasu.

Co więcej, pomyśl, że zajdzie nagle potrzeba usunięcia lub dodania jakiegoś VLAN-u lub chociażby zmiany nazwy jednego z nich. W takim przypadku musiałbyś wykonać swój obchód ponownie za każdym razem, gdy zaistnieje konieczność dokonania zmiany w konfiguracji.

Już słyszę, jak mówisz: „Mogę przecież użyć telnetu do połączenia z każdym przełącznikiem i dokonania zmian w konfiguracji!". Owszem, ale jeśli dokonasz zmian w parametrach VLAN-u, przez który się łączysz, zostaniesz odcięty i będziesz musiał wrócić do zarządzania za pomocą konsoli — tym razem Twój przełożony będzie groził Ci dowolnym przedmiotem, jaki będzie miał pod ręką, ponieważ sieć przestała działać, kiedy zacząłeś przy niej majstrować (nie przejmuj się jednak; takie rzeczy prawie nigdy się nie zdarzają… więcej niż raz).

Telnet jest oczywiście jakąś opcją, pod warunkiem, że bardzo uważasz przy wprowadzaniu poleceń i ustrzeżesz się literówek. Błąd człowieka jest główną przyczyną awarii na całym świecie. Na szczęście istnieje lepsze rozwiązanie: VTP.

VTP umożliwia konfigurację VLAN-ów za pośrednictwem jednego przełącznika. Wszystkie wprowadzane zmiany są wtedy propagowane na każdy switch w domenie VTP. **Domenę VTP** tworzy grupa połączonych ze sobą przełączników, na których skonfigurowano tę samą domenę. Jeśli przełączniki są między sobą połączone, ale są skonfigurowane w różnych domenach VTP, nie będą wymieniać informacji o VLAN-ach. Każdy switch może przynależeć tylko do jednej domeny VTP. Domyślnie domena VTP nie jest skonfigurowana, czyli jest zerowa (ang. *null*).

Switche z różnych domen VTP nie będą negocjować między sobą protokołu trunku. Aby w takim przypadku skonfigurować kanał trunkowy, musisz na portach łącza wybrać tryb `trunk`. Więcej szczegółów na ten temat znajdziesz w rozdziale 5.

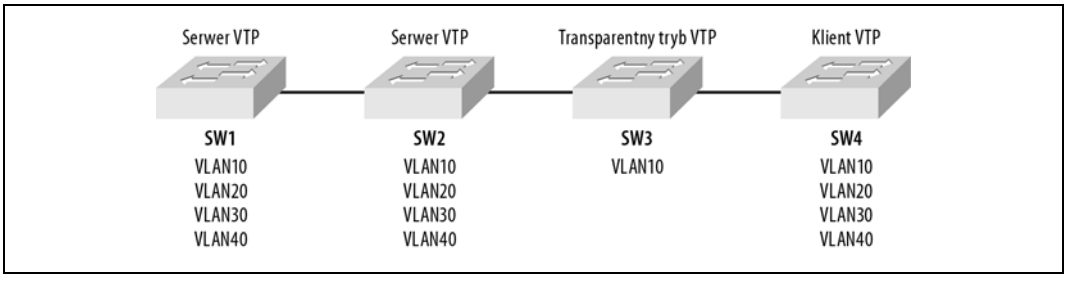
Główną koncepcją VTP jest dokonywanie zmian konfiguracyjnych na **serwerach VTP** (ang. *VTP servers*). Następnie wprowadzone zmiany propagowane są do **klientów VTP** (ang. *VTP clients*)

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

oraz pozostałych serwerów VTP w domenie. Przełączniki mogą być ręcznie skonfigurowane w trybie serwera, w trybie klienta lub jako trzecia możliwość, w **trybie transparentnym** (ang. *VTP transparent*). W trybie transparentnym switch odbiera i przekazuje dalej aktualizacje VTP, ale nie zmienia według nich własnej konfiguracji. Domyślnie jedne przełączniki ustawione są jako serwery VTP, a inne w trybie transparentnym. Na przełącznikach w trybie klienta VTP nie można lokalnie konfigurować VLAN-ów.

Jest właściwie jeszcze czwarty tryb ustawień VTP na przełączniku: `off` (wyłączone). Przełącznik w trybie VTP `off` nie będzie akceptował pakietów VTP, nie będzie też z tego względu przesyłał ich dalej. Może się to okazać pomocne, jeśli chcesz zatrzymać przesyłanie aktualizacji VTP w pewnym punkcie sieci. Tryb VTP `off` nie jest dostępny na wszystkich switchach. Znajdziesz go zazwyczaj w dużych przełącznikach typu *chassis*, jak seria 6500 czy Nexus 7000.

Rysunek 6.2 przedstawia schemat prostej sieci z czterema przełącznikami. SW1 i SW2 są ustawione jako serwery VTP. SW3 skonfigurowano w trybie transparentnym, a SW4 w trybie klienta VTP. Każda zmiana w konfiguracji VLAN-ów na SW1 będzie propagowana do SW2 i SW4. Informacje będą przechodziły przez SW3, lecz jego ustawienia nie będą modyfikowane. Ponieważ ten przełącznik nie uwzględnia aktualizacji VTP, jego VLAN-y muszą zostać skonfigurowane ręcznie, jeśli podłączeni do niego użytkownicy mają komunikować się z resztą sieci.



Rysunek 6.2. Działanie trybów VTP

Istotne jest, abyś analizując rysunek 6.2, zdawał sobie sprawę, że obydwa serwery VTP mogą przyjmować i rozpowszechniać informacje o VLAN-ach. Prowadzi to do powstania interesującego problemu. Jeśli ktoś dokonuje zmian na SW1, podczas gdy w tym samym momencie inna osoba wprowadza zmiany na SW2, to kto „wygrywa"?

Za każdym razem, kiedy dokonujesz zmian na serwerze VTP, tworzona jest nowa wersja konfiguracji. Każda kolejna wersja jest opatrywana numerem o jeden wyższym od poprzedniej. Następnie serwer przesyła do pozostałych urządzeń aktualizacje VTP, zwane też **rozgłoszeniami zbiorczymi** (ang. *summary advertisements*), zawierające aktualny numer wersji konfiguracji. Po rozgłoszeniach zbiorczych wysyłane są **rozgłoszenia podzbioru** (ang. *subset advertisements*), które zawierają informacje o konkretnych VLAN-ach.

Kiedy przełącznik odbiera pakiet aktualizacji VTP, porównuje najpierw jego nazwę domeny VTP do swojej. Jeśli domena się różni, pakiet zostaje zignorowany, natomiast gdy domena się zgadza, switch przechodzi do sprawdzenia numeru wersji konfiguracji. Jeżeli numer wersji jest niższy lub równy tej już zainstalowanej, pakiet jest ignorowany. W przypadku stwierdzenia, że otrzymana wersja jest nowsza, przełącznik wysyła **żądanie rozgłoszenia** (ang. *advertisement*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

request). W odpowiedzi na żądanie otrzymuje kolejne rozgłoszenie zbiorcze, a następnie rozgłoszenie podzbioru. Po otrzymaniu rozgłoszenia podsieci przełącznik ma wszelkie niezbędne informacje do zaimplementowania zmian w konfiguracji VLAN-ów.

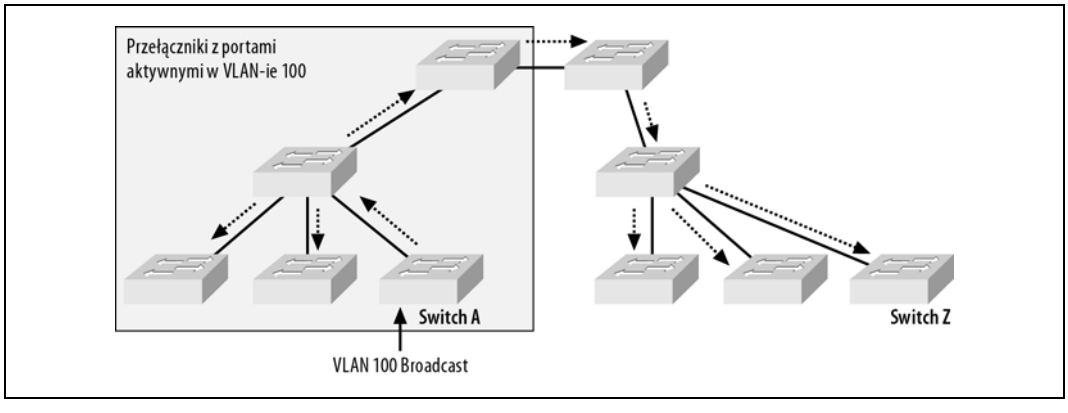
W sytuacji gdy przełącznik z zerową domeną VTP otrzyma pakiet aktualizacji VTP poprzez łącze trunkowe, odziedziczy on nazwę domeny VTP oraz konfigurację VLAN-ów od przełącznika po drugiej stronie kanału. Ma to jednak miejsce jedynie w przypadku ręcznie skonfigurowanych trunków — wynegocjowanie trunku przez DTP nie jest możliwe bez uprzedniego skonfigurowania nazwy domeny VTP. Uważaj na to zachowanie przełączników, gdyż może ono doprowadzić do poważnego bólu serca, nudności, a nawet utraty pracy, jeśli nie będziesz wystarczająco ostrożny (lub gdy nieostrożny był Twój poprzednik).

Przełączniki wysyłają również żądanie rozgłoszenia, jeśli zostały zresetowane lub zmieniła się nazwa ich domeny VTP.

Aby odpowiedzieć na postawione wcześniej pytanie, zakładając, że w danym momencie zarówno SW1, jak i SW2 miały tę samą wersję konfiguracji, ten serwer, który zapisze jako pierwszy dokonane w konfiguracji zmiany — „wygrywa". Jako pierwszy będzie więc mógł rozgłosić wyższy numer wersji konfiguracji i roześle zmiany do całej domeny. Zmiany wprowadzane na drugim przełączniku zostaną utracone, gdyż zostaną w skuteczny sposób nadpisane. Nic nie będzie też wskazywało, że zostały one utracone lub w ogóle kiedykolwiek wprowadzone.

VTP Pruning

W dużych lub zatłoczonych sieciach VTP może być przyczyną pewnych problemów w przypadku nadmiernego i przy tym często niepotrzebnego natężenia ruchu na trunkach. Spójrz na przykład na schemat sieci przedstawiony na rysunku 6.3. Wszystkie przełączniki w zaznaczonym szarym kolorem obszarze posiadają porty przynależące do VLAN-u 100, natomiast pozostałe switche nie. Przy aktywnym VTP wszystkie przełączniki będą jednak miały skonfigurowany VLAN 100, więc będą też otrzymywać pakiety broadcast zainicjowane na tym VLAN-ie. Te z nich, które nie mają portów przypisanych do VLAN-u 100, nie mają tymczasem potrzeby otrzymywania tych transmisji.

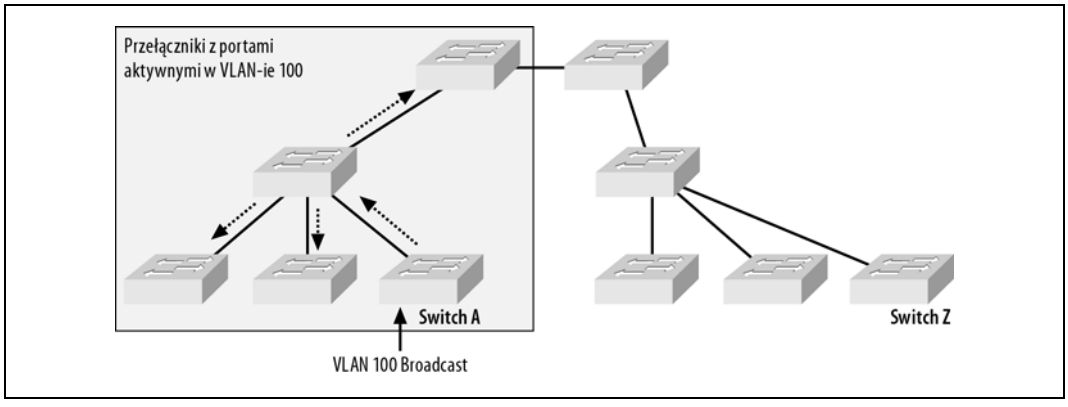


Rysunek 6.3. Broadcasty wysyłane do wszystkich przełączników w domenie VTP

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W VLAN-ie o dużym ruchu sieciowym broadcasty mogą stanowić znaczący procent całego ruchu. W takim przypadku pakiety te są niepotrzebnie rozsyłane po całej sieci i konsumują przy tym istotną część pasma kanałów trunkowych między przełącznikami.

VTP Pruning zapobiega propagowaniu transmisji przynależącej do konkretnego VLAN-u na przełączniki, w których VLAN ten nie jest **aktywny** (np. przełączniki nieposiadające podłączonych lub skonfigurowanych do tego VLAN-u portów). Przy włączonej funkcji VTP Pruning broadcasty VLAN-u 100 będą ograniczone do switchy, na których VLAN 100 jest aktywnie w użyciu, tak jak pokazuje to rysunek 6.4.



Rysunek 6.4. VTP Pruning ogranicza transmisję do przełączników z aktywnymi portami w VLAN-ie

Dokumentacja Cisco stwierdza, że pruning nie jest przeznaczony do uruchamiania na przełącznikach ustawionych w trybie transparentnym VTP: `transparent mode`.

Zagrożenia VTP

VTP, oferując wiele udogodnień, niesie ze sobą jednak kilka poważnych zagrożeń, zwłaszcza jeśli nie jesteś ostrożny.

Wyobraź sobie sieć w aktywnie działającym biurze. Biuro znajduje się na Manhattanie i zajmuje 12 pięter w drapaczu chmur. W rdzeniu struktury sieciowej są dwa przełączniki 6509, po dwa 4507R znajdują się na każdym piętrze w warstwie dystrybucji, a przełączniki 3750 tworzą warstwę dostępu w danym środowisku. Całkowita liczba switchy jest bliska 100. Protokół VTP jest wdrożony na przełącznikach rdzenia (6509) będących jedynymi serwerami VTP. Pozostałe switche skonfigurowane są w trybach klientów VTP. W sumie jest to całkiem porządnie zbudowany system, bardzo podobny do tego z rysunku 6.1 (tylko dużo większy).

Teraz załóżmy, że jeden z pracowników potrzebował switchy do działu laboratorium i zorganizował na swoje potrzeby dwa przełączniki 3750. Zostały one zainstalowane w laboratorium, ale nie podłączono ich do firmowej sieci. Przez kilka miesięcy wykorzystywano je jako wolno stojące urządzenia połączone trunkiem tylko między sobą. Konfiguracja VLAN-ów była często zmieniana, jak to zwykle ma miejsce w tego typu zastosowaniach. Co ważniejsze, konfiguracja laboratorium powstała jako lustrzane odbicie sieci firmowej, łącznie z zastosowaniem tej samej domeny VTP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Nadszedł jednak dzień, w którym pewien jajogłowy zdecydował, żeby podłączyć laboratorium do ogólnej sieci firmy. Udało mu się z powodzeniem zestawić trunk z jednym z przełączników 4507R warstwy dystrybucji. W przeciągu minuty cała sieć padła. Pamiętasz jeszcze przełożonego grożącego Ci kluczem do rur? To jeszcze nic w porównaniu z dyrektorem działu technologii instytucji finansowej, który właśnie wpadł w szał!

Co poszło nie tak? Pamiętaj, że wiele przełączników jest domyślnie ustawionych w trybie serwera VTP. Ponadto przełącznik w domenie VTP, gdy otrzymuje pakiet aktualizacyjny z wyższą wersją konfiguracji niż jego własna, dokona stosownych zmian w ustawieniach. W naszym scenariuszu przełączniki z laboratorium działały jako oddzielna sieć, ale posiadały tę samą domenę co regularna sieć firmowa. Wielokrotne wprowadzanie zmian do konfiguracji ich VLAN-ów doprowadziło do powstania pliku konfiguracji z wysokim numerem wersji. Kiedy te przełączniki, będące serwerami VTP, zostały podłączone do rzadziej konfigurowanej sieci firmowej, automatycznie rozesłały pakiety aktualizacyjne. Każdy z przełączników w głównej sieci, łącznie z przełącznikami 6509 warstwy rdzenia, otrzymał aktualizację z wyższym numerem wersji niż jego obecna konfiguracja. W konsekwencji wszystkie switche wygenerowały żądanie przesłania nowej konfiguracji VLAN-ów z podłączonych właśnie przełączników 3750 i wprowadziły wszystkie zmiany.

Szczególnie przerażające w tym scenariuszu jest to, że ludzie z laboratorium mogli być zupełnie niezaangażowani w proces podłączania ich przełączników do ogólnej sieci. Ponieważ gdyby byli, pewnie szybko rozpoznaliby, że przyczyną problemu może być schemat konfiguracji VLAN-ów w ich laboratorium. W przeciwnym przypadku zdiagnozowanie problemu mogłoby zająć sporo czasu.

Wnioskiem z tej lekcji jest, że VTP może być niebezpieczny, jeśli nie jest właściwie zarządzany. Niekiedy, jak w przypadku stabilnych, niewielkich sieci, w ogóle nie powinno się stosować VTP. Dobrą ilustracją tego typu sieci jest witryna e-commerce. Zmiany VLAN-ów będą tu występować rzadko, jeśli w ogóle, więc wdrażanie VTP nie niesie za sobą zbyt wielu korzyści.

W większych, bardziej dynamicznych środowiskach wykorzystujących VTP musisz przestrzegać odpowiednich procedur, aby zmniejszyć ryzyko wystąpienia niezamierzonych problemów. W przedstawionym przykładzie środki bezpieczeństwa, takie jak włączenie hasła VTP (omówię to w następnej sekcji), prawdopodobnie zapobiegłyby katastrofie. Co ważniejsze, być może należałoby przyjąć zasadę, że podłączanie do firmowej sieci przypadkowych, nieznanych urządzeń powinno być poprzedzone określonymi procedurami kontrolno-sprawdzającymi. Dobrym sposobem uniknięcia podłączenia niepowołanych przełączników jest wyłączenie wszystkich nieobsadzonych portów switchy w całej strukturze sieci. Wymusi to zgłaszanie przez pracowników potrzeby włączenia konkretnego portu, gdy będą chcieli podpiąć nowe urządzenie.

Konfiguracja VTP

Aby skorzystać z protokołu VTP, musisz najpierw skonfigurować domenę VTP. Musisz również wiedzieć, jak ustawić odpowiedni tryb na przełącznikach. Kolejne opcje konfiguracyjne obejmują ustalanie hasła VTP oraz uruchamianie VTP Pruning.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Domeny VTP

Domyślnie domena VTP jest zerowa. Miej to na uwadze, wdrażając ten protokół, gdyż połączenia trunkowe wynegocjowane z zerową domeną zostaną zerwane, jeśli przypiszesz inną domenę po jednej ze stron.

W zależności od modelu przełącznika możesz spotkać się z odmiennym zachowaniem na tym etapie konfiguracji. Przykładowo przełącznik Catalyst 5000 nie będzie negocjował trunków, dopóki domena VTP nie zostanie ustalona na każdym switchu.

W niektórych przełącznikach, takich jak Cisco 6500, domena zerowa zostanie nadpisana, jeśli rozgłoszenie VTP zostanie odebrane na kanale trunkowym, a przełącznik odziedziczy nazwę domeny z rozgłoszenia. Nie zdarzy się tak jednak, gdy nazwa domeny została już wcześniej skonfigurowana.

Po ustawieniu nazwy domeny na inną niż zerowa nie można jej już przywrócić do tej wartości tylko poprzez wykasowanie konfiguracji i zrestartowanie systemu.

IOS

Ustawienie lub zmiana nazwy domeny VTP odbywa się w IOS za pomocą polecenia `vtp domain`:

```
3750-IOS(config)#vtp domain GAD-Lab
Changing VTP domain name from NULL to GAD-Lab
2w0d: %DTP-5-DOMAINMISMATCH: Unable to perform trunk negotiation on port Gi1/0/20 because of VTP
↪domain mismatch.
```

W tym przypadku zmiana nazwy domeny spowodowała brak możliwości negocjacji trunku na porcie G1/0/20 z uwagi na niezgodność domen VTP.

CatOS

W CatOS w tym samym celu stosuje się komendę `set vtp domain`:

```
Switch1-CatOS# (enable)set vtp domain GAD-Lab
VTP domain GAD-Lab modified
```

Tym sposobem rozwiązałem problemy z trunkiem, więc nie pojawił się komunikat o błędzie. Gdyby wprowadzona zmiana skutkowała niezgodnością nazw domen VTP, przełącznik zaalarmowałby mnie, wyświetlając wiadomość podobną jak w IOS.

NX-OS

Jeśli chodzi o NX-OS, funkcja VTP musi zostać wpierw włączona (więcej szczegółów na ten temat znajdziesz w rozdziale 18.):

```
NX-7K-1-Daisy(config)# feature vtp
```

Od tej chwili Nexus 7000 będzie działał jedynie w trybie `transparent` (transparentnym), a Nexus 5000 w trybie VTP `off` (wyłączone). Ponieważ przełączniki te nie biorą aktywnego udziału w działaniu protokołu VTP, nie ma potrzeby ustanawiania hasła:

```
NX-7K-1-Daisy(config)# vtp mode ?
  transparent  Set the device to transparent mode
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W dokumentacji Cisco (*http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/configuration/guide/cli/VLANs.html*) można znaleźć następujące informacje na temat protokołu VTP dotyczące serii Nexus 5000:

> Protokół VTP jest wyłączony (OFF). Ramki VTP są odrzucane na wszystkich interfejsach przełączników z serii Cisco Nexus 5000, co powoduje przedzielanie domen tworzonych przez inne switche, na których protokół VTP jest włączony.

Komenda `show vtp status` umożliwia w przypadku Nexus 7000 podejrzenie statusu ustawień VTP:

```
NX-7K-1-Daisy(config)# sho vtp status
VTP Version                     : 2
Configuration Revision          : 0
Maximum VLANs supported locally : 1005
VTP Operating Mode              : Transparent
VTP Domain Name                 : accounting
VTP Pruning Mode                : Disabled
VTP V2 Mode                     : Enabled
VTP Traps Generation            : Disabled
```

Tryb VTP

Prawdopodobnie będziesz zmuszony zmienić domyślny tryb VTP na jednym lub kilku przełącznikach w domenie. W tym celu niezbędna będzie znajomość poleceń dla IOS i CatOS.

IOS

W switchach z systemem IOS dostępne są trzy tryby VTP: `server`, `client` oraz `transparent`. Do wybrania trybu służy polecenie `vtp mode`:

```
3750-IOS(config)#vtp mode ?
  client       Set the device to client mode. // ustawia tryb klient VTP
  server       Set the device to server mode. // ustawia tryb serwer VTP
  transparent  Set the device to transparent mode. // ustawia tryb transparentny VTP
```

Próba wybrania trybu, który jest już ustawiony, skutkuje komunikatem błędu:

```
3750-IOS(config)#vtp mode server
Device mode already VTP SERVER.
```

Przy udanej zmianie trybu uzyskasz krótką informację potwierdzającą wykonanie polecenia:

```
3750-IOS(config)#vtp mode transparent
Setting device to VTP TRANSPARENT mode.
```

CatOS

CatOS posiada jeszcze dodatkowy tryb: `off`. Przypomina on tryb `transparent`, w którym rozgłoszenia są ignorowane. Nie są one jednak przesyłane do innych przełączników, jak to miałoby miejsce w trybie `transparent`. Tryby konfiguruje się poleceniem `set vtp mode`:

```
Switch1-CatOS# (enable)set vtp mode ?
  client                    VTP client mode
  off                       VTP off
  server                    VTP server mode
  transparent               VTP transparent mode
```

Zmiana trybu w CatOS powoduje wyświetlenie informacji o statusie, która uwzględnia modyfikacje w domenie VTP:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Switch1-CatOS# (enable)set vtp mode transparent
Changing VTP mode for all features // Zmienia tryb VTP dla wszystkich funkcji
VTP domain GAD-Lab modified // Domena VTP GAD-Lab została zmodyfikowana
```

W przeciwieństwie do IOS próba zmiany trybu na ten, który został już ustawiony, nie powoduje wystąpienia komunikatu o błędzie.

NX-OS

Jedyne tryby dostępne w przełącznikach Nexus to `transparent` i `off`. Modele z serii 5000 działają tylko w trybie `off`, a urządzenia 7000 tylko w trybie `transparent`. W rezultacie nie ma faktycznie możliwości zmiany trybu:

```
NX-7K-1-Cozy(config)# vtp mode ?
  transparent  Set the device to transparent mode // Ustawia urządzenie w trybie transparentnym VTP
```

Hasło VTP

Ustalenie hasła VTP zapewnia, że tylko przełączniki skonfigurowane z tym samym hasłem będą uwzględniały rozgłoszenia VTP.

IOS

W IOS hasło dla VTP możesz skonfigurować komendą `vtp password`:

```
3750-IOS(config)#vtp password MilkBottle
Setting device VLAN database password to MilkBottle // Hasło bazy danych VLAN-ów ustawione na MilkBottle
```

Nie ma możliwości zaszyfrowania hasła, ale nie wyświetla się ono w opcjach konfiguracyjnych. Aby podejrzeć wprowadzone hasło, wyjdź z trybu konfiguracyjnego i użyj polecenia `show vtp password`:

```
3750-IOS#sho vtp password
VTP Password: MilkBottle
```

Aby usunąć hasło, zaneguj pierwotną komendę:

```
3750-IOS(config)#no vtp password
Clearing device VLAN database password. // Hasło bazy danych VLAN-ów zostało usunięte
```

CatOS

Do ustanowienia hasła w CatOS służy polecenie `set vtp passwd`:

```
Switch1-CatOS# (enable)set vtp passwd MilkBottle
Generating the secret associated to the password. // Hasło zostało ustanowione
VTP domain GAD-Lab modified
```

Aby zaszyfrować hasło tak, by nie było widoczne w konfiguracji, dodaj `hidden` do sekwencji komendy:

```
Switch1-CatOS# (enable)set vtp passwd MilkBottle hidden
Generating the secret associated to the password.
The VTP password will not be shown in the configuration. // Hasło VTP nie będzie widoczne
                                                         // w konfiguracji
VTP domain GAD-Lab modified
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli chcesz wyczyścić hasło w CatOS, ustaw jako hasło cyfrę zero:

```
Switch1-CatOS# (enable)set vtp passwd 0
Resetting the password to Default value. // Hasło zostało zresetowane do wartości domyślnej
VTP domain GAD-Lab modified
```

NX-OS

NX-OS nie obsługuje haseł VTP, ponieważ przełączniki te pracują jedynie w trybie `transpa ↪rent` i `off`.

VTP Pruning

VTP Pruning musi być włączony lub wyłączony na wszystkich urządzeniach w domenie VTP. Błąd w konfiguracji tej funkcji może prowadzić do niestabilnego działania całej sieci.

Domyślnie pruning może być stosowany dla wszystkich VLAN-ów, aż do VLAN-u 1001, ale z wyjątkiem VLAN-u 1. VTP nie obsługuje VLAN-ów o numerze wyższym niż 1001, nie można więc również włączyć dla nich funkcji pruning. CatOS obsługuje pruning dla VLAN-ów 2 – 1000.

Jeśli uruchomisz VTP Pruning na serwerze VTP, funkcja ta zostanie automatycznie włączona dla całej domeny.

IOS

VTP Pruning w IOS włącza się poleceniem `vtp pruning`. Można to zrobić jedynie na przełączniku ustawionym w trybie serwera VTP:

```
3750-IOS(config)#vtp pruning
Pruning switched on
```

Funkcję wyłącza się przez zanegowanie pierwotnej komendy (`no vtp pruning`).

Żeby sprawdzić, które VLAN-y umożliwiają zastosowanie funkcji pruning na trunku, wpisz komendę `interface` *`id-interfejsu`* `switchport`:

```
3750-IOS#sho int g1/0/15 switchport
Name: Gi1/0/15
Switchport: Enabled
Administrative Mode: dynamic auto
Operational Mode: down
Administrative Trunking Encapsulation: negotiate
Negotiation of Trunking: On
Access Mode VLAN: 1 (default)
Trunking Native Mode VLAN: 1 (default)
Administrative Native VLAN tagging: enabled
Voice VLAN: none
Administrative private-vlan host-association: none
Administrative private-vlan mapping: none
Administrative private-vlan trunk native VLAN: none
Administrative private-vlan trunk Native VLAN tagging: enabled
Administrative private-vlan trunk encapsulation: dot1q
Administrative private-vlan trunk normal VLANs: none
Administrative private-vlan trunk associations: none
Administrative private-vlan trunk mappings: none
Operational private-vlan: none
Trunking VLANs Enabled: ALL
Pruning VLANs Enabled: 2-1001
Capture Mode Disabled
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Capture VLANs Allowed: ALL

Protected: false
Unknown unicast blocked: disabled
Unknown multicast blocked: disabled
Appliance trust: none
```

Listę VLAN-ów, dla których dopuszczalny będzie pruning, konfigurujesz w IOS na poziomie interfejsu. Użyj komendy `switchport trunk pruning vlan` na każdym interfejsie trunkingowym przełącznika, na którym chcesz włączyć pruning:

```
3750-IOS(config-if)#switchport trunk pruning vlan ?
  WORD    VLAN IDs of the allowed VLANs when this port is in trunking mode
  add     add VLANs to the current list
  except  all VLANs except the following
      none    no VLANs
  remove  remove VLANs from the current list
```

Domyślnie wszystkie VLAN-y dopuszczają pruning. Jeśli wydasz polecenie, aby VLAN 100 dopuszczał pruning, dla IOS będzie to oznaczać, że jest to **jedyny** VLAN, na którym ta funkcja może być zastosowana:

```
3750-IOS(config-if)#switchport trunk pruning vlan 100
```

Nie zostanie przy tym wyświetlona żadna informacja, że wyłączono pruning dla VLAN-ów 2 – 99 i 101 – 1001. Musisz w tym celu ponownie podejrzeć konfigurację interfejsu:

```
3750-IOS#sho int g1/0/15 switchport | inc Pruning
Pruning VLANs Enabled: 100
```

Możesz dodać VLAN-y do listy VLAN-ów dopuszczających pruning, stosując dla komendy parametr `add`, a parametr `remove`, by je usunąć:

```
3750-IOS(config-if)#switchport trunk pruning vlan add 20-30
3750-IOS(config-if)#switchport trunk pruning vlan remove 30
```

Inna możliwość to dopuścić pruning na wszystkich VLAN-ach oprócz jednego lub kilku wybranych VLAN-ów. Służy do tego polecenie `switchport trunk pruning vlan except` *`id-vlan`*.

Pamiętaj, żeby sprawdzić zmiany w konfiguracji, których dokonałeś. Użyj komendy `show interface` *`id-interfejsu`* `switchport`. Dodawanie i usuwanie VLAN-ów może szybko doprowadzić do tego, że się pogubisz. Może się tak stać szczególnie w przypadku IOS, w którym zarządzanie funkcją VTP Pruning odbywa się na poziomie interfejsów.

CatOS

W CatOS otrzymasz ostrzeżenie dotyczące poprawności działania protokołu VTP w całej domenie, gdy będziesz chciał włączyć VTP Pruning. Prunig uruchamia się poleceniem `set vtp pruning enable`:

```
Switch1-CatOS# (enable)set vtp pruning enable
This command will enable the pruning function in the entire management domain.
                          // Ta komenda włączy funkcję pruning dla całej domeny.
All devices in the management domain should be pruning-capable before enabling.
                          // Upewnij się, że wszystkie urządzenia w domenie obsługują tę funkcję, zanim ją włączysz.
Do you want to continue (y/n) [n]? y // Czy chcesz kontynuować?
VTP domain GAD-Lab modified
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Wyłączenie pruningu skutkuje podobnym ostrzeżeniem. Aby wyłączyć VTP Pruning w CatOS, użyj polecenia `set vtp pruning disable`:

```
Switch1-CatOS# (enable)set vtp pruning disable
This command will disable the pruning function in the entire management domain.
                                                    // To polecenie wyłączy funkcję pruning dla całej domeny.
Do you want to continue (y/n) [n]? y
VTP domain GAD-Lab modified
```

Gdy funkcja pruning zostanie włączona, obejmie domyślnie VLAN-y 2 – 1000. Aby usunąć VLAN z listy, użyj komendy `clear vtp pruneeligible`. CatOS zarządza pruningiem VLAN-ów na poziomie przełącznika w przeciwieństwie do IOS, w którym odbywa się to na poziomie interfejsów:

```
Switch1-CatOS# (enable)clear vtp pruneeligible 100
Vlans 1,100,1001-4094 will not be pruned on this device.
VTP domain GAD-Lab modified.
```

Jeśli chcesz na powrót dodać VLAN-y do listy, wpisz komendę `set vtp pruneeligible`:

```
Switch1-CatOS# (enable)set vtp pruneeligible 100
Vlans 2-1000 eligible for pruning on this device.
VTP domain GAD-Lab modified.
```

NX-OS

NX-OS nie obsługuje VTP Pruning, gdyż przełączniki te pracują jedynie w trybach `transpa ↪rent` i `off`.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 7.

Agregacja portów

EtherChannel

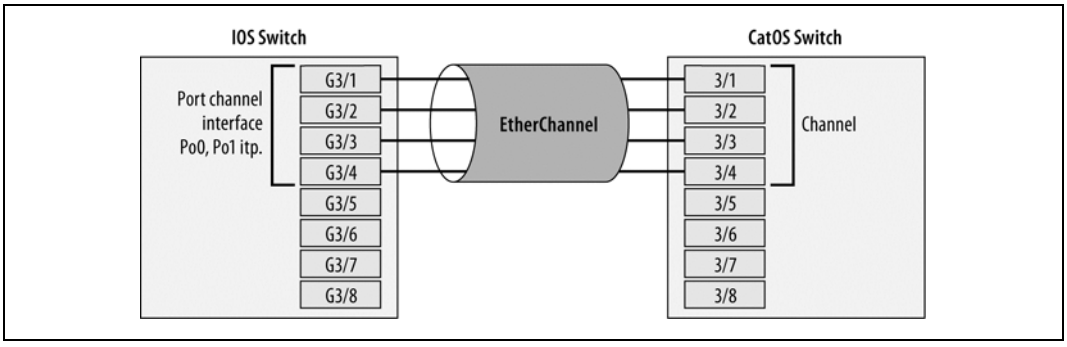
EtherChannel to termin firmy Cisco określający technologię, która umożliwia połączenie ze sobą do ośmiu ethernetowych łączy fizycznych w jedno łącze logiczne. Pierwotnie Cisco stosowało dla EtherChannel nazwę **Fast EtherChannel** (FEC), gdyż był on dostępny tylko w standardzie Fast Ethernet (100 Mb/s). W zależności od prędkości transmisji stosuje się również określenie **Gigabit EtherChannel** (GEC) lub bardziej ogólnie po prostu **Port Channel**. Inni producenci używają zazwyczaj pojęcia **agregacja portów** (ang. *Link Aggregation*), czyli LAG.

W przypadku EtherChannel prędkość pojedynczego logicznego łącza jest równa zagregowanym prędkościom wszystkich interfejsów fizycznych, które zostały użyte. Przykładowo, jeśli tworzysz EtherChannel z czterech portów Ethernet 100 Mb/s, uzyskasz kanał o prędkości 400 Mb/s. Brzmi świetnie i tak jest, jednak koncepcja ta nie jest pozbawiona wad. Po pierwsze, uzyskana przepustowość nie zawsze faktycznie odpowiada sumie prędkości łączy fizycznych. Ilustracją może być tutaj EtherChannel złożony z czterech linków 1 Gb/s, gdzie każda transmisja będzie domyślnie nadal ograniczona do prędkości 1 Gb/s.

Standardowym zachowaniem jest przydzielanie jednego fizycznego łącza do każdego pakietu, który przemierza EtherChannel, na podstawie docelowego adresu Mac dla tego pakietu. Oznacza to, że jeśli pojedyncza stacja robocza komunikuje się z pojedynczym serwerem poprzez EtherChannel, tylko jedno fizyczne łącze zostanie wykorzystane. W rzeczywistości cały ruch skierowany do serwera będzie przechodził przez pojedynczy fizyczny kanał w EtherChannel. Innymi słowy, każdemu z użytkowników w danym momencie zostanie udostępniony tylko 1 Gb/s z EtherChannel. Można zmienić to zachowanie tak, aby kolejne pakiety wysyłane były poprzez różne linki fizyczne. Jednak jak sam zobaczysz, istnieją ograniczenia dotyczące poprawności działania tego typu konfiguracji dla zastosowań takich jak VoIP. Korzyści są widoczne, gdy masz do czynienia z wieloma punktami docelowymi transmisji, z których każdy może wykorzystać osobną ścieżkę.

W zależności od tego, jaki jest system operacyjny przełącznika, EtherChannel może być różnie nazywany. Jak widać na rysunku 7.1, w CatOS EtherChannel określa się mianem **channel**, podczas gdy w IOS lub NX-OS — **port channel interface**. W przypadku CatOS komenda umożliwiająca konfigurację EtherChannel to `set port channel`, a polecenia `show port channel` i `show channel` służą do sprawdzania ustawień.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 7.1. EtherChannel w IOS i CatOS

Kanały EtherChannel w przełącznikach IOS i NX-OS są interfejsami wirtualnymi, więc należy odnosić się do nich jak do każdego innego interfejsu (np. `interface port-channel 0` lub `int po0`).

Jest jeszcze jeden problem dotyczący terminologii, który może przysporzyć licznych bólów głowy administratorom sieci. O ile w żargonie Cisco grupa fizycznych interfejsów Ethernet połączonych w wiązkę zwana jest **EtherChannel**, o tyle administratorzy Uniksa stosują czasem do tej samej konfiguracji nazwę **trunk**. Oczywiście w środowisku Cisco termin „trunk" oznacza coś zupełnie innego, czyli kanał tagujący ramki znacznikami VLAN-ów, tak aby wiele VLAN-ów mogło być transmitowanych równolegle tym łączem. Niektórzy nowocześni Uniksowcy będą używać niekiedy terminu interfejsu **bond** dla agregacji portów, a administratorzy Windowsa określenia **grupowanie** (ang. *teaming*).

Rozłożenie obciążenia w EtherChannel

Jak już wspomniałem wcześniej, EtherChannel nie zapewnia tak naprawdę całkowitej agregacji prędkości połączonych interfejsów fizycznych. Poprzez przesyłanie określonych pakietów przez wybrane kanały fizyczne EtherChannel daje **wrażenie** prędkości łącznej sprzężonych interfejsów. Domyślnie fizyczny link dla każdego pakietu jest określany na podstawie adresu MAC jego przeznaczenia. Algorytm ten jest własnością Cisco. Gwarantuje on, że pakiety z określonym adresem MAC przeznaczenia będą zawsze przesyłane tym samym łączem fizycznym. Takie rozwiązanie zapewnia, że pakiety przeznaczone dla tego adresu MAC będą zawsze docierać we właściwej kolejności.

Algorytm haszujący stosowany do określania fizycznego połączenia nie musi zawsze być algorytmem publicznym, ale obciążenie wykorzystywanych przez niego łączy jest publikowane. Istotna jest tutaj informacja, że idealne rozłożenie obciążenia między fizycznymi linkami nie musi być zapewnione.

Działanie algorytmu haszującego polega na przyporządkowaniu docelowemu adresowi MAC (lub, jak zobaczysz później, innemu parametrowi) cyfry z zakresu 0 – 7. Ten sam zakres wartości stosowany jest za każdym razem, bez względu na to, ile połączeń jest użytych w EtherChannel. Każdemu fizycznemu kanałowi przyporządkowana jest jedna lub więcej z tych wartości, w zależności od tego, ile łączy ma EtherChannel. Z tych względów EtherChannel dopuszcza maksymalnie osiem fizycznych połączeń.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Rysunek 7.2 przedstawia sposób dystrybucji pakietów stosowany w tej metodzie. Zauważ, że dystrybucja nie rozkłada się zawsze równomiernie. Ważne, byś to zrozumiał, ponieważ wykażą to statystyki łączy, a szczególnie wykresy graficzne.

Numer łącza \ Liczba łączy fizycznych	8	7	6	5	4	3	2
1	1	2	2	2	2	3	4
2	1	1	2	2	2	3	4
3	1	1	1	2	2	2	
4	1	1	1	1	2		
5	1	1	1	1			
6	1	1	1				
7	1	1					
8	1						

Rysunek 7.2. Rozłożenie obciążenia fizycznych łączy w EtherChannel

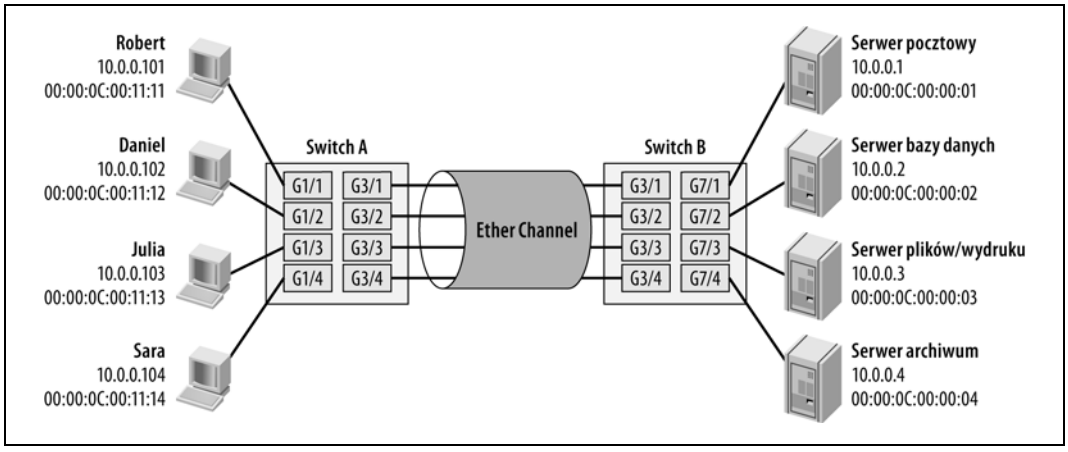
W EtherChannel, który składa się z ośmiu kanałów fizycznych, każdemu z tych kanałów przypisana jest jedna wartość. Gdy w EtherChannel jest sześć kanałów, dwa z nich otrzymają po dwie wartości, a pozostałe po jednej. Oznacza to, że te dwa łącza (oczywiście tylko w idealnym modelu teoretycznym) będą musiały obsłużyć dwukrotnie większy ruch niż pozostałe cztery. Nie wszystkie połączenia w EtherChannel będą więc wykorzystywane w równym stopniu. Po przeanalizowaniu rysunku 7.2 powinno stać się oczywiste, że jedynym sposobem zrównoważenia dystrybucji ruchu na wszystkich łączach EtherChannel (w przypadku idealnego rozłożenia obciążenia) jest wykorzystanie do jego budowy ośmiu, czterech lub dwóch łączy fizycznych. Bez względu na informację służącą do określania łącza metoda nadal będzie polegała na przyporządkowaniu tym wartościom cyfry z zakresu 0 – 7, która zostanie wykorzystana do obciążenia łączy według przedstawionej tabeli.

Możesz zmienić metodę, która jest stosowana w przełączniku do określania ścieżki dla transmisji pakietów. Domyślnie wykorzystuje się docelowy adres MAC. Jednak w zależności od wersji systemu operacyjnego oraz sprzętu możesz napotkać następujące opcje:

- źródłowy adres MAC;
- docelowy adres MAC;
- źródłowy i docelowy adres MAC;
- źródłowy adres IP;
- docelowy adres IP;
- źródłowy i docelowy adres IP;
- port źródłowy;
- port docelowy;
- port źródłowy i docelowy.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Powody, dla których zmienia się domyślne ustawienia, zależeć będą od okoliczności. Rysunek 7.3 przedstawia stosunkowo częsty układ: grupa użytkowników podłączonych do Switcha A komunikuje się poprzez EtherChannel z grupą serwerów na przełączniku B. Domyślnie metoda rozkładania obciążenia będzie oparta na docelowych adresach MAC każdego z pakietów. Problemem jest tutaj sposób wykorzystania kanału. Z uwagi na unikatowość adresów MAC mógłbyś sądzić, że łącza fizyczne kanału będą obciążone równomiernie. W rzeczywistości jednak większa część generowanego ruchu przeznaczona jest do jednego z serwerów, podczas gdy pozostałe otrzymują znacznie mniej pakietów.



Rysunek 7.3. Czynniki rozkładu obciążenia dla EtherChannel

Załóżmy, że serwer pocztowy w tej sieci otrzymuje ponad 1 Gb/s danych, podczas gdy pozostałe serwery średnio po około 50 Mb/s. Metoda wykorzystująca docelowy adres MAC doprowadzi do utraty części pakietów na EtherChannel, ponieważ każdy z pakietów skierowany do adresu MAC serwera pocztowego będzie transmitowany tym samym łączem fizycznym EtherChannel. Niemieszczące się pakiety nie będą kierowane na inne łącza — przy zapełnieniu łącza będą one po prostu odrzucane.

Przypadek ten zasługuje na specjalny komentarz. Spotkałem wielu inżynierów sieciowych, niektórych z wieloletnim doświadczeniem, którzy mocno wierzą, że EtherChannel automatycznie przesunie ruch sieciowy z jednego łącza fizycznego na drugie, jeśli to pierwsze okaże się przepełnione. W tej sytuacji tak się nie stanie.

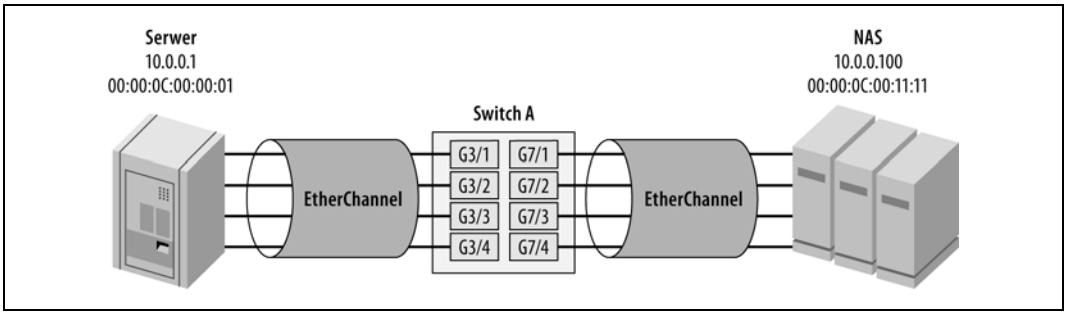
W scenariuszu, w którym jeden z serwerów otrzymuje lwią część pakietów w sieci, rozłożenie obciążenia bazujące na docelowym adresie MAC nie ma większego sensu. Bardziej rozsądne wydaje się zastosowanie metody opartej na źródłowym adresie MAC.

Kolejną istotną kwestią wartą zapamiętania jest to, że metoda równoważenia obciążenia stosowana jest tylko do pakietów **wysyłanych** przez EtherChannel. Nie jest to funkcja dwukierunkowa. O ile zmiana metody na Switchu A na źródłowy adres MAC może być dobrym pomysłem, o tyle w przypadku Switcha B byłoby to koszmarne rozwiązanie, ponieważ serwer pocztowy jest najbardziej obciążonym serwerem. Pamiętaj, że źródłowym adresem MAC pakietów zwracanych przez serwer pocztowy jest właśnie adres tego serwera. Więc jeśli wykorzystamy źródłowy adres MAC do zrównoważenia obciążenia na Switchu B, napotkamy ten sam problem, który właśnie próbujemy rozwiązać.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W tych okolicznościach rozwiązaniem jest zastosowanie na Switchu A metody rozkładania obciążenia opartej na źródłowym adresie MAC, a na Switchu B na docelowym adresie MAC. Jeśli wszystkie serwery podłączone są do jednego przełącznika, a wszyscy użytkownicy do drugiego, taka konfiguracja powinna zdać egzamin. Niestety w praktyce problemy, które napotkasz, są najczęściej bardziej złożone. Prawdopodobniejszym scenariuszem jest, że wszystkie urządzenia w sieci są podłączone do jednego dużego przełącznika, jak np. 6509. Zmiana algorytmu rozłożenia obciążenia odbywa się na poziomie magistrali całego przełącznika, więc jeśli wszystkie urządzenia są do niego podłączone, to masz pecha.

Rysunek 7.4 przedstawia interesujący problem. Mamy tu pojedynczy serwer podłączony do Switcha A poprzez EtherCahnnel oraz pojedyncze sieciowe urządzenie magazynujące typu NAS (ang. *Network Attached Storage*), również podłączone do Switcha A poprzez EtherChannel. Wszystkie systemy plików dla serwera znajdują się na urządzeniu NAS, a sam serwer jest mocno obciążony — jako serwer bazy danych obsługuje ponad 5000 użytkowników jednocześnie. Przepustowość wymagana pomiędzy serwerem i NAS-em przekracza 2 GB/s.



Rysunek 7.4. Pojedynczy serwer połączony z urządzeniem NAS

Niestety nie istnieje proste rozwiązanie tego problemu. Nie możemy wykorzystać docelowego ani źródłowego adresu MAC do metody rozkładania obciążenia, ponieważ w każdym przypadku jest tylko jeden adres, zawsze taki sam. Z tych samych względów nie możemy wykorzystać kombinacji źródłowego i docelowego adresu MAC lub źródłowego i/lub docelowego adresu IP. Nie możemy też użyć źródłowych lub docelowych numerów portów, ponieważ po wynegocjowaniu pozostają niezmienne. Jedną z możliwości, zakładając, że sterowniki ją obsługują, jest zmiana serwera i/lub urządzenia NAS tak, aby każdy link miał swój własny adres MAC. Dla pakietów jednak nadal źródłem i przeznaczeniem będzie tylko jeden z tych adresów.

Jedyne dostępne rozwiązania w tej sytuacji to ręczne rozłożenie obciążenia lub szybsze łącza. Podzielenie kanału na cztery łącza po 1 Gb/s, każde w osobnej sieci IP, oraz zamontowanie różnych systemów plików dla każdego łącza rozwiązuje problem. Jest to jednak zbyt skomplikowane jak na mój gust. Lepszym wyjściem, o ile jest ono dostępne, może być użycie szybszych łączy fizycznych, takich jak 10 Gb Ethernet.

EtherChannel — konfiguracja i zarządzanie

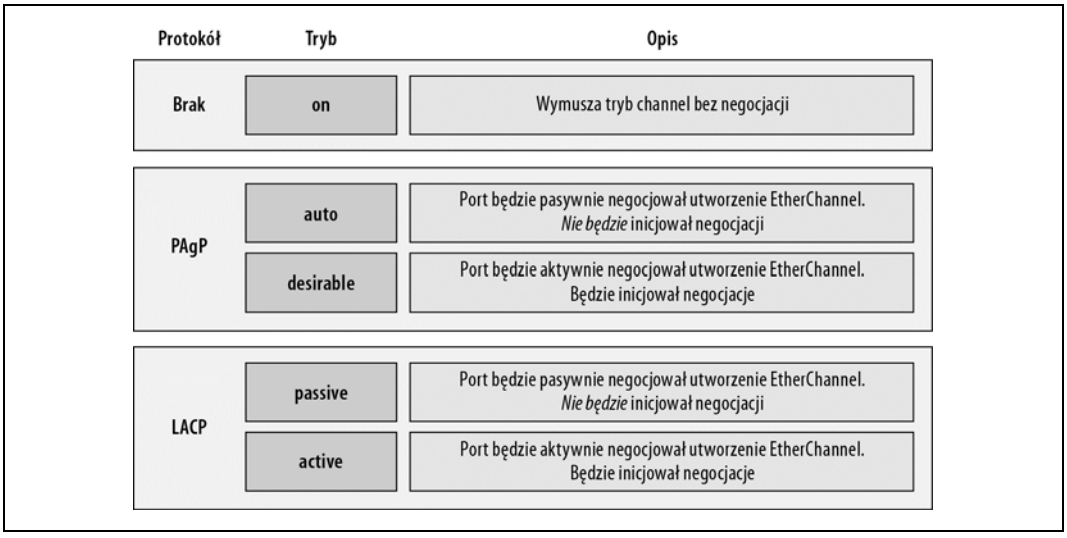
Zazwyczaj to urządzenie znajdujące się po jednej ze stron EtherChannel determinuje jego konfigurację. Jedną z zasad, do których musisz się zawsze stosować, jest to, że każde łącze w EtherChannel musi mieć tę samą konfigurację. Opisy łączy mogą być różne, ale wszystkie łącza fizyczne

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

muszą być tego samego typu oraz mieć tę samą prędkość. Muszą być również w tym samym VLAN-ie. Jeśli łącza są trunkami, powinny też posiadać te same parametry trunku.

Protokoły EtherChannel

EtherChannel może negocjować z urządzeniem znajdującym się po drugiej stronie łącza. W urządzeniach Cisco obsługiwane są dwa protokoły. Pierwszym z nich jest LACP (ang. *Link Aggregation Control Protocol*), zdefiniowany w specyfikacji IEEE 803.ad. LACP jest stosowany w przypadku łączenia z urządzeniami innej marki niż Cisco, np. z serwerami. Drugi protokół wykorzystywany do negocjacji kanałów EtherChannel to PAgP (ang. *Port Aggregation Control Protocol*). Ponieważ PAgP jest własnością firmy Cisco, stosowany jest tylko do łączenia dwóch urządzeń Cisco poprzez EtherChannel. Każdy z protokołów obsługuje dwa tryby: tryb pasywny (nazywany `auto` w PAgP i `passive` w LACP) oraz tryb aktywny (`desirable` dla PAgP i `active` dla LACP). Alternatywnie możesz ustawić tryb `on`, wymuszając utworzenie EtherChannel. Dostępne protokoły i tryby przedstawione są na rysunku 7.5.



Rysunek 7.5. Protokoły EtherChannel i ich dostępne tryby

Zasadniczo, gdy konfigurujesz EtherChannel pomiędzy dwoma urządzeniami Cisco, użyte porty zostaną przypisane do EtherChannel na cały czas funkcjonowania danej instalacji. Sensowne jest ustawienie wszystkich interfejsów po obu stronach EtherChannel w trybie `desirable`. Przy łączeniu Cisco z urządzeniami innego typu, jak np. z maszynami Uniksowymi, skorzystaj z trybu `active` protokołu LACP. Miej też na uwadze, że niektóre urządzenia stosują własne metody łączenia kanałów, więc wymagają, aby po stronie urządzeń Cisco ustawiono tryb `on`, ponieważ nie negocjują one z drugą stroną łącza.

Na przykładzie CatOS

Tworzenie EtherChannel w CatOS jest stosunkowo proste, jeśli wiesz, czego potrzebujesz. Dla przykładu utworzymy EtherChannel pokazany na rysunku 7.1. Ponieważ urządzenia po obu stronach kanału to przełączniki Cisco, skonfigurujemy wszystkie porty w trybie `desirable` (wykorzystamy protokół PAgP oraz inicjowanie negocjacji):

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
set port name       3/1  Link #1 in Channel
set port name       3/2  Link #2 in Channel
set port name       3/3  Link #3 in Channel
set port name       3/4  Link #4 in Channel

set vlan 20    3/1-4

set port channel 3/1-4 mode desirable
```

Zakładając, że druga strona jest ustawiona właściwie, to wszystko, czego potrzebujemy, by uruchomić EtherChannel. Przypisywanie nazw nie jest konieczne, jednak dobrze jest mieć wszystko oznaczone i wprowadzone do dokumentacji, nawet jeśli nie ma takiego wymagania.

Kiedy już skonfigurowaliśmy EtherChannel, musimy mieć możliwość sprawdzenia jego statusu. Spójrzmy na wynik użycia komendy `show port channel`:

```
Switch-2-CatOS: (enable)sho port channel
Port  Status     Channel              Admin Ch
                 Mode                 Group Id
----- ---------- -------------------- ----- -----
 3/1 connected  desirable               74   770
 3/2 connected  desirable               74   770
 3/3 connected  desirable               74   770
 3/4 connected  desirable               74   770
```

`Show channel info` wyświetla podobne zestawienie co `show port channel`, zawierające jednak nieco więcej informacji. W większości przypadków to polecenie jest bardziej użyteczne, ponieważ pozwala uzyskać jednocześnie również takie dane jak: ID kanału (ang. *channel ID*), grupa administracyjna (ang. *admin group*), typ interfejsu (ang. *interface type*), tryb dupleksu (ang. *duplex mode*) oraz przypisany VLAN:

```
Switch-2-CatOS: (enable)sho channel info
Chan Port  Status     Channel              Admin Speed Duplex Vlan
id                    mode                 group
---- ----- ---------- -------------------- ----- ----- ------ ----
 770  3/1  connected  desirable               74 a-1Gb a-full   20
 770  3/2  connected  desirable               74 a-1Gb a-full   20
 770  3/3  connected  desirable               74 a-1Gb a-full   20
 770  3/4  connected  desirable               74 a-1Gb a-full   20
```

Komenda `show channel` pozwala odczytać bardzo skróconą informację na temat tego, które porty przypisane są do których kanałów:

```
Switch-2-CatOS: (enable)sho channel
Channel Id   Ports
-----------  ----------------------------------------------
770          3/1-4
```

Kolejnym użytecznym pleceniem jest `show channel traffic`. Przedstawia ono wykorzystanie łączy, z dystrybucją obciążenia podaną w procentach:

```
Switch-2-CatOS: (enable)sho channel traffic
ChanId Port  Rx-Ucst Tx-Ucst Rx-Mcst Tx-Mcst Rx-Bcst Tx-Bcst
------ ----- ------- ------- ------- ------- ------- -------
   770  3/1   21.80%  18.44%  87.48%  87.70%  26.49%  21.20%
   770  3/2   34.49%  37.97%   4.02%   4.98%  19.38%  11.73%
   770  3/3   21.01%  23.47%   3.99%   3.81%  29.46%  28.60%
   770  3/4   22.66%  20.06%   4.13%   2.79%  23.69%  38.32%
```

Zauważ, że suma wartości procentowych nie zawsze wynosi 100%. To narzędzie służy raczej do pokazania pewnych tendencji.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Na przykładzie IOS

Konfiguracja EtherChannel w przełącznikach z systemem IOS nie jest trudna, jednak, jak już wcześniej pisałem, dla tych, którzy są przyzwyczajeni do CatOS, terminologia może wydawać się nieco dziwna. Zasadniczą różnicą jest tworzenie wirtualnego interfejsu port channel. Daje Ci to dużą swobodę: jeśli zechcesz, możesz skonfigurować interfejs z adresem IP (zakładając, że przełącznik obsługuje warstwę trzecią) lub po prostu pozostawić jako zwykły port switcha. Pamietaj, że każdy interfejs musi mieć takie same ustawienia, wyjątkiem jest tu jedynie opis. Osobiście staram się nadawać na wszystkich fizycznych portach opisy, które mają określone znaczenie. Pozwala mi to sledzić przypisanie danego interfejsu, gdyż komenda `show interface` nie wskazuje, czy interfejs przynależy do EtherChannel.

Ponownie utworzymy dla przykładu EtherChannel z rysunku 7.1, tak więc mamy przełączniki Cisco po obu stronach kanału:

```
interface Port-channel1
 description 4G Etherchannel Po1
 no ip address
 switchport
 switchport access vlan 20

interface GigabitEthernet3/1
 description Link #1 in Po1
 no ip address
 switchport
 channel-group 1 mode desirable

interface GigabitEthernet3/2
 description Link #2 in Po1
 no ip address
 switchport
 channel-group 1 mode desirable

interface GigabitEthernet3/3
 description Link #3 in Po1
 no ip address
 switchport
 channel-group 1 mode desirable

interface GigabitEthernet3/4
 description Link #4 in Po1
 no ip address
 switchport
 channel-group 1 mode desirable
```

Szybkim sposobem na sprawdzenie statusu EtherChannel w IOS jest zastosowanie komendy `show etherchannel summary`. Użytkownicy CatOS mogą poczuć się sfrustrowani złożonością informacji wynikowej. Na początek musisz się zapoznać z kodami opisanymi w legendzie. Potem będziesz mógł określić status EtherChannel. W tym przypadku mamy do czynienia z EtherChannel w warstwie drugiej i kanał ten jest w użyciu (`Po1(SU)`). Wszystkie indywidualne łącza fizyczne są aktywne, ponieważ mają oznaczenie (`P`) obok numeru portu:

```
Switch-1-IOS#sho etherchannel summary
Flags:  D - down        P - in port-channel // Flagi: D — wyłączony, P — w port channel
        I - stand-alone s - suspended // I — samodzielny, s — zawieszony
        H - Hot-standby (LACP only) // H — w trybie gotowości (tylko LACP)
        R - Layer3      S - Layer2 // R — warstwy trzeciej, S — warstwy drugiej
        U - in use      f - failed to allocate aggregator
                                   // U — w użyciu, f — nie udało się zlokalizować agregacji
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
        u - unsuitable for bundling // u — nieodpowiedni do łączenia
Number of channel-groups in use: 1 // Liczba grup kanałów w użyciu: 1
Number of aggregators:           1 // Liczba agregacji
Group  Port-channel  Protocol    Ports
------+-------------+-----------+-----------------------------------------------
1      Po1(SU)         PAgP      Gi3/1(P)   Gi3/2(P)    Gi3/3(P)     Gi3/4(P)
```

Bardziej użyteczną komendą, mimo że nie pokazuje ona rzeczywistego statusu interfejsów, jest `show etherchannel`. Jest to interesująca komenda, gdyż informuje o liczbie bitów wykorzystanych w algorytmie haszującym dla każdego interfejsu fizycznego, jak to zilustrowano uprzednio na rysunku 7.3. Co ciekawe, pojawia się też tutaj informacja, od jak dawna interfejs jest przyłączony do EtherChannel.

```
Switch-1-IOS #sho etherchannel 1 port-channel
                Port-channels in the group:
                ----------------------
Port-channel: Po1
------------
Age of the Port-channel   = 1d:09h:22m:37s
Logical slot/port   = 14/6            Number of ports = 4
GC                  = 0x00580001      HotStandBy port = null
Port state          = Port-channel Ag-Inuse
Protocol            =   PAgP
Ports in the Port-channel:
Index   Load   Port     EC state        No of bits
------+------+------+------------------+-----------
  1     11     Gi3/1    Desirable-Sl    2
  2     22     Gi3/2    Desirable-Sl    2
  0     44     Gi3/3    Desirable-Sl    2
  3     88     Gi3/4    Desirable-Sl    2
Time since last port bundled:    1d:09h:21m:08s    Gi3/4
```

Ponieważ kanały EtherChannel są w IOS przypisane do wirtualnych interfejsów, możesz wyświetlić informacje o danym interfejsie, jakby był interfejsem fizycznym lub wirtualnym. Zauważ, że szerokość pasma ustawiono na sumę prędkości użytych łączy, ale poniżej linia dupleksu wskazuje konfigurację `Full-duplex, 1000 Mb/s`. Fizycznie do sprzętowego EtherChannel należą interfejsy Gi3/1, Gi3/2, Gi3/3 oraz Gi3/4:

```
Switch-1-IOS#sho int port-channel 1
Port-channel1 is up, line protocol is up (connected)
  Hardware is EtherChannel, address is 0011.720a.711d (bia 0011.720a.71
  Description: 4G Etherchannel Po1
  MTU 1500 bytes, BW 4000000 Kbit, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Full-duplex, 1000Mb/s
  input flow-control is off, output flow-control is unsupported
  Members in this channel: Gi3/1 Gi3/2 Gi3/3 Gi3/4
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input never, output never, output hang never
  Last clearing of "show interface" counters 30w6d
  Input queue: 0/2000/1951/0 (size/max/drops/flushes); Total output drops: 139
  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  5 minute input rate 3906000 bits/sec, 628 packets/sec
  5 minute output rate 256000 bits/sec, 185 packets/sec
     377045550610 packets input, 410236657639149 bytes, 0 no buffer
     Received 66730119 broadcasts (5743298 multicast)
     0 runts, 0 giants, 0 throttles
     0 input errors, 0 CRC, 0 frame, 1951 overrun, 0 ignored
     0 watchdog, 0 multicast, 0 pause input
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
     0 input packets with dribble condition detected
     255121177828 packets output, 159098829342337 bytes, 0 underruns
     0 output errors, 0 collisions, 0 interface resets
     0 babbles, 0 late collision, 0 deferred
     0 lost carrier, 0 no carrier, 0 PAUSE output
     0 output buffer failures, 0 output buffers swapped out
```

Ponieważ indywidualne łącza są łączami fizycznymi, interfejsy te mogą być przedstawione w IOS jako interfejs fizyczny za pomocą komendy `show interface`:

```
Switch-1-IOS#sho int g3/1
GigabitEthernet5/1 is up, line protocol is up (connected)
  Hardware is C6k 1000Mb 802.3, address is 0011.7f1a.791c (bia 0011.7f1a.791c)
  Description: Link #1 in Po1
  MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Full-duplex, 1000Mb/s
  input flow-control is off, output flow-control is off
  Clock mode is auto
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:00:45, output 00:00:03, output hang never
  Last clearing of "show interface" counters 30w6d
  Input queue: 0/2000/1054/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  5 minute input rate 924000 bits/sec, 187 packets/sec
  5 minute output rate 86000 bits/sec, 70 packets/sec
     190820216609 packets input, 207901078937384 bytes, 0 no buffer
     Received 48248427 broadcasts (1757046 multicast)
     0 runts, 0 giants, 0 throttles
     0 input errors, 0 CRC, 0 frame, 1054 overrun, 0 ignored
     0 watchdog, 0 multicast, 0 pause input
     0 input packets with dribble condition detected
     129274163672 packets output, 80449383231904 bytes, 0 underruns
     0 output errors, 0 collisions, 0 interface resets
     0 babbles, 0 late collision, 0 deferred
     0 lost carrier, 0 no carrier, 0 PAUSE output
     0 output buffer failures, 0 output buffers swapped out
```

Zwróć uwagę, że oprócz informacji w opisie nie ma tu żadnego wskazania, iż dany interfejs należy do EtherChannel. Wzmacnia to moje przekonanie, że wszystkie porty powinny być oznakowane za pomocą polecenia `description`.

Na przykładzie NX-OS

Tworzenie i zarządzanie interfejsami typu port channel odbywa się w NX-OS prawie tak samo jak w IOS:

```
NX-7K-1-Cozy(config)# int e3/10-11
NX-7K-1-Cozy(config-if-range)# channel-group 10
```

Kiedy już grupa kanałów (ang. *channel group*) zostanie utworzona, możesz podejrzeć jej status poleceniem `show interface`:

```
NX-7K-1-Cozy(config-if-range)# sho int po 10
port-channel10 is down (No operational members)
  Hardware: Port-Channel, address: 0026.9807.95c2 (bia 0000.0000.0000)
  MTU 1500 bytes, BW 100000 Kbit, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA
  auto-duplex, auto-speed
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  Input flow-control is off, output flow-control is off
  Switchport monitor is off
  Members in this channel: Eth3/10, Eth3/11
  Last clearing of "show interface" counters never
  30 seconds input rate 0 bits/sec, 0 packets/sec
  30 seconds output rate 0 bits/sec, 0 packets/sec
  Load-Interval #2: 5 minute (300 seconds)
    input rate 0 bps, 0 pps; output rate 0 bps, 0 pps
  L3 in Switched:
    ucast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes
  L3 out Switched:
    ucast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes
  RX
    0 unicast packets  0 multicast packets  0 broadcast packets
    0 input packets  0 bytes
    0 jumbo packets  0 storm suppression packets
    0 runts  0 giants  0 CRC  0 no buffer
    0 input error  0 short frame  0 overrun   0 underrun  0 ignored
    0 watchdog  0 bad etype drop  0 bad proto drop  0 if down drop
    0 input with dribble  0 input discard
    0 Rx pause
  TX
    0 unicast packets  0 multicast packets  0 broadcast packets
    0 output packets  0 bytes
    0 jumbo packets
    0 output error  0 collision  0 deferred  0 late collision
    0 lost carrier  0 no carrier  0 babble
    0 Tx pause
  0 interface resets
```

Ciekawą komendą jest `show interface brief`, umożliwiająca sprawdzenie, które łącza przynależą do danego port channel. Nie jest to komenda `show ip interface brief`, do której mogłeś się już przyzwyczaić. Spójrz na ostatnią kolumnę — wskazuje ona, do którego port channel jest przypisany konkretny interfejs:

```
NX-7K-1-Cozy(config-if-range)# sho int brie
--------------------------------------------------------------------------------
Ethernet      VLAN   Type Mode   Status  Reason                   Speed     Port
Interface                                                                    Ch #
--------------------------------------------------------------------------------
Eth3/1        1      eth  trunk  down    Link not connected         auto(D) --
Eth3/2        10     eth  access down    Link not connected         auto(D) --
Eth3/3        --     eth  routed down    Administratively down      auto(D) --
Eth3/4        --     eth  routed down    Administratively down      auto(D) --
Eth3/5        --     eth  routed down    Administratively down      auto(D) --
Eth3/6        --     eth  routed down    Administratively down      auto(D) --
Eth3/7        --     eth  routed down    Administratively down      auto(D) --
Eth3/8        --     eth  routed down    Administratively down      auto(D) --
Eth3/9        --     eth  routed down    Administratively down      auto(D) --
Eth3/10       --     eth  routed down    Administratively down      auto(D) 10
Eth3/11       --     eth  routed down    Administratively down      auto(D) 10
Eth3/12       --     eth  routed down    Administratively down      auto(D) --
[fragment usunięty]
```

Są to świetne polecenia, ale oczywiście potrzeba trochę czasu, by je opanować. Komendy `show etherchannel` z IOS, do których mogłeś się przyzwyczaić, nie są dostępne w aktualnej wersji NX-OS.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Cross-stack EtherChannel

Czasem pojawia się możliwość utworzenia EtherChannel „w poprzek" dwóch fizycznych przełączników. Przykładowo dwa połączone w stos (zestakowane) przełączniki 3750 stają się logiczną całością, a każdy z nich działa jak moduł typu „blade". Konfiguracja EtherChannel w zestakowanych przełącznikach jest możliwa, a nawet zalecana w celu wyeliminowania pojedynczych punktów stwarzających ryzyko awarii. Jeśli stos został prawidłowo skonfigurowany, nie wymaga to żadnych specjalnych ustawień — po prostu przypisz każdy z portów do tej samej grupy port channel:

```
3750-IOS(config)#int range g1/0/1 , g2/0/1
3750-IOS(config-if-range)#channel-group 10 mode on
Creating a port-channel interface Port-channel 10
```

Cisco nazywa to **cross-stack EtherChannel** („EtherChannel w stosie"). Oferuje on kilka ciekawych możliwości, ale ma też swoje ograniczenia. Weźmy dla przykładu cross-stack EtherChannel z 16 portami, z których osiem jest aktywnych (ang. *active*), a kolejne 8 w trybie gotowości (ang. *standby*). Nie będziesz mógł jednak użyć w tym przypadku PAgP — możesz jedynie wybrać tryb `active`, `passive` lub `on`.

Multichassis EtherChannel (MEC)

Multichassis EtherChannel jest funkcją co do zasady bardzo podobną do cross-stack EtherChannel. Różnica polega głównie na tym, że jest ona dostępna jedynie w przełącznikach Catalyst 6500 VSS. VSS (Virtual Switching System) jest technologią tworzenia klastrów umożliwiającą skonfigurowanie kilku przełączników 6500 typu *chassis* w jeden logiczny switch. Działa to analogicznie do połączonych w stos przełączników 3750. Warunkiem zastosowania MEC jest właściwa konfiguracja VSS, a przełączniki muszą być aktywne w klastrze VSS. Bardziej szczegółowe informacje na temat konfiguracji VSS znajdziesz w rozdziale 17.

Kiedy uruchomione jest VSS, konfiguracja MEC jest równie prosta jak w przypadku 3750. Poniżej utworzyłem 20 Gb kanał MEC pomiędzy dwoma przełącznikami 6500 w klastrze VSS:

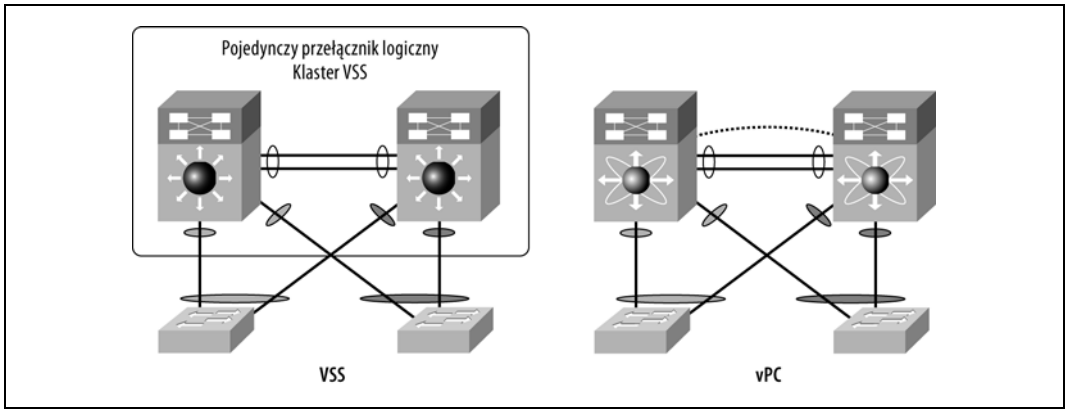
```
VSS(config)#int range ten 1/3/2 , ten 2/3/2
VSS(config-if-range)#channel-group 500 mode on
```

Virtual Port Channel

Virtual Port Channel (vPC) jest rozwiązaniem typu multichassis dla interfejsów port channel przełączników Nexus. O ile rezultat jest bardzo zbliżony, o tyle zastosowana metoda znacznie różni się od MEC. Podstawowa różnica wiąże się z tym, że nie możesz połączyć w klaster kilku przełączników Nexus, tak jak dzieje się to w przypadku 6500 VSS. Posiadanie wielu łączy fizycznych traktowanych jako jedno łącze logiczne jest dość skomplikowane i może upłynąć nieco czasu, zanim przyzwyczaisz się do rozwiązań stosowanych w tej technologii. Różnice logiczne przedstawia rysunek 7.6.

VSS i vPC pozwalają osiągnąć ten sam cel, ale przy zastosowaniu różnych metod. W przypadku VSS klaster jest tak skonfigurowany, że dwa fizyczne switche 6500 zachowują się jak jeden przełącznik logiczny (jedna magistrala sterowania i jedna magistrala danych). Dla vPC

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 7.6. VSS MEC kontra Nexus vPC

w Nexus 5000 lub 7000 przełączniki pozostają oddzielnymi podmiotami fizycznymi i logicznymi (oddzielne magistrale sterowania, jedna magistrala danych). EtherChannel kończy się na dwóch fizycznych przełącznikach w sieci, a jednocześnie wygląda to tak, jakby był podłączony do jednego przełącznika fizycznego. Jeśli wydaje Ci się to nieco pogmatwane, to dlatego, że takie właśnie jest. Dobra wiadomość jest taka, że jeśli już „zaskoczysz", o co chodzi, to okaże się to bardzo przystępne i łatwe w konfiguracji.

Podczas gdy MEC wymaga uruchomienia i skonfigurowania klastra VSS, dla vPC należy włączyć i skonfigurować funkcję vPC na przełącznikach Nexus. Pierwotna konfiguracja musi być wykonana tylko raz dla każdego przełącznika w parze. Potem już tworzenie kolejnych vPC jest relatywnie proste.

Wstępna konfiguracja vPC

Poniższe czynności musisz wykonać dla każdego z dwóch przełączników Nexus w parze tworzącej vPC. Aktualna wersja systemu operacyjnego dopuszcza jedynie vPC składające się z dwóch switchy Nexus.

Na początek włącz funkcję vPC poleceniem `feature vpc`:

```
NX-5k-1(config)# feature vpc
```

Zanim przejdziesz do dalszych czynności, powinieneś utworzyć domenę vPC. Numer domeny musi być taki sam dla wszystkich przełączników ze skonfigurowanymi kanałami vPC:

```
NX-5K-1(config-if)# vpc domain 1
```

Para przełączników połączonych funkcją vPC wymaga jeszcze dodatkowego linku, zwanego **peer keepalive link**. Jest on dodatkowy w stosunku do kanału port channel, przez który przechodzą trunki VLAN-ów. Keepalive link musi funkcjonować w warstwie trzeciej, ponieważ na każdym z jego końców stosuje się adres IP. Łącze to może działać w ramach istniejącej sieci IP, Cisco zaleca jednak jego tworzenie na poziomie interfejsu zarządzania. Jest to szczególnie ważne w przypadku przełączników z serii Nexus 5000, ponieważ nie posiadają one innych interfejsów warstwy trzeciej poza interfejsem zarządzania. Nie jestem wprawdzie fanem stosowania kryzysowych (dla obserwatora) rozwiązań do zarządzania siecią, jednak przykład pokazuje, że tak należy zrobić. VFR zarządzający jest w przełącznikach Nexus 5000 VRF-em domyślnym (bardziej szczegółowe informacje dotyczące VRF-u znajdziesz w rozdziale 18.):

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-5K-1(config-vpc)# peer-keepalive destination 10.10.10.2
```

W przypadku Nexus 7000 tym łączem może być dowolny interfejs, ponieważ wszystkie interfejsy są i tak ustawione domyślnie w trybie routingu. Oto konfiguracja połączenia punkt-punkt pomiędzy dwoma switchami Nexus 7000. Na początek NX-7K-1:

```
NX-7K-1(config-if-range)# int e3/1
NX-7K-1(config-if)# ip address 10.255.255.1/24

NX-7K-1 (config-vpc-domain)# peer-keepalive destination 10.255.255.2
```

A teraz konfiguracja drugiego przełącznika, o nazwie NX-7K-2:

```
NX-7K-2(config-if-range)# int e3/1
NX-7K-2(config-if)# ip address 10.255.255.2/24

NX-7K-2 (config-vpc-domain)# peer-keepalive destination 10.255.255.1
```

vPC wymaga, aby łącze — najlepiej trunk — pomiędzy parą vPC zostało zdefiniowane jako łącze równorzędne (ang. *peer link*). Łącze to zostanie użyte do przenoszenia ruchu, co jest niezbędne do właściwego wykorzystania EtherChannel. Generalnie łącze to jest głównym ogniwem przełączającym pomiędzy dwoma switchami. Zgodnie z dokumentacją vPC peer link musi być multi-10 Gb/s kanałem EtherChannel, choć oczywiście będzie działać tylko z jednym fizycznym łączem. Stosowana komenda jest taka sama w przełącznikach 5000 i 7000. Tutaj port channel 1 został użyty jako peer link:

```
NX-5K-1(config-if)# int po 1
NX-5K-1(config-if)# vpc peer-link

Please note that spanning tree port type is changed to "network" port type on vPC peer-link. This will
enable spanning tree Bridge Assurance on vPC peer-link provided the STP Bridge Assurance (which is
enabled by default) is not disabled.
```

// Zauważ, że port spanning tree został zmieniony na port typu „network"na łączu równorzędnym vPC.
// Uruchomi to funkcję Spanning Tree Bridge Assurance na łączu równorzędnym vPC, pod warunkiem,
// że STP Bridge Assurance (domyślnie włączone) nie zostało wcześniej wyłączone.

Kanały vPC działają zawsze w obrębie domeny vPC. Każdy przełącznik podłączony przez vPC musi mieć taką samą nazwę (numer) domeny. Pamiętaj też, że w domenie mogą być tylko dwa przełączniki. W naszym przykładzie każdy 5K jest przypisany do vPC Domain 1.

Może się to wydawać dość trudne, ale jest też dobra wiadomość. Wszystkie polecenia vPC muszą być wprowadzone tylko raz. Kiedy już para przełączników Nexus zostanie skonfigurowana do obsługi vPC, to dodawanie kolejnych kanałów vPC okaże się bardzo łatwe.

Dodawanie vPC

Kiedy wszystko inne jest już skonfigurowane, kanały EtherChannel na każdym z przełączników 5K, które mają być połączone pojedynczym vPC, muszą zostać otagowane znacznikiem vPC ID. ID możesz wybrać dowolnie, jednak dobrą praktyką jest znakowanie EtherChannel tym samym numerem, którym oznaczony jest sam kanał.

Teoretycznie mógłbyś połączyć ze sobą port channel 10 na przełączniku 5K1 z port channel 20 na 5K2 i zastosować przy tym znacznik vPC 30. Jeśli jednak tak zrobisz, to każdy inżynier sieciowy, który będzie administrował tymi urządzeniami po Tobie, będzie chciał Cię odnaleźć i udusić. Staraj się więc utrzymać spójne nazewnictwo.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Na początek konfiguracja na przełączniku NX-5K-1:

```
NX-5K-1(config)# int po 10
NX-5K-1(config-if)#vpc 10
```

I ta sama konfiguracja na NX-5K-2:

```
NX-5K-2(config)# int po 10
NX-5K-2(config-if)#vpc 10
```

Chociaż ogólna koncepcja może być trudna do ogarnięcia, to kiedy już skonfigurujesz pierwotną architekturę vPC, sposób dodawania kolejnych kanałów jest właśnie tak prosty. Istnieją pewne ograniczenia dotyczące vPC, które opisałem w rozdziale 18. Jednym z ważniejszych jest sposób podłączania modułów lub urządzeń rozszerzających infrastrukturę sieci.

Ponieważ vPC tworzy pojedyncze łącze danych pomiędzy dwoma osobnymi przełącznikami, muszą one być zgodne co do możliwości ustanowienia vPC. W tym celu switche komunikują się ze sobą i jeśli są skonfigurowane odmiennie, następuje odmowa połączenia vPC. Na szczęście istnieje szybki sposób sprawdzenia, czy występuje tego typu problem. Zastosuj polecenie `show vpc consistency-parameters`, które pozwoli wykryć niezgodności uniemożliwiające połączenie. Dodając parametr `global`, wyświetlisz te ustawienia, które powinny być dopasowane:

```
NX-5K-1# sho vpc consistency-parameters global
    Legend:
        Type 1 : vPC will be suspended in case of mismatch
                // Parametr 1 w kolumnie Type: kanał vPC zostanie zawieszony w przypadku braku zgodności ustawień
Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
QoS                         1     ([], [3], [], [], [],  ([], [3], [], [], [],
                                  [])                    [])
Network QoS (MTU)           1     (1538, 2240, 0, 0, 0,  (1538, 2240, 0, 0, 0,
                                  0)                     0)
Network Qos (Pause)         1     (F, T, F, F, F, F)     (F, T, F, F, F, F)
Input Queuing (Bandwidth)   1     (50, 50, 0, 0, 0, 0)   (50, 50, 0, 0, 0, 0)
Input Queuing (Absolute     1     (F, F, F, F, F, F)     (F, F, F, F, F, F)
Priority)
Output Queuing (Bandwidth)  1     (50, 50, 0, 0, 0, 0)   (50, 50, 0, 0, 0, 0)
Output Queuing (Absolute    1     (F, F, F, F, F, F)     (F, F, F, F, F, F)
Priority)
STP Mode                    1     Rapid-PVST             Rapid-PVST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type, Edge         1     Normal, Disabled,      Normal, Disabled,
BPDUFilter, Edge BPDUGuard        Disabled               Disabled
STP MST Simulate PVST       1     Enabled                Enabled
Allowed VLANs               -     1-10,50-90,94-97       1-10,50-90,94-97
```

Możesz też w tym poleceniu wykorzystać parametr `interface` *`nazwa-interfejsu`*, aby uzyskać informacje dotyczące zgodności interfejsu w odniesieniu do tworzenia vPC:

```
NX-5K-1# sho vpc consistency-parameters interface po4
    Legend:
        Type 1 : vPC will be suspended in case of mismatch
                // Parametr 1 w kolumnie Type: kanał vPC zostanie zawieszony w przypadku braku zgodności ustawień
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
STP Port Type               1     Edge Trunk Port        Edge Trunk Port
STP Port Guard              1     None                   None
STP MST Simulate PVST       1     Default                Default
lag-id                      1     [(1, 2-a0-98-13-6-10,  [(1, 2-a0-98-13-6-10,
                                  1, 0, 0), (7f9b,       1, 0, 0), (7f9b,
                                  0-23-4-ee-be-1, 8005,  0-23-4-ee-be-1, 8005,
                                  0, 0)]                 0, 0)]
mode                        1     active                 active
Speed                       1     10 Gb/s                10 Gb/s
Duplex                      1     full                   full
Port Mode                   1     trunk                  trunk
Native Vlan                 1     1                      1
Shut Lan                    1     No                     No
Allowed VLANs               -     3-7,9                  3-7,9
```

Jeśli chcesz sprawdzić status wszystkich kanałów vPC, wykorzystaj komendę `show vpc`:

```
NX-5K-1# sho vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
                                                  // (*) — lokalny vPC jest wyłączony, przesyłanie via vPC peer-link
vPC domain id                     : 1
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status: success
vPC role                          : secondary
vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ -------------------------------------------------
1    Po1    up     1-100,300-400
vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
2      Po2         up     success     success                    1-10,50-90
                                                                 ,94-97
3      Po12        up     success     success                    1-10,50-90
                                                                 ,94-97
4      Po13        up     success     success                    3-7,9
5      Po14        up     success     success                    3-7,9
6      Po15        up     success     success                    3-7,9
7      Po16        down*  success     success
```

Jeśli sądzisz, że dopasowanie wszystkich tych parametrów konfiguracyjnych może być bolesne, to masz rację. Jednym z największych problemów, jakie napotkałem, dotyczących vPC na przełącznikach Nexus, było działanie w środowiskach, w których konieczne są częste zmiany w konfiguracji. Pracowałem w centrum danych, gdzie VLAN-y były stale dodawane i usuwane z trunków, ponieważ kolejne wirtualne maszyny były dodawane do serwerów typu „blade". Aby dodać VLAN do trunku, stosuje się polecenie `switchport trunk allowed vlan` *`numer-vlan`*. Jak widać na poprzednich listingach, liczba dozwolonych VLAN-ów jest jednym z parametrów, których zgodność jest wymagana dla vPC. Oznacza to, że jeśli dodasz VLAN do przełącznika NX-5K-1, kanał vPC zostanie zerwany do czasu dodania tego VLAN-u również na drugim przełączniku z pary. Nawet jeśli otworzysz równolegle dwa okna SSH, po jednym do każdego przełącznika, i naciśniesz *Enter* w tym samym momencie — vPC i tak zostanie zerwane.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Poniżej podobny przykład, w którym zmieniłem parametr „spanning tree” na porcie kanału vPC. Zanim przekonfigurowałem również drugi przełącznik, vPC został zerwany i pojawił się poniższy komunikat:

```
2010 May 11 19:51:38 NX-5k-1 %ETHPORT-5-IF_ERROR_VLANS_SUSPENDED: VLANs 3-7,9 on
↪Interface port-channel14 are being suspended. (Reason: vPC type-1 configuration
↪incompatible - STP interface port type inconsistent)
        // VLAN-y 3-7,9 na interfejsie port channel14 zostały zawieszone. (Powód: konfiguracja
        // vPC typ-1 wykazuje brak kompatybilności — typ portu interfejsu STP jest niewłaściwy)
```

Odpowiedzią Cisco na występowanie tego problemu jest funkcja **config-sync**. Więcej informacji na ten temat znajdziesz w rozdziale 18.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 8.

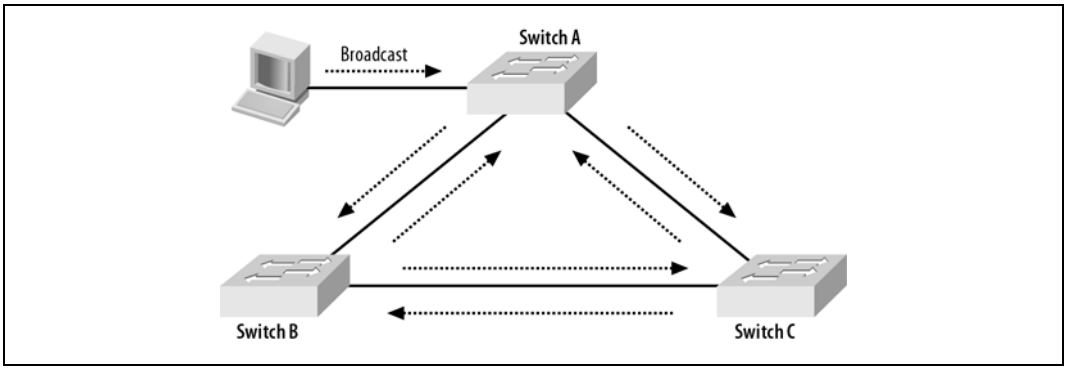
Spanning Tree

Protokół STP (ang. *Spanning Tree Protocol*), zwany po polsku protokołem drzewa rozpinającego, jest stosowany do eliminowania problemu pętli powstających w warstwie drugiej w sieciach lokalnych. Jak przeczytasz w tym rozdziale, pętle warstwy drugiej mogą powodować spustoszenie.

Spanning Tree został zaprojektowany, aby zapobiegać powstawaniu pętli pomiędzy mostkami. **Mostek** (ang. *bridge*) to urządzenie, które łączy wiele segmentów sieci, tworząc jedną domenę kolizyjną. Przełączniki są traktowane jak mostki — huby nie. O ile dokumentacja Spanning Tree zawsze odnosi się ogólnie do mostków, o tyle ja w przykładach wykorzystam przełączniki. To właśnie przełączniki są zasadniczo tymi urządzeniami, których dotyczy kwestia Spanning Tree.

Kiedy switch otrzymuje pakiet broadcastowy (inaczej rozsiewczy lub rozgłoszeniowy), przekazuje go na każdy port (z wyjątkiem portu źródłowego). W przypadku powstania pętli broadcasty są powielane w nieskończoność. Efektem tego jest **burza broadcastowa** (także rozsiewcza lub rozgłoszeniowa), która w krótkim czasie prowadzi do zatrzymania działania całej sieci.

Rysunek 8.1 ilustruje, co się dzieje, jeśli w sieci powstanie pętla.



Rysunek 8.1. Burza broadcastowa

Komputer podłączony do Switcha A wysyła ramkę rozgłoszeniową. Switch A powiela tę ramkę do Switchy B i C. Z kolei Switch B przekazuje broadcast do Switcha C, a ten powiela

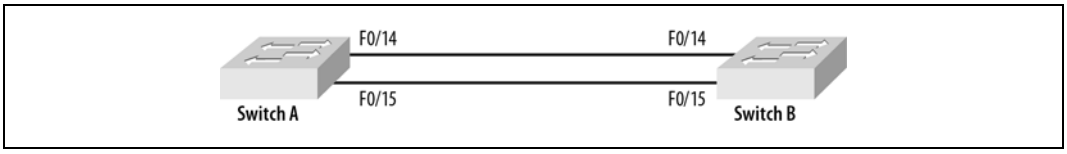
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

go do Switcha B; Switche B i C powielają również broadcast z powrotem do Switcha A. Następnie Switch A powiela do Switcha C broadcast otrzymany od Switcha B oraz do Switcha B broadcast otrzymany od Switcha C. Ta progresja może się ciągnąć w nieskończoność, dopóki pętla nie zostanie w jakiś sposób przerwana. Spanning Tree to automatyczny mechanizm wykorzystywany do wykrywania i przerywania tego typu pętli.

Protokół Spanning Tree został opracowany przez dr Radię Perlman z Sun Microsystems, która podsumowała tę ideę w wierszu zatytułowanym *Algorhyme*, będącym parafrazą utworu *Trees* Joyce'a Kilmera. Wiersz jest dostępny na stronie *http://etherealmind.com/algorhyme-radia-perlman*.

Burza broadcastowa

W sieci przedstawionej na rysunku 8.2 wystąpiła prosta pętla między dwoma przełącznikami. Switche A i B są połączone ze sobą dwoma linkami: F0/14 i F0/15 na Switchu A podpięte są do analogicznych portów na Switchu B. Wyłączyłem Spanning Tree, który jest domyślnie uruchomiony, aby zademonstrować potęgę burzy broadcastowej. Oba porty są interfejsami trunkowymi. Do pozostałych portów tych dwóch przełączników podpięte są różne urządzenia generujące standardowe pakiety rozgłoszeniowe (takie jak ARP czy DHCP). Sieć nie wyróżnia się niczym szczególnym oprócz tego, że wyłączony jest protokół drzewa rozpinającego.



Rysunek 8.2. Prosta pętla warstwy drugiej

Interfejs F0/15 został już skonfigurowany i działa poprawnie. Po wpisaniu polecenia `show interface f0/15` otrzymasz listing, który pokazuje, że ilość danych wchodzących (ang. *input rate*) oraz wychodzących (ang. *output rate*) jest dość mała (obydwie wartości w przybliżeniu 1000 bitów na sekundę oraz 2 – 3 pakiety na sekundę):

```
3550-IOS#sho int f0/15
FastEthernet0/15 is up, line protocol is up (connected)
  Hardware is Fast Ethernet, address is 000f.8f5c.5a0f (bia 000f.8f5c.5a0f)
  MTU 1500 bytes, BW 100000 Kbit, DLY 100 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  Full-duplex, 100Mb/s, media type is 10/100BaseTX
  input flow-control is off, output flow-control is unsupported
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:00:10, output 00:00:00, output hang never
  Last clearing of "show interface" counters never
  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  5 minute input rate 1000 bits/sec, 2 packets/sec
  5 minute output rate 1000 bits/sec, 3 packets/sec
     5778444 packets input, 427859448 bytes, 0 no buffer
     Received 5707586 broadcasts (0 multicast)
     0 runts, 0 giants, 0 throttles
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
     0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
     0 watchdog, 5707585 multicast, 0 pause input
     0 input packets with dribble condition detected
     2597516 packets output, 213866427 bytes, 0 underruns
```

Przy diagnozowaniu problemów związanych z burzą broadcastową przydatnym narzędziem jest polecenie `show processes cpu history`. Wyświetla ono histogram ASCII przedstawiający użycie procesora w ciągu ostatnich 72 godzin. Histogram zawiera 3 wykresy:

- Wykorzystanie procesora w procentach na sekundę (ostatnie 60 sekund) — „CPU percent per second (last 60 seconds)".
- Wykorzystanie procesora w procentach na minutę (ostatnie 60 minut) — „CPU percent per minute (last 60 minutes)".
- Wykorzystanie procesora w procentach na godzinę (ostatnie 72 godziny) — „CPU percent per hour (last 72 hours)".

Poniżej dane otrzymane po wprowadzeniu na Switchu B komendy `show process cpu history`. Wykres pokazuje wykorzystanie procesora na poziomie 0 – 3 procent w ciągu ostatniej minuty (pozostałe wykresy zostały usunięte dla zwiększenia czytelności):

```
3550-IOS#sho proc cpu history
      11111          33333          11111                    1
100
 90
 80
 70
 60
 50
 40
 30
 20
 10
   0....5....1....1....2....2....3....3....4....4....5....5....
             0    5    0    5    0    5    0    5    0    5
              CPU% per second (last 60 seconds)
```

Liczby po lewej stronie wykresu (w osi pionowej) to poziom procentowego użycia CPU. Na dole wykresu (na osi poziomej) przedstawiono upływ czasu w sekundach (0 = moment wykonania komendy). Liczby na górze wykresu pokazują całkowite wartości określające wykorzystanie procesora w danym momencie. Zgodnie z wykresem użycie procesora wzrosło od 0 do 1 procenta w piątej sekundzie, a następnie do 3 procent w dwudziestej sekundzie. Jeśli wartości wzrosłyby powyżej 10 procent, zobaczyłbyś wizualizację tych szczytów bezpośrednio na wykresie.

Jest to przełącznik 3550 obsługujący dynamiczny routing w sieci (w tym przypadku EIGRP). Dlatego też jest ważnym urządzeniem zapewniającym funkcjonalność warstwy trzeciej:

```
3550-IOS#sho ip eigrp neighbors
IP-EIGRP neighbors for process 55
H   Address              Interface     Hold Uptime   SRTT   RTO  Q    Seq Type
                                       (sec)         (ms)        Cnt  Num
0   10.55.1.10           Fa0/13        14 00:25:30      1   200  0    27
2   10.55.10.3           Vl10          13 1w0d         18   200  0    25
```

Teraz ustawiam interfejs F0/14 jako trunk:

```
3550-IOS(config)#int f0/14
3550-IOS(config-if)# switchport
3550-IOS(config-if)# switchport trunk encapsulation dot1q
3550-IOS(config-if)# switchport mode trunk
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W tym momencie istnieją dwa trunki łączące Switche A i B. Pamiętaj, że wyłączyłem Spanning Tree. Kilka sekund po tym, gdy przekształciłem F0/14 w kanał trunkowy, ilość danych wchodzących i wychodzących na interfejsie F0/15 skoczyła z 1000 bitów na sekundę i 2 – 3 pakiety na sekundę do 815 000 bitów na sekundę i 1561 pakietów na sekundę:

```
3550-IOS#sho int f0/15 | include minute
  5 minute input rate 815000 bits/sec, 1565 packets/sec
  5 minute output rate 812000 bits/sec, 1561 packets/sec
```

Dziesięć sekund później wartości te wynoszą już ponad 2,7 Mb/s i 4500 pakietów na sekundę:

```
3550-IOS#sho int f0/15 | include minute
  5 minute input rate 2744000 bits/sec, 4591 packets/sec
  5 minute output rate 2741000 bits/sec, 4587 packets/sec
```

Teraz zaczynam na konsoli otrzymywać komunikaty ostrzegawcze. Sąsiedzi EIGRP zaczynają odbijać zapytania:

```
1w0d: %DUAL-5-NBRCHANGE: IP-EIGRP(0) 55: Neighbor 10.55.1.10 (FastEthernet0/13) is down: holding time
➥expire
1w0d: %DUAL-5-NBRCHANGE: IP-EIGRP(0) 55: Neighbor 10.55.1.10 (FastEthernet0/13) is up: new adjacency
1w0d: %DUAL-5-NBRCHANGE: IP-EIGRP(0) 55: Neighbor 10.55.10.3 (Vlan10) is down: retry limit exceeded
1w0d: %DUAL-5-NBRCHANGE: IP-EIGRP(0) 55: Neighbor 10.55.10.3 (Vlan10) is up: new adjacency
```

Szybki rzut oka na statystyki wykorzystania procesora z ostatniej minuty wszystko wyjaśnia. Histogram pokazuje, że przez ostatnie 25 sekund użycie procesora było na poziomie 99%. Problem zaczął się 40 sekund temu, kiedy uruchomiłem drugi trunk:

```
3550-IOS#sho proc cpu hist
     9999999999999999999999999999333339999966666
     9999999999999999999999999999999999999333311111
100  *************************     *****
 90  *************************     *****
 80  *************************     *****
 70  *************************     *****
 60  *************************     **********
 50  *************************     **********
 40  ****************************************
 30  ****************************************
 20  ****************************************
 10  ****************************************
    0....5....1....1....2....2....3....3....4....4....5....5....
              0    5    0    5    0    5    0    5    0    5
               CPU% per second (last 60 seconds)
```

Użycie procesora skoczyło gwałtownie do 99% w przeciągu kilku sekund od momentu powstania pętli. Przełącznik wykorzystuje większość swojej mocy obliczeniowej na przesyłanie broadcastów w pętli. W rezultacie cierpią na tym wszystkie inne procesy. Zastosowanie telnetu lub SSH do administrowania przełącznikiem jest już w tej chwili praktycznie niemożliwe. Nawet konsola zaczyna wykazywać objawy spowolnienia działania.

W ciągu pięciu minut poziom wykorzystania obydwu łączy osiągnął 50%. Wysyłają i otrzymują ponad 70 000 pakietów na sekundę:

```
3550-IOS#sho int f0/15
FastEthernet0/15 is up, line protocol is up (connected)
  Hardware is Fast Ethernet, address is 000f.8f5c.5a0f (bia 000f.8f5c.5a0f)
  MTU 1500 bytes, BW 100000 Kbit, DLY 100 usec,
     reliability 255/255, txload 143/255, rxload 143/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Full-duplex, 100Mb/s, media type is 10/100BaseTX
input flow-control is off, output flow-control is unsupported
ARP type: ARPA, ARP Timeout 04:00:00
Last input 00:00:00, output 00:00:00, output hang never
Last clearing of "show interface" counters never
Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
Queueing strategy: fifo
Output queue: 0/40 (size/max)
5 minute input rate 56185000 bits/sec, 71160 packets/sec
5 minute output rate 56277000 bits/sec, 70608 packets/sec
   48383882 packets input, 112738864 bytes, 0 no buffer
   Received 48311185 broadcasts (0 multicast)
   0 runts, 0 giants, 0 throttles
   0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
   0 watchdog, 36608855 multicast, 0 pause input
   0 input packets with dribble condition detected
   45107032 packets output, 4212355164 bytes, 356 underruns
   0 output errors, 0 collisions, 1 interface resets
   0 babbles, 0 late collision, 0 deferred
   0 lost carrier, 0 no carrier, 0 PAUSE output
   361 output buffer failures, 0 output buffers swapped out
```

Pamiętaj, że pierwotnym obciążeniem dla tych interfejsów, będącym punktem odniesienia, jest 1000 bitów na sekundę oraz 2 – 3 pakiety na sekundę. Przy obecnym natężeniu ruchu 99,9% broadcastów to pakiety przetwarzane po raz kolejny. To jest właśnie burza broadcastowa. Możesz być pewien, że doświadczenie tego problemu na własnej skórze — szczególnie w przypadku sieci, za którą jesteś odpowiedzialny — raczej nie jest zabawne. Burza pojawia się nagle i natychmiast dotyka wszystkie usługi w sieci.

Ponieważ nasz eksperyment miał miejsce w środowisku laboratoryjnym, to dobrze się bawiłem (chociaż moja żona z pewnością miałaby jakieś ciekawe uwagi odnośnie do mojego poczucia humoru). Jednak wszystko, co dobre, szybko się kończy. W ciągu kilku sekund od wyłączenia jednego z trunków EIGRP się stabilizuje, a problemy działania sieci znikają. Mogę teraz bez problemu skomunikować się z przełącznikiem poprzez telnet. Poziom wykorzystania procesora prawie natychmiast wraca do standardowych wartości 0 – 1%:

```
3550-IOS(config)#int f0/14
3550-IOS(config-if)# no switchport
3550-IOS# sho proc cpu hist
                 666669999999999999999999999999999999999999999999
            111111111199999999999999999999999999999999999999999
100                 *********************************************
 90                 *********************************************
 80                 *********************************************
 70                 *********************************************
 60             ************************************************
 50             ************************************************
 40             ************************************************
 30             ************************************************
 20             ************************************************
 10             ************************************************
    0....5....1....1....2....2....3....3....4....4....5....5....
              0    5    0    5    0    5    0    5    0    5
              CPU% per second (last 60 seconds)
```

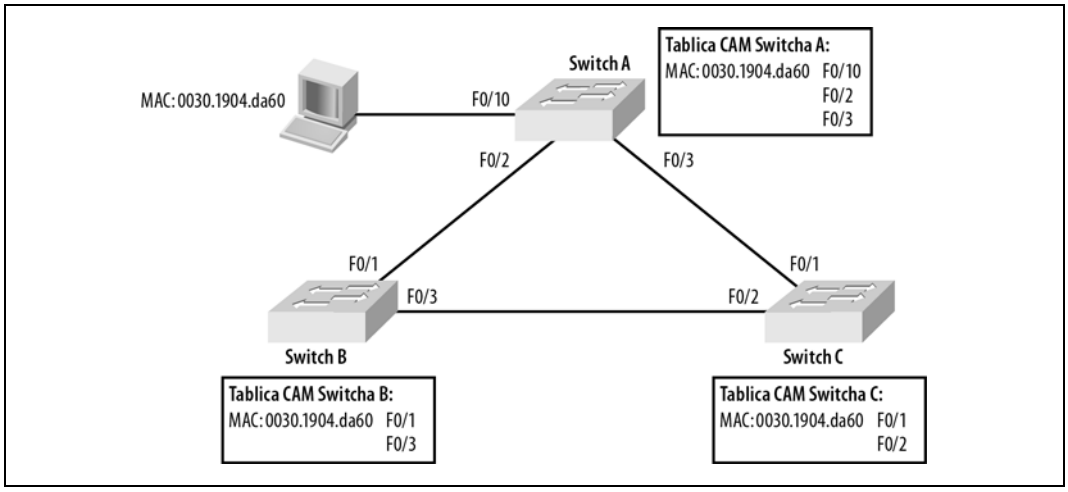
Ten przykład pokazuje, jak szkodliwa w skutkach może być burza rozgłoszeniowa. Kiedy przełączniki w jej obszarze przestają odpowiadać, zdiagnozowanie burzy może stać się bardzo trudne. W przypadku braku dostępu do przełączników poprzez konsolę, SSH czy telnet jedynym sposobem na przerwanie pętli jest odcięcie linków powodujących problem. Jeśli będziesz miał

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

szczęście, diody aktywności portów tworzących pętle będą migać z większą częstotliwością niż diody pozostałych portów. Tak czy inaczej, nie będzie to Twój dobry dzień.

Niestabilność tablicy adresów MAC

Kolejnym problemem związanym z powstaniem pętli jest nieustanne aktualizowanie tablic adresów MAC (tablic CAM w przypadku CatOS). Przyjrzyj się przykładowi sieci przedstawionej na rysunku 8.3.



Rysunek 8.3. Niespójności tablicy adresów MAC

Jeśli wszystkie przełączniki są połączone między sobą, a protokół Spanning Tree jest wyłączony, Switch A będzie zakładał, że źródłem adresu MAC komputera podłączonego bezpośrednio do niego jest inny przełącznik. Dochodzi do tego bardzo szybko podczas burzy broadcastowej, a w rzadkich przypadkach, kiedy spostrzeżesz takie zachowanie bez burzy broadcastowej, będzie to oznaczać, że już za chwilę sprawy potoczą się bardzo złym torem.

W przypadku tej sieci z wyłączonym Spanning Tree szukałem adresu MAC komputera na Switchu A, korzystając z komendy `show mac-address-table | include 0030.1904.da60`. Powtarzałem to polecenie kilkakrotnie, tak szybko, jak potrafiłem, i oto rezultat:

```
3550-IOS#sho mac-address-table | include 0030.1904.da60
   1    0030.1904.da60    DYNAMIC     Fa0/10
3550-IOS# sho mac-address-table | include 0030.1904.da60
   1    0030.1904.da60    DYNAMIC     Fa0/10
3550-IOS# sho mac-address-table | include 0030.1904.da60
   1    0030.1904.da60    DYNAMIC     Fa0/2
3550-IOS# sho mac-address-table | include 0030.1904.da60
   1    0030.1904.da60    DYNAMIC     Fa0/3
3550-IOS# sho mac-address-table | include 0030.1904.da60
   1    0030.1904.da60    DYNAMIC     Fa0/2
3550-IOS# sho mac-address-table | include 0030.1904.da60
```

Przełącznik ten jest połączony bezpośrednio z urządzeniem, o które pytałem, wpisując komendę. Jednak za którymś razem switch zaczyna informować, że najlepsza ścieżka do urządzenia prowadzi przez Switch B lub C.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Pamiętaj, że switch analizuje każdy pakiet otrzymywany na danym porcie i przypisuje źródłowy adres MAC pakietu do tego portu w tablicy adresów MAC/tablicy CAM. Podłączane do przełącznika urządzenia mogą być przenoszone do innej lokalizacji i często tak się dzieje. Dlatego przełącznik „zakłada", że ostatni port, na którym zarejestrowany został adres MAC urządzenia, jest miejscem, do którego adres przynależy. Ponieważ pakiety broadcastowe z urządzenia są nieustannie w obiegu zapętlonej sieci, za każdym razem, kiedy switch otrzyma taki pakiet na jeden z portów, będzie zakładał, że tam przynależy dany adres MAC.

Zapobieganie powstawaniu pętli przez Spanning Tree

Oczywistym sposobem zapobiegania powstawaniu pętli jest przestrzeganie tej samej rady, którą dałby Ci lekarz, gdybyś się poskarżył: „boli mnie, kiedy to robię" — więc tego nie rób! W rzeczywistych warunkach istnieje jednak wiele zmiennych, na które nie masz żadnego wpływu. Nieraz byłem świadkiem tego, że jakaś sieć padła, ponieważ ktoś postanowił spiąć ze sobą pod biurkiem dwie nieobsadzone końcówki kabli za pomocą małego przełącznika przyniesionego z domu. Do cholery, widziałem nawet, jak robią to sami administratorzy sieci.

Spinanie ze sobą przełączników więcej niż jednym linkiem jest dobrym pomysłem umożliwiającym redundancję łączy — jest wręcz zalecane. Sztuka polega na tym, żeby w danym momencie tylko jedno z łączy było aktywne (lub aby obydwa z nich tworzyły jeden logiczny link). Jeśli połączysz przełączniki dwoma linkami, ale jeden z nich wyłączysz, zlikwidujesz ryzyko powstania pętli. Jednak gdy aktywny link zawiedzie, będziesz musiał ręcznie uruchomić drugi.

Spanning Tree jest protokołem zaprojektowanym w celu wykrywania pętli w sieci i przerywania ich, zanim spowodują wymierne szkody. Prawidłowo skonfigurowany, Spanning Tree jest doskonałym narzędziem, które powinno być stosowane w każdej sieci. Skonfigurowany niepoprawnie może jednak stać się przyczyną pewnych problemów, które trudno będzie zdiagnozować.

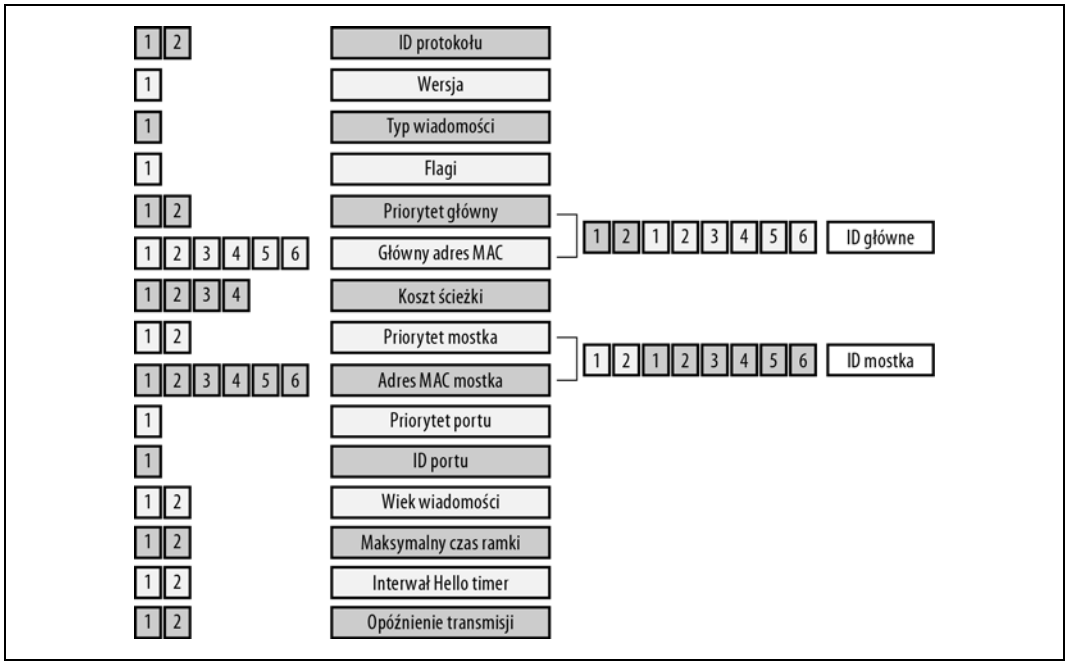
Jak działa Spanning Tree

Spanning Tree wybiera w sieci jeden z mostków (przełączników), który będzie **mostkiem głównym** (ang. *root bridge*). W założeniu pozostałe mostki muszą łączyć się z mostkiem głównym najkrótszą z możliwych ścieżek. Spanning Tree analizuje koszty możliwych ścieżek pomiędzy każdym z mostków a mostkiem głównym. Najlepsza z nich, o najniższym koszcie, jest utrzymywana jako aktywne łącze. Pozostałe ścieżki są zrywane. STP zrywa ścieżki przez **zablokowanie** (ang. *blocking*) odpowiednich portów.

Każdy mostek w sieci obsługujący protokół Spanning Tree wysyła co dwie sekundy ramki **BPDU** (ang. *bridge protocol data units*). Format ramki BPDU przedstawia rysunek 8.4.

Ramki te zawierają informacje niezbędne każdemu przełącznikowi w sieci do wykonywania poniższych zadań:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 8.4. Format ramki BPDU

Wybór mostka głównego

Kiedy przełącznik jest uruchamiany, przyjmuje założenie, że jest mostkiem głównym. Dopisuje wtedy ID główne do ID mostka lokalnego w wysyłanych ramkach BPDU. Jeśli otrzyma od innego przełącznika ramkę BPDU, która ma niższe ID główne, uzna go za przełącznik główny (ang. *root switch*). Lokalny przełącznik zaczyna wtedy używać właśnie tego ID głównego (niższego) w wysyłanych ramkach BPDU.

Każdy mostek posiada **ID mostka** (ang. *bridge ID*). ID mostka jest kombinacją priorytetu mostka i jego adresu MAC. Priorytet mostka jest konfigurowalnym dwubajtowym polem o domyślnej wartości 32768. Im mniejsza wartość ID mostka, tym większe prawdopodobieństwo, że zostanie on mostkiem głównym (ang. *root bridge*) (mostkiem głównym zostaje ten o najniższym ID mostka).

ID główne (*ang. root ID*) jest analogicznie zbudowane z dwóch pól: priorytet główny (ang. *root priority*) oraz główny adres MAC (ang. *root MAC address*). Pole priorytet główny jest konfigurowalne i posiada domyślną wartość 32768 (0×8000). Jeśli priorytety główne są ze sobą związane, niższy główny adres MAC jest używany do rozłączenia priorytetów.

Określenie najlepszej ścieżki do mostka głównego

Jeśli ramki BPDU z mostka głównego otrzymywane są na więcej niż jednym porcie przełacznika, oznacza to, że istnieje kilka ścieżek do mostka głównego. Najlepsza ścieżka prowadzi przez ten port, na którym odebrane zostały ramki BPDU z najniższym **kosztem ścieżki** (ang. *path cost*).

Koszty ścieżki są ustalane przez dodawanie priorytetu portu każdego mostka do początkowego kosztu ścieżki, kiedy ramka wędruje od mostka do mostka.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ustalenie portu głównego na każdym mostku

Port główny (ang. *root port*) to ten port na przełączniku, od którego prowadzi najkrótsza ścieżka do mostka głównego. Mostek główny nie posiada portów głównych; ma jedynie porty wyznaczone.

Określenie portu wyznaczonego w każdym segmencie

Port wyznaczony (ang. *designated port*) to ten port w danym segmencie sieci, który ma najkrótszą ścieżkę do mostka głównego. W segmentach, które są bezpośrednio podłączone do mostka głównego, to jego porty są portami wyznaczonymi.

Wybór mostka wyznaczonego w każdym segmencie

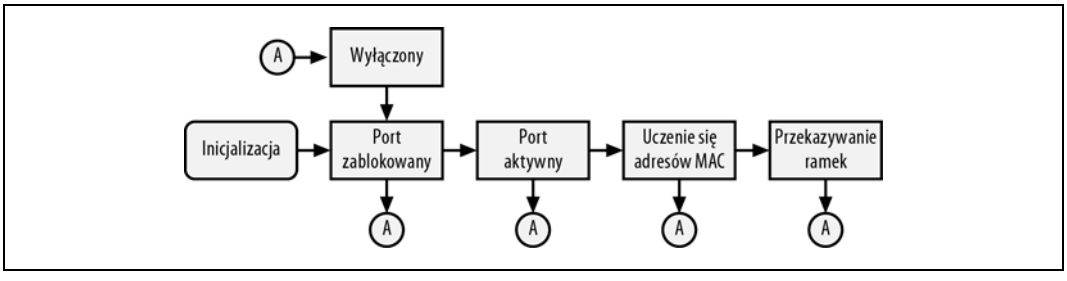
W danym segmencie mostek, który posiada porty wyznaczone, jest traktowany jak **mostek wyznaczony** (ang. *designated bridge*). Mostek główny jest również mostkiem wyznaczonym dla wszystkich segmentów podłączonych bezpośrednio. W przypadku gdy dwa mostki w segmencie posiadają porty główne, mostek z najniższym ID mostka zostaje mostkiem wyznaczonym.

Blokowanie nietransmitujących portów

Porty, które otrzymały ramki BPDU, a nie są ani portami wyznaczonymi, ani głównymi, są ustawiane w trybie **blokowania** (ang. *blocking*). Pod względem administracyjnym są one włączone, nie mogą jednak przekazywać pakietów (mimo to nadal odbierają i przetwarzają ramki BPDU).

Zawsze konfiguruj przełącznik tak, aby był mostkiem głównym. Umożliwienie przełącznikom autokonfiguracji jest niebezpieczne, gdyż wybiorą one z reguły switch o najniższym adresie MAC, który w większości przypadków będzie niewłaściwy. Główną zasadą, którą powinieneś stosować, jest niedopuszczanie, aby urządzenia sieciowe podejmowały kluczowe decyzje na podstawie wartości domyślnych. Spowoduje to, że sieć będzie się zachowywać w sposób nieoczekiwany, a Ty oblejesz egzaminy na certyfikaty wyższego poziomu. Są one tak skonstruowane, żeby złapać Cię właśnie na takich błędach. Zazwyczaj wybór urządzenia, które ma być mostkiem głównym, jest dość oczywisty. Najczęściej mostkiem głównym powinien być jeden z przełączników warstwy rdzenia w schemacie Twojej sieci.

Każdy port na przełączniku od momentu podłączenia może przyjmować różne stany protokołu Spanning Tree, które przedstawia schemat blokowy z rysunku 8.5. Stany te układają się w określony wzór w zależności od informacji uzyskanych z ramek BPDU odbieranych na porcie.



Rysunek 8.5. Stany portu w protokole Spanning Tree

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Oto stany portów protokołu Spanning Tree:

Inicjalizacja (ang. *initializing*)
Port jest w stanie inicjalizacji, jeśli właśnie zostało włączone zasilanie lub port został przestawiony przez administratora ze stanu wyłączonego do włączonego.

Port zablokowany (ang. *blocking*)
Taki port jest w zasadzie nieużywany. Nie przekazuje i nie odbiera ramek poza pewnymi wyjątkami:

- port odbiera i przetwarza ramki BPDU;
- port odbiera wiadomości związane z zarządzaniem siecią i odpowiada na nie.

Port aktywny (ang. *listening*)
Port aktywny zachowuje się podobnie do portu zablokowanego, z tym wyjątkiem, że zarówno odbiera, jak i wysyła ramki BPDU. Przekazywanie pakietów jest nadal zabronione. Port nie uczy się też żadnych adresów.

Uczenie się adresów MAC (ang. *learning*)
W tym stanie port nadal nie przekazuje ramek. Analizuje jednak ramki, które odbiera, aby odczytać ich adresy MAC i zapisać je w tablicy adresów MAC/tablicy CAM. Po przeanalizowaniu ramki są odrzucane (ang. *discarded*).

Przekazywanie ramek (ang. *forwarding*)
Ten stan wielu określa jako stan „normalny". Port odbiera i wysyła ramki BPDU, analizuje przychodzące pakiety pod kątem adresów MAC i przekazuje ramki otrzymane z portów innych przełączników. Kiedy port jest w stanie przekazywania ramek, urządzenie lub sieć do niego podłączona są aktywne i mogą się komunikować.

Port wyłączony (ang. *disabled*)
Port wyłączony nie przekazuje ramek i nie uczestniczy w Spanning Tree. Odbiera i odpowiada jedynie na wiadomości związane z zarządzaniem siecią.

Per-VLAN Spanning Tree

Ponieważ VLAN-y mogą być przycinane na trunkach (omówiłem to w rozdziale 6.), możliwe jest, że niektóre z nich będą formować pętle, a inne nie. Z tego powodu przełączniki Cisco stosują teraz domyślnie odmianę Spanning Tree obsługującą VLAN-y, zwaną Per-VLAN Spanning Tree (PVST). PVST umożliwia uruchamianie oddzielnej instancji Spanning Tree dla każdego VLAN-u przy wykorzystaniu trunków ISL. Per-VLAN Spanning Tree Plus (PVST+) oferuje tę samą funkcję dla protokołu trunkingu 802.1Q.

Domyślnie wszystkie VLAN-y odziedziczą te same wartości dla wszystkich konfiguracji Spanning Tree. Każdy VLAN może być jednak skonfigurowany inaczej. Przykładowo każdy VLAN może mieć inny mostek główny Spanning Tree. Więcej szczegółów na temat różnych typów Spanning Tree w przełącznikach Cisco zamieściłem na końcu tego rozdziału.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zarządzanie Spanning Tree

Protokół drzewa rozpinającego jest domyślnie włączony. Aby podejrzeć jego status, użyj polecenia `show spanning-tree` dla IOS i NX-OS:

```
Cat-3550#sho spanning-tree
VLAN0001
  Spanning tree enabled protocol idee
  Root ID    Priority    24577
             Address     0009.43b5.0f80
             Cost        23
             Port        20 (FastEthernet0/20)
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
  Bridge ID  Priority    32769  (priority 32768 sys-id-ext 1)
             Address     000f.8f5c.5a00
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
             Aging Time 300
Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Fa0/13           Altn BLK 19        128.13   P2p
Fa0/14           Altn BLK 19        128.14   P2p
Fa0/15           Altn BLK 19        128.15   P2p
Fa0/20           Root FWD 19        128.20   P2p
Fa0/23           Desg FWD 19        128.23   P2p
[fragment usunięty]
```

Wytłuszczone fragmenty listingu pokazują priorytet i adres MAC mostka głównego, a także port przełącznika, który tworzy ścieżkę dostępu (port główny). To bardzo przydatne informacje, kiedy próbujesz zlokalizować w sieci mostek główny. Wykonując tę komendę na każdym przełączniku w sieci, będziesz mógł sporządzić mapę połączeń i zorientować się, który przełącznik jest główny.

Poniżej listing z wykonania polecenia `show spanning-tree` na przełączniku Nexus 7000, który działa jako mostek główny STP:

```
NX-7K-1-Cozy# sho spanning
VLAN0001
  Spanning tree enabled protocol rstp
  Root ID    Priority    32769
             Address     0026.9807.95c2
             This bridge is the root
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec
  Bridge ID  Priority    32769  (priority 32768 sys-id-ext 1)
             Address     0026.9807.95c2
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Po1              Desg FWD 1         128.4096 (vPC peer-link) Network P2p
Po10             Desg FWD 1         128.4105 (vPC) P2p
[fragment usunięty]
```

Szczegółowe informacje dotyczące przełącznika umieszczone są poniżej informacji głównych. Pod nimi z kolei znajdują się dane dotyczące każdego portu przełącznika aktywnie uczestniczącego w Spanning Tree. Są one powtórzone dla każdego VLAN-u.

W CatOS odpowiednikiem tego polecenia jest `show spantree`. Wyświetlone zostaną bardzo podobne informacje w nieco innym układzie:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
CatOS-6509: (enable)sho spantree
VLAN 1
Spanning tree mode          RAPID-PVST+
Spanning tree type          ieee
Spanning tree enabled
Designated Root             00-00-00-00-00-00
Designated Root Priority    0
Designated Root Cost        0
Designated Root Port        1/0
Root Max Age   0  sec   Hello Time 0  sec   Forward Delay 0  sec
Bridge ID MAC ADDR          00-00-00-00-00-00
Bridge ID Priority          32768
Bridge Max Age 20 sec   Hello Time 2  sec   Forward Delay 15 sec
Port                      State         Role Cost      Prio Type
------------------------ ------------- ---- --------- ---- --------------------
 1/1                      not-connected  -           4   32
 1/2                      not-connected  -           4   32
 2/1                      not-connected  -           4   32
 2/2                      not-connected  -           4   32
[fragment usunięty]
```

Zauważ, że wyznaczony adres MAC główny jest przedstawiony w postaci samych zer. Oznacza to, że przełącznik uważa się za mostek główny.

W IOS, aby uzyskać podsumowanie Spanning Tree, wprowadź polecenie `show spanning-tree summary`. Wyświetlony zostanie status funkcji takich jak UplinkFast i BackboneFast (omówię je w następnej sekcji):

```
Cat-3550#sho spanning-tree summary
Switch is in pvst mode
Root bridge for: VLAN0002, VLAN0200
Extended system ID           is enabled
Portfast Default             is disabled
PortFast BPDU Guard Default  is disabled
Portfast BPDU Filter Default is disabled
Loopguard Default            is disabled
EtherChannel misconfig guard is enabled
UplinkFast                   is disabled
BackboneFast                 is disabled
Configured Pathcost method used is short
Name                   Blocking Listening Learning Forwarding STP Active
---------------------- -------- --------- -------- ---------- ----------
VLAN0001                     3         0        0          2          5
VLAN0002                     0         0        0          5          5
VLAN0003                     2         0        0          2          4
VLAN0004                     2         0        0          2          4
VLAN0010                     2         0        0          2          4
VLAN0100                     2         0        0          2          4
VLAN0200                     0         0        0          4          4
---------------------- -------- --------- -------- ---------- ----------
7 vlans                     11         0        0         19         30
```

Teraz ta sama komenda na przykładzie Nexus 5000:

```
NX-5K-1# sho spanning-tree summ
Switch is in rapid-pvst mode
Root bridge for: none
Port Type Default                        is disable
Edge Port [PortFast] BPDU Guard Default  is disabled
Edge Port [PortFast] BPDU Filter Default is disabled
Bridge Assurance                         is enabled
Loopguard Default                        is disabled
Pathcost method used                     is short
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Name                 Blocking Listening Learning Forwarding STP Active
-------------------- -------- --------- -------- ---------- ----------
VLAN0001                    0         0        0          1          1
VLAN0100                    0         0        0          2          2
-------------------- -------- --------- -------- ---------- ----------
2 vlans                     0         0        0          3          3
```

Dla CatOS komenda podsumowująca to `show spantree summary`:

```
CatOS-6509: (enable)sho spantree summ
Spanning tree mode: RAPID-PVST+
Runtime MAC address reduction: disabled
Configured MAC address reduction: disabled
Root switch for vlans: 20.
Global loopguard is disabled on the switch.
Global portfast is disabled on the switch.
BPDU skewing detection disabled for the bridge.
BPDU skewed for vlans:  none.
Portfast bpdu-guard disabled for bridge.
Portfast bpdu-filter disabled for bridge.
Uplinkfast disabled for bridge.
Backbonefast disabled for bridge.

Summary of connected spanning tree ports by vlan

VLAN  Blocking Listening Learning Forwarding STP Active
----- -------- --------- -------- ---------- ----------
  20         0         0        0          1          1
      Blocking Listening Learning Forwarding STP Active
----- -------- --------- -------- ---------- ----------
Total        0         0        0          1          1
```

Doskonałym poleceniem IOS jest `show spanning-tree root`, które podaje informacje dotyczące mostka głównego dla każdego VLAN-u:

```
Cat-3550#sho spanning-tree root
                                        Root    Hello Max Fwd
Vlan                 Root ID            Cost    Time  Age Dly  Root Port
---------------- -------------------- --------- ----- --- ---  ------------
VLAN0001         24577 0009.43b5.0f80        23    2   20  15  Fa0/20
VLAN0002         32770 000f.8f5c.5a00         0    2   20  15
VLAN0003         32771 000d.edc2.0000        19    2   20  15  Fa0/13
VLAN0004         32772 000d.edc2.0000        19    2   20  15  Fa0/13
VLAN0010         32778 000d.edc2.0000        19    2   20  15  Fa0/13
VLAN0100         32868 000d.edc2.0000        19    2   20  15  Fa0/13
VLAN0200         32968 000f.8f5c.5a00         0    2   20  15
```

I to samo polecenie dla Nexus 7000:

```
NX-7K-1-Cozy# sho spanning-tree root
                                        Root  Hello Max Fwd
Vlan                 Root ID            Cost  Time  Age Dly  Root Port
---------------- -------------------- ------- ----- --- ---  ----------------
VLAN0001         32769 0026.9807.95c2       0    2   20  15  This bridge is root
VLAN0010         32778 0026.9807.95c2       0    2   20  15  This bridge is root
```

Dla systemu CatOS nie istnieje analogiczna komenda.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Dodatkowe funkcje Spanning Tree

Pierwotnie Spanning Tree był projektowany dla mostków z niewielką liczbą portów. Wraz z pojawieniem się przełączników ethernetowych do protokołu wprowadzono pewne zmiany. Pomogły one w widoczny sposób usprawnić działanie Spanning Tree przez zmniejszenie czasu oczekiwania hosta na port oraz zmniejszenie czasu konwergencji w warstwie drugiej sieci.

PortFast

PortFast to funkcja przełączników Cisco, która umożliwia wybranemu portowi pominięcie kilku stanów protokołu Spanning Tree (zobacz rysunek 8.5) i przejście od razu do stanu przekazywania ramek. PortFast należy uruchamiać jedynie na portach, do których nie będą podłączone żadne przełączniki. STP potrzebuje około 30 sekund, aby przesunąć normalny port do stanu przekazywania ramek. W przypadku systemów korzystających z DHCP może to spowodować przekroczenie czasu (ang. *time out*) i w rezultacie nieprzydzielenie adresu IP (dla maszyn z systemem Windows zostanie wtedy przydzielony domyślny adres IP). Włączenie PortFast może umożliwić uniknięcie tego problemu, ale musisz być bardzo ostrożny, stosując tę funkcję. Jeśli do portu z aktywnym PortFast zostanie podłączony switch, może powstać pętla, która nie zostanie wykryta.

Aby uruchomić PortFast na przełączniku z IOS, użyj polecenia interfejsu `spanning-tree portfast`. Zostanie przy tym wyświetlone ostrzeżenie o zagrożeniach wynikających z włączenia tej funkcji:

```
Cat-3550(config-if)#spanning-tree portfast
%Warning: portfast should only be enabled on ports connected to a single host.
➥Connecting hubs, concentrators, switches, bridges, etc... to this interface when
➥portfast is enabled, can cause temporary bridging loops. Use with CAUTION
// Uwaga: portfast należy uruchamiać jedynie na portach
// podłączonych do pojedynczych hostów. Podłączenie hubów, koncentratorów, przełączników czy mostków (oraz innych
// urządzeń tego typu) może doprowadzić do powstania tymczasowych pętli. Stosuj ostrożnie.
%Portfast has been configured on FastEthernet0/20 but will only have effect when the interface is in
➥a non-trunking mode. // Portfast został skonfigurowany na interfejsie FastEthernet0/20,
      // ale będzie działał jedynie wtedy, gdy interfejs jest w trybie nietrunkingowym.
```

Aby wyłączyć PortFast na interfejsie w IOS, zastosuj negację pierwotnej komendy. Nie wyświetlą się już żadne dodatkowe informacje:

```
Cat-3550(config-if)#no spanning-tree portfast
Cat-3550(config-if)#
```

W przypadku CatOS polecenie włączające funkcję PortFast to `set spantree portfast <mod/port> enable`. I tutaj napotkasz podobne ostrzeżenie:

```
CatOS-6509: (enable)set spantree portfast 3/10 enable
Warning: Connecting Layer 2 devices to a fast start port can cause temporary spanning tree loops. Use
with caution. // Uwaga: podłączenie urządzeń warstwy drugiej do portu z włączoną tą
                                  // funkcją może spowodować powstanie tymczasowych pętli. Stosuj ostrożnie.
Spantree port  3/10 fast start enabled.
```

Aby wyłączyć PortFast, użyj tej samej komendy, zmieniając jedynie `enable` na `disable`:

```
CatOS-6509: (enable)set spantree portfast 3/10 disable
Spantree port  3/10 fast start disabled.
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W przełącznikach Nexus stosuje się komendę interfejsu `spanning-tree port type edge` do uruchomienia PortFast:

```
NX-5K-1(config-if)# spanning-tree port type edge
Warning: Edge port type (portfast) should only be enabled on ports connected to
➥a single host. Connecting hubs, concentrators, switches, bridges, etc... to this interface when
➥edge port type (portfast) is enabled, can cause temporary bridging loops. Use with CAUTION
➥Edge Port Type (Portfast) has been configured on Ethernet100/1/12 but will only have effect when the
➥interface is in a non-trunking mode.
```

Dla serwerów wymagających trunków przeznaczony jest specjalny typ funkcji PortFast. Aby skonfigurować dla serwera trunk z funkcją PortFast, wprowadź polecenie interfejsu `spanning-➥tree port type edge trunk`:

```
NX-5K-1(config-if)# spanning-tree port type edge trunk
Warning: Edge port type (portfast) should only be enabled on ports connected to a single host.
➥Connecting hubs, concentrators, switches, bridges, etc... to this interface when edge port type
➥(portfast) is enabled, can cause temporary bridging loops. Use with CAUTION
```

BPDU Guard

Porty z włączoną funkcją PortFast nie będą otrzymywać ramek BPDU, o ile podłączone są do urządzeń innych niż przełączniki/mostki. Jeśli jednak taki port zostanie podłączony do przełącznika, wystąpi pętla między mostkami. Aby temu zapobiec, Cisco wprowadziło funkcję **BPDU Guard**. BPDU Guard automatycznie wyłącza port ze skonfigurowaną funkcją PortFast, w przypadku gdy zacznie on otrzymywać ramki BPDU. Port nie jest ustawiany w stan zablokowany, ale w stan **ErrDisable**. Musi przy tym zostać zresetowany. W przypadku IOS BPDU Guard jest uruchamiany poleceniem `spanning-tree bpduguard enable`:

```
Cat-3550(config-if)#spanning-tree bpduguard enable
```

Aby wyłączyć funkcję, zmieniamy w komendzie `enable` na `disable`.

W CatOS użyj polecenia `set spantree bpdu-guard <mod/port> enable` (lub `disable`):

```
CatOS-6509: (enable)set spantree bpdu-guard 3/10 enable
Spantree port  3/10 bpdu guard enabled.
```

Karty rozszerzeń typu FEX (ang. *fabric extender*) w systemach Nexus mają domyślnie włączoną funkcję BPDU Guard i nie powinna ona być wyłączana. Zapobiega to podłączeniu przełącznika do kart FEX, które zaprojektowane zostały tylko do podłączania hostów.

UplinkFast

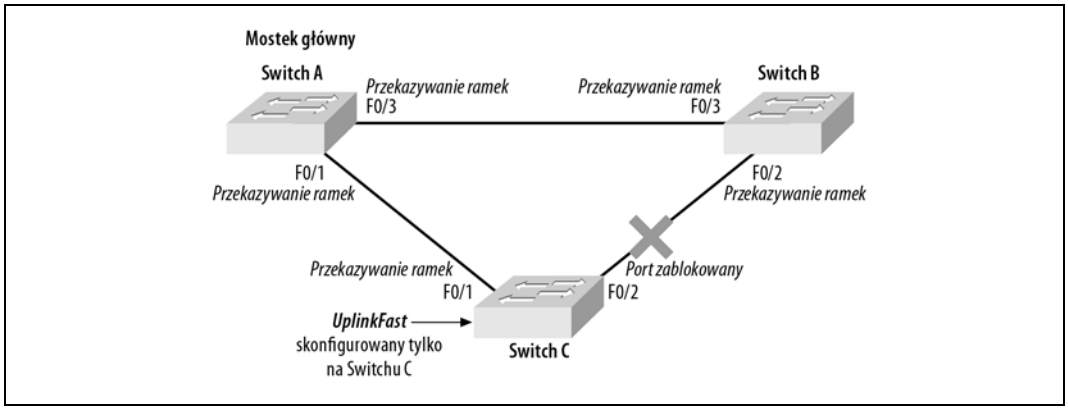
UplinkFast to funkcja przeznaczona dla przełączników pracujących w warstwie dostępowej. Przełączniki te są z reguły zlinkowane z innymi przełącznikami, aby mieć połączenie z warstwą dystrybucyjną. Standardowo w przypadku awarii łącza na porcie wyznaczonym jego funkcję musi przejąć port z alternatywną ścieżką do mostka głównego. Port ten musi najpierw przejść przez stany „port aktywny" i „uczenie się adresów MAC", zanim powróci do stanu „przekazywanie ramek". Tylko wtedy może normalnie przekazywać pakiety. Cały proces może trwać 45 sekund lub dłużej.

UplinkFast pozwala blokowanemu portowi uplinkowemu pominąć stany „port aktywny" i „uczenie się adresów MAC" w przypadku awarii portu wyznaczonego. Taka procedura umożliwia szybkie przywrócenie działania sieci, w czasie krótszym nawet niż 5 sekund. Funkcja

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ma wpływ na wszystkie VLAN-y na przełączniku. Ustawia też priorytet mostka na 49152, co zapewni, że switch nie zostanie mostkiem głównym.

Rysunek 8.6 przedstawia sposób zastosowania UplinkFast w prostej sieci. Switche A i B mogą być zarówno w warstwie rdzenia, jak i w warstwie dystrybucyjnej. Switch C jest przełącznikiem warstwy dostępowej. Jedyne łącza Switcha C z innymi przełącznikami prowadzą do warstwy dystrybucyjnej lub rdzenia.



Rysunek 8.6. Przykład zastosowania UplinkFast

Funkcja UplinkFast powinna być uruchamiana jedynie na przełącznikach warstwy dostępowej. Nigdy nie należy jej włączać na przełącznikach warstwy dystrybucyjnej lub rdzenia, ponieważ nie pozwala ona, by dany przełącznik został mostkiem głównym. Nie jest to wskazane dla przełączników warstwy rdzenia.

Aby włączyć UplinkFast na przełączniku z CatOS, zastosuj komendę `set spantree uplinkfast enable`:

```
CatOS-6509: (enable)set spantree uplinkfast enable
VLANs 1-4094 bridge priority set to 49152.
The port cost and portvlancost of all ports set to above 3000.
Station update rate set to 15 packets/100ms.
uplinkfast all-protocols field set to off.
uplinkfast enabled for bridge.
```

Bądź ostrożny, wyłączając funkcję UplinkFast, i pamiętaj, że wiele parametrów zostało zmodyfikowanych, kiedy ją uruchomiłeś:

```
CatOS-6509: (enable)set spantree uplinkfast disable
uplinkfast disabled for bridge.
Use clear spantree uplinkfast to return stp parameters to default.
                    // Użyj polecenia clear spantree uplinkfast, aby przywrócić domyślne ustawienia stp
```

Ostatnia linijka listingu jest dość istotna. Jeśli usuniesz przełącznik z warstwy dostępowej i będziesz chciał ustawić go jako mostek główny, będziesz musiał przywrócić wszystkie priorytety do wartości domyślnych.

Oto, co się stanie, jeśli posłuchasz rady otrzymanej od przełącznika:

```
CatOS-6509: (enable)clear spantree uplinkfast
This command will cause all portcosts, portvlancosts, and the bridge priority on all vlans to be set
to default. // To polecenie spowoduje przywrócenie wartości domyślnych wszystkich protokołów, kosztów portvlan oraz
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
                    // priorytetów mostka na wszystkich vlan-ach.
Do you want to continue (y/n) [n]? y
VLANs 1-4094 bridge priority set to 32768.
The port cost of all bridge ports set to default value.
The portvlancost of all bridge ports set to default value.
uplinkfast all-protocols field set to off.
uplinkfast disabled for bridge.
```

Wartości parametrów nie zostały przywrócone do tych sprzed włączenia funkcji UplinkFast — zostały przywrócone do ustawień domyślnych przełącznika.

Podczas konfiguracji UplinkFast w przełącznikach z IOS nie pojawiają się niepokojące komunikaty, tak jak miało to miejsce w przypadku CatOS:

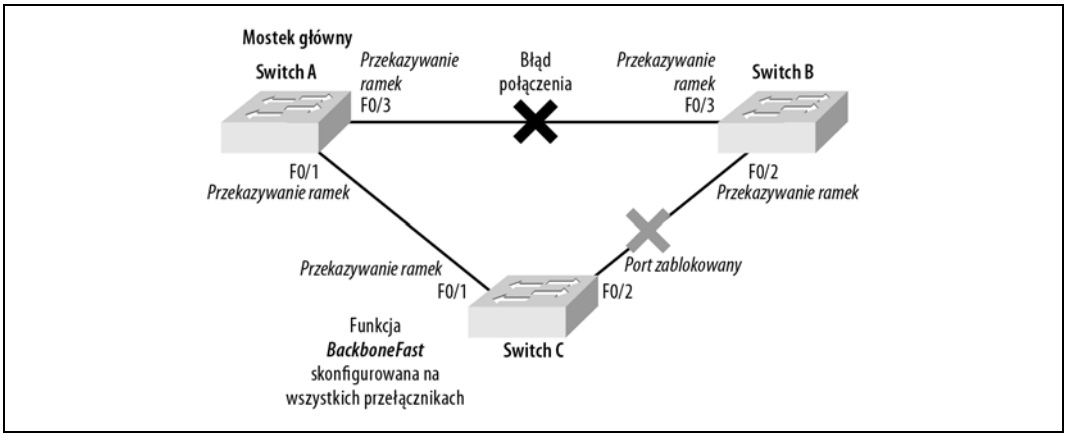
```
Cat-3550(config)#spanning-tree uplinkfast
Cat-3550(config)#
```

I tyle. Proste! Ale niech ta prostota Cię nie zwiedzie — włączenie UplinkFast zmienia priorytety również w IOS. Odmiennie jednak, niż ma to miejsce w CatOS, wyłączenie tej funkcji (komendą `no spannig-tree uplinkfast`) automatycznie przywraca domyślne wartości priorytetów. To nie zawsze może być tym, czego oczekujesz. Bądź więc ostrożny.

BackboneFast

Kiedy przełącznik otrzymuje komunikat BPDU próbujący narzucić mniej optymalny mostek główny niż ten obecnie wykorzystywany, odrzuca taką ramkę BPDU. Dzieje się tak, dopóki przełącznik nie wykryje lepszego mostka głównego. Jeśli ramki BPDU od bardziej optymalnego mostka przestaną napływać, przełącznik będzie uważał go za najlepszy mostek główny do czasu przekroczenia domyślnej wartości parametru `max_age` (maksymalny czas trwania ramki). Wartość domyślna `max_age` to 20 sekund.

Rysunek 8.7 przedstawia sieć z trzema przełącznikami. Wszystkie switche pracują w warstwie rdzenia lub w warstwie dystrybucyjnej, choć mogłyby być również przełącznikami warstwy dostępowej. W normalnej konwergencji port F0/2 na Switchu C jest portem blokowanym, podczas gdy pozostałe porty przekazują ramki.



Rysunek 8.7. Przykład BackboneFast

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Powiedzmy, że przytrafia się awaria, która wyłącza link F0/3 pomiędzy Switchem A i B. Link ten nie łączy się bezpośrednio ze Switchem C. W rezultacie mamy do czynienia z **awarią pośredniego linku** (ang. *indirect link failure*) na Switchu C. Kiedy Switch B rozpozna awarię linku, wykrywa, że nie ma ścieżki do mostka głównego, więc zaczyna rozgłaszać, iż sam jest mostkiem głównym. Przedtem Switch B wysyłał ramki BPDU z informacją, że najbardziej optymalnym mostkiem głównym jest Switch A. Taką też informację nadal pamięta Switch C i nie zmieni on konfiguracji, dopóki nie upłynie domyślny czas określony dla parametru `max_age`.

Po 20 sekundach Switch C zaakceptuje ramki BPDU z komunikatem od Switcha B i zacznie wysyłać do niego własne komunikaty BPDU. Kiedy Switch B zacznie otrzymywać ramki BPDU od Switcha C, wykryje, że istnieje ścieżka do Switcha A (który jest najbardziej optymalnym mostkiem głównym) poprzez Switch C, i ponownie zaakceptuje Switch A jako mostek główny. Przy domyślnych ustawieniach liczników Spanning Tree proces ten zajmie co najmniej 50 sekund.

BackboneFast dodaje funkcjonalność umożliwiającą wykrywanie awarii pośrednich linków. Po awarii łącza stara się aktywnie odnaleźć ścieżkę do mostka głównego, wysyłając pakiety **root link query**. Kiedy odnajdzie ścieżkę, ustawia licznik `max_age` na 0, aby port mógł przejść kolejne stany Spanning Tree bez oczekiwania przez kolejne 20 sekund.

Jeśli chcesz zastosować funkcję BackboneFast, musisz ją włączyć na wszystkich przełącznikach w sieci.

Włączenie BackboneFast na przełączniku z IOS jest łatwe. Wystarczy ogólna komenda `spanning-tree backbonefast`:

```
Cat-3550(config)#spanning-tree backbonefast
Cat-3550(config)#
```

Negacja komendy wyłącza tę funkcję.

W przypadku przełączników z CatOS skorzystaj z polecenia `set spantree backbonefast enable` do uruchomienia funkcji:

```
CatOS-6509: (enable)set spantree backbonefast enable
Backbonefast enabled for all VLANs
```

Aby ją wyłączyć, zamień w poleceniu parametr `enable` na `disable`.

Najczęstsze problemy ze Spanning Tree

Spanning Tree może być czasem nie lada wyzwaniem dla administratora, kiedy nie działa w oczekiwany sposób. Co więcej, zdiagnozowanie związanych z tym problemów może być bardzo trudne, jeśli sieć nie jest prawidłowo zaprojektowana. Oto kilka typowych problemów i sposoby ich unikania.

Niezgodność dupleksu

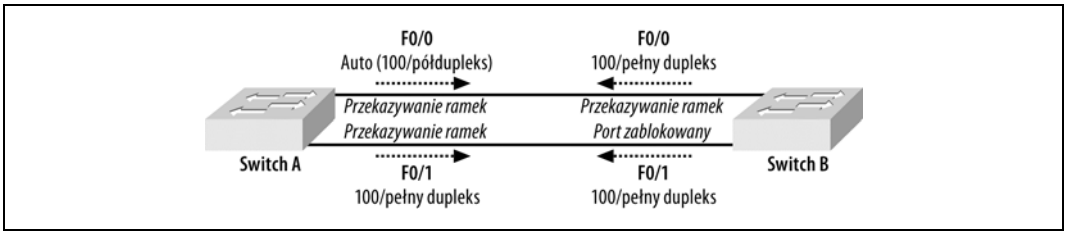
Mostek otrzymuje i przetwarza ramki BPDU na portach, nawet jeśli są one zablokowane. W przypadku awarii podstawowej ścieżki pozwala to mostkowi określić, czy istnieje inna ścieżka do mostka głównego.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli port w stanie zablokowanym przestałby otrzymywać ramki BPDU, mostek nie brałby dłużej pod uwagę tego portu jako możliwej ścieżki do mostka głównego. W takim przypadku nie

ma już konieczności blokowania tego portu, więc może on zostać przesunięty w stan przekazywania ramek. Czy taka sytuacja jest faktycznie możliwa? Mnie przydarzyło się to więcej niż raz.

Typowy problem ze Spanning Tree przedstawiony jest na rysunku 8.8. Są tu dwa przełączniki połączone dwoma linkami: F0/0 na Switchu A jest połączony z F0/0 na Switchu B, a F0/1 na Switchu A jest połączony z F0/1 na Switchu B. Switch A jest mostkiem głównym. Wszystkie porty ustawione są w stan przekazywania ramek, z wyjątkiem F0/1 na Switchu B, który jest portem zablokowanym. Sieć działa stabilnie, ponieważ Spanning Tree zerwał potencjalne pętle. Strzałki pokazują wysyłanie ramek BPDU.



Rysunek 8.8. Problem półdupleksu w Spanning Tree

Port F0/0 na Switchu A to jedyny port ustawiony w trybie autonegocjacji. Autonegocjacja określiła dla portu parametry 100 Mb/s w półdupleksie. Pozostałe porty zostały ręcznie skonfigurowane do 100 Mb/s w pełnym dupleksie. Spanning Tree wysyła ramki BPDU ze wszystkich portów i na wszystkich portach je odbiera — uczestniczy w tym również port zablokowany.

Zawsze upewnij się, że obie strony łącza Ethernet są skonfigurowane w ten sam sposób pod względem prędkości i trybu dupleksu. Szczegóły znajdziesz w rozdziale 3.

Kiedy port jest w półdupleksie, przed rozpoczęciem transmisji nasłuchuje ewentualnych kolizji. Port w pełnym dupleksie tego nie robi. Port w pełnym dupleksie połączony z portem w półdupleksie będzie wysyłał ramki nieustannie, co spowoduje, że ten drugi port zarejestruje wiele kolizji. Po wykryciu kolizji port uruchomi algorytm „back-off" i poczeka z ponownym przesłaniem kolidującego pakietu. W naszym przykładzie port w półdupleksie jest aktywnym linkiem, przez który transmitowane są dane. Gdy natężenie ruchu jest wysokie, problem kolizji robi się poważny, co skutkuje zrzucaniem ramek — również ramek BPDU.

Switch B będzie nasłuchiwał ramek na obu linkach pokazanych na rysunku 8.8. Jeśli przez określony czas nie otrzyma żadnej ramki BPDU na interfejsie F0/0, to nie będzie dłużej traktował tego łącza jako właściwej ścieżki do mostka głównego. Ponieważ była to podstawowa ścieżka do mostka głównego, przełącznik nie będzie w tym momencie widział go w sieci i zmieni interfejs F0/1 ze stanu zablokowanego do stanu przekazywania ramek, aby odzyskać ścieżkę do mostka głównego. W tym momencie nie ma już zablokowanych portów na żadnym z dwóch linków łączących oba przełączniki i pojawia się pętla między mostkami.

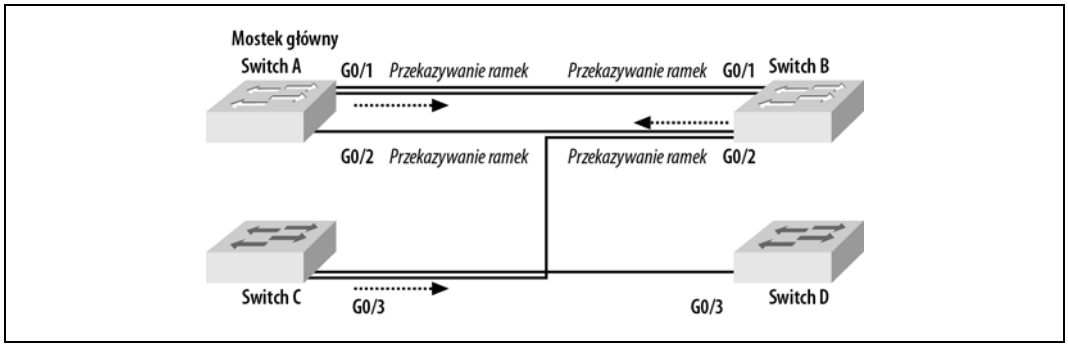
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Linki jednokierunkowe

Jeśli link może transmitować tylko w jednym kierunku, określany jest jako link **jednokierunkowy** (ang. *unidirectional*). Może się to przydarzyć w przypadku miedzianego Ethernetu, jednak najczęściej problem ten występuje w sieciach światłowodowych.

Podczas instalowania światłowodów istotną rzeczą jest przekrosowanie pojedynczych włókien pary światłowodu na przeciwległych stronach łącza. Jeżeli włókna w parze są rozdzielone, może się to skończyć tym, że każde z włókien zostanie podłączone do innego portu lub nawet przełącznika.

Rysunek 8.9 przedstawia cztery przełączniki. Switch A powinien być podłączony do Switcha B dwoma parami światłowodów — jeden pomiędzy interfejsami G0/1, a drugi pomiędzy interfejsami G0/2 na obu przełącznikach. Gdzieś w strukturze okablowania para włókien światłowodu dla linku G0/2 została rozdzielona. Pomimo prawidłowego podłączenia końcówek na Switchu B do Switcha A poprowadzono tylko jedno włókno z pary. Drugie włókno zostało skierowane do Switcha C i podłączone do interfejsu G0/3.



Rysunek 8.9. Problem linku jednokierunkowego

Interfejsy światłowodowe testują integralność łącza przez ustalenie, czy jest sygnał po stronie RX (odbiorczej) danej pary. Port G0/2 na Switchu B wykrywa integralność łącza, ponieważ RX jest aktywne od strony Switcha C. Port G0/2 na Switchu A wykrywa integralność łącza, ponieważ RX jest aktywne od strony Switcha B. I w końcu interfejs G0/3 na Switchu C wykrywa integralność łącza, gdyż RX jest aktywne od strony Switcha D.

Spanning Tree wysyła ramki BPDU z każdego interfejsu na Switchach A i B, ponieważ łącza są aktywne. Switch A jest mostkiem głównym. Switch B odbiera ramki od mostka głównego tylko na jednym porcie: G0/1. Ponieważ Switch B nie otrzymuje ramek BPDU na porcie G0/2, Spanning Tree nie blokuje tego portu. Broadcasty otrzymane przez Switch B na porcie G0/1 będą przekazywane dalej przez port G0/2. W ten sposób powstaje pętla.

Problem ten może być trudny do wykrycia. Pętla między dwoma mostkami powoduje chaos w sieci, ponieważ wykorzystanie procesora w urządzeniach może szybko osiągnąć poziom 100%, co spowoduje przestoje. Pierwszą czynnością, jaką w takich przypadkach wykonują niedoświadczeni inżynierowie, jest zrestartowanie jednego lub kilku urządzeń sieciowych. Jeśli masz do czynienia z linkiem jednokierunkowym, nie przyniesie to pozytywnego rezultatu. Gdy restart nie pomaga, zazwyczaj pętla zostaje wykryta, a inżynier wyłącza jeden z linków. Niestety powoduje to najczęściej utratę możliwości zdiagnozowania linku jednokierunkowego.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Warstwa fizyczna w pierwszej kolejności! Diagnozując problemy z łącznością, zawsze bierz pod uwagę, że problem może być natury fizycznej. Może to zaoszczędzić Ci wielu godzin bólu głowy, zwłaszcza kiedy wszystkie pozostałe oznaki nie wydają się sugerować niczego sensownego. Nie zakładaj również, że coś będzie działać dzisiaj, tylko dlatego, że działało wczoraj. Nie potrzeba wiele, aby przypadkiem zgnieść włókno światłowodu, zamykając drzwi szafy.

W najnowszych wersjach systemów IOS i CatOS problemy z jednokierunkowym linkiem zostały rozwiązane przez zastosowanie protokołu UDLD (ang. *Unidirectional Link Detection*). UDLD jest domyślnie włączony i tak powinieneś go pozostawić. W przypadku wykrycia błędów UDLD szukaj problemów podobnych do tych, które właśnie opisałem.

Bridge Assurance

W przełącznikach Nexus wprowadzono pewną zabawną poprawkę, zwaną **Bridge Assurance** (BA), czyli asekuracją mostka. Została ona zaprojektowana w celu zapobiegania powstawaniu pętli w sieci. Chwileczkę, czyż do tego nie służy właśnie Spanning Tree? Tak, jednak Bridge Assurance sprawi, że sieć będzie idiotoodporna! Jeśli w to uwierzysz, mogę Ci sprzedać mostek wraz z książką o tym, czemu frajerzy nie powinni używać kalamburów.

Asekuracja mostka ma chronić przed jednokierunkowymi linkami, działa jednak inaczej niż UDLD.

Zacznijmy od marketingowej notatki zamieszczonej w dokumentacji przełączników Cisco Nexus:

> Bridge Assurance jest włączone domyślnie i może być wyłączone tylko wtedy, jeśli zostanie to zrobione dla całej sieci. Ponadto Bridge Assurance może być włączone tylko na portach, na których działa Spanning Tree i które są linkami typu punkt-punkt. I wreszcie, oba końce łącza muszą mieć włączone Bridge Assurance.
>
> Przy włączonym Bridge Assurance ramki BPDU są wysyłane przez wszystkie działające w sieci porty, również porty alternatywne i zapasowe, co ustalony interwał czasu zwany „hello time". Jeśli port nie otrzyma ramki BPDU w określonym interwale czasowym, zostaje przeniesiony w stan zablokowania i nie jest brany pod uwagę podczas wyboru portu głównego. Jeśli po tej przerwie port zacznie znów otrzymywać ramki BPDU, zostanie przywrócony do normalnego obiegu Spanning Tree.

Oto, co powinieneś zapamiętać: BA jest zawsze włączone, chyba że zostało wyłączone globalnie, a przy włączonym BA ramki BPDU są wysyłane przez wszystkie działające porty sieciowe. Co więcej, jeśli nie zarządzasz BA właściwie, to przerwie ono działanie Twojej sieci. Choć jest to prawda również w przypadku Spannig Tree, to przyzwyczailiśmy się już do sposobu, w jaki działa. BA zmienia nieco reguły gry.

BA obsługiwane jest tylko przez Rapid PVST+ i MST. Jest to ważne, ponieważ aby BA działało poprawnie, musi być włączone i skonfigurowane po obu stronach linku. Jeśli podłączysz przełącznik Nexus 5K do przełącznika 2960, na którym uruchomione jest PVST, BA może wyłączyć link do 2960.

W przypadku przełączników Nexus porty Spanning Tree identyfikowane są jako **normal** (normalne), **host** (porty hosta) lub **network** (porty sieciowe). Port hosta jest analogiczny do `spanning-tree portfast` w IOS. Port sieciowy nie jest podłączony do żadnego hosta, ale może być podłączony do innego przełącznika z funkcją BA. Bądź ostrożny, jeśli kusi Cię, aby wszystkie porty trunkowe uczynić portami sieciowymi Spanning Tree. Miałem klienta, dla którego podłączyliśmy kiedyś przełącznik Nexus 5K do 6509. 6509 obsługiwał Bridge Assurance,

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ale uważaliśmy, odpowiednio konfigurując łącze między przełącznikami. Po stronie Nexusa skonfigurowaliśmy dla Spanning Tree port typu **network** (ostatecznie był on portem typu network), ale nie spodziewaliśmy się, że 6509 miał uruchomione PVST. W rezultacie Nexus zablokował większość naszych VLAN-ów. Zmiana portu trunkowego Spanning Tree na typ **normal** rozwiązała problem.

Zasady projektowania umożliwiające uniknięcie problemów ze Spanning Tree

Właściwy projekt może zminimalizować powstawanie problemów związanych z działaniem protokołu Spanning Tree. Najprostszym sposobem ograniczenia komplikacji jest prowadzenie dokumentacji i dobra znajomość własnej sieci. Jeśli zaczniesz się zastanawiać, jak działa Twoja sieć, dopiero gdy pojawi się problem, to problem ten może mieć dłuższy żywot niż Twoje zatrudnienie.

W systemach redundancji stosuj routing zamiast przełączania

Niegdyś zwykło się powiadać: „Przełączaj, kiedy możesz, i rutuj, kiedy musisz". W dzisiejszym świecie szybkich przełączników warstwy trzeciej mantra ta nie ma dłużej zastosowania. W przełącznikach warstwy trzeciej możesz rutować z prędkościami przełączania.

Dla wielu osób zrozumienie redundancji warstwy trzeciej jest łatwiejsze niż zrozumienie redundancji warstwy drugiej. Dopóki zaspokajane są potrzeby biznesowe, a rezultat jest taki sam, wykorzystanie routingu do rozwiązywania problemów redundancji jest całkowicie akceptowalne.

Jeśli zdecydujesz się użyć routingu zamiast przełączania, nie wyłączaj Spanning Tree. Spanning Tree będzie nadal chronić przed pętlami, które mogłeś przeoczyć. Gdy stosujesz przełączniki — nawet przełączniki warstwy trzeciej — Spanning Tree może ocalić Ci życie, jeśli ktoś podepnie przełącznik tam, gdzie nie powinno go być.

Zawsze konfiguruj mostek główny

Nie pozwalaj, by Spanning Tree dynamicznie wybierał mostek główny. Sam zdecyduj, który przełącznik w sieci powinien nim być, i ustaw dla niego priorytet mostka 1. Jeśli pozwolisz, aby przełączniki decydowały, to nie dość, że często ten wybór nie będzie miał sensu, ale i dołączane potem kolejne przełączniki będą przyjmowały, że są mostkiem głównym. Doprowadzi to do przekalkulowania zasad działania całej sieci i zmiany stanów portów, gdy sieć określi ścieżki do nowego mostka głównego.

Widziałem kiedyś zamieszanie w dużej sieci firmy finansowej spowodowane tym, że mostkiem głównym nagle okazał się przełącznik pod czyimś biurkiem. Mostek główny w tej sieci nie został uprzednio ręcznie skonfigurowany, więc kiedy ktoś podłączył jakiś przypadkowy switch, pozostałe przełączniki skierowały się do niego, ponieważ miał on najniższy adres MAC wśród wszystkich switchy. Choć na pierwszy rzut oka może się to nie wydawać dużym problemem, to w konsekwencji główny trunk pomiędzy przełącznikami warstwy rdzenia został zablokowany. Okazał się on bowiem najdalszym linkiem od mostka głównego.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 9.

Routing i routery

Termin **routing**, inaczej trasowanie, może posiadać wiele znaczeń w zależności od dyscypliny, w której jest stosowany. Zasadniczo odnosi się on do sposobu wyznaczania ścieżki lub trasy wykorzystywanej dla określonych celów. W telekomunikacji rozmowa może być rutowana na podstawie wybranego numeru lub jakiegoś innego identyfikatora. W każdym z przypadków chodzi o określenie ścieżki dla tej rozmowy.

Poczta jest również rutowana — i nie mówię tu o poczcie elektronicznej (chociaż e-mail oczywiście także jest rutowany) — ale raczej o tradycyjnej poczcie. Kiedy piszesz na kopercie adres i kod pocztowy, dostarczasz parametry umożliwiające poczcie rutowanie Twojego listu. Podajesz adres docelowy i zazwyczaj adres nadawcy, a poczta wyznacza najlepszą trasę dla przesyłki. Jeśli istnieje problem z dostarczeniem listu, adres zwrotny wykorzystywany jest w celu rutowania go z powrotem do Ciebie. Sama ścieżka, którą list jest dostarczany od źródła do celu, nie ma praktycznie znaczenia. Dla Ciebie liczy się jedynie (i całe szczęście) dostarczenie przesyłki w odpowiednim czasie i w jednym kawałku.

W środowisku IP pakiety lub ramki w sieci lokalnej przekazywane są przez przełączniki, huby lub mostki. Jeśli adres docelowy znajduje się poza obrębem sieci lokalnej, pakiet musi być przekazany do **bramy** (ang. *gateway*). Brama odpowiada za określenie sposobu dostarczenia pakietu tam, gdzie powinien się znaleźć. Publikacja RFC 791, zatytułowana PROTOKÓŁ INTERNETOWY (ang. *Internet Protocol*), definiuje bramę następująco:

> 2.4. Bramy
>
> Bramy wykorzystują protokół internetowy do przekazywania datagramów między sieciami. Stosują też protokół GGP (ang. *Gateway-to-Gateway Protocol*) do koordynowania routingu i innych internetowych informacji kontrolnych.
>
> W bramie nie muszą być implementowane protokoły wyższej warstwy, a funkcje GGP są dodawane do modułu IP.

Kiedy stacja w sieci wysyła pakiet do bramy, nie przejmuje się tym, **w jaki sposób** zostanie on dostarczony do celu — ważne jest tylko, żeby został dostarczony (przynajmniej w przypadku protokołu TCP). Podobnie jak w przypadku tradycyjnej przesyłki pocztowej każdy pakiet zawiera informacje o adresie źródłowym i docelowym, więc routing nie powinien stanowić problemu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W sferze semantyki oraz IP brama jest urządzeniem, które przekazuje pakiety do miejsca przeznaczenia innego niż sieć lokalna. Z wielu względów praktycznych brama jest routerem. Generalnie używam w tej książce terminu router, choć napotkasz jeszcze określenie **brama domyślna** (ang. *default gateway*).

W dawnych czasach w dziedzinie transmisji danych „bramą" nazywano urządzenie tłumaczące jeden protokół na drugi. Przykładowo urządzenie konwertujące łącze szeregowe na łącze równoległe mogło być określane jako brama. Podobnie było w przypadku urządzenia konwertującego Ethernet na Token Ring. Obecnie takie urządzenia nazywa się **media-konwerterami** (ang. *media converters*). (To nie będzie pierwszy raz, kiedy można mnie oskarżyć o nawiązywanie do starych dobrych czasów — usiądź wygodnie w bujanym fotelu na werandzie, popijając miętową lemoniadę, a ja opowiem Ci historyjkę).

Routery zazwyczaj komunikują się ze sobą za pomocą jednego lub kilku **protokołów routingu** (ang. *routing protocols*), zwanych też protokołami trasowania. Protokoły te umożliwiają routerom pozyskiwanie informacji o sieciach innych niż ta, do której są podłączone.

Urządzenia sieciowe ograniczały się niegdyś do mostków i routerów. Mostki, huby i przełączniki operowały jedynie w warstwie drugiej modelu ISO/OSI, a routery tylko w warstwie trzeciej. Obecnie te poszczególne urządzenia często połączone są w jedno, a routery i przełączniki mogą pracować we wszystkich siedmiu warstwach modelu ISO/OSI.

W dzisiejszym świecie każde urządzenie ma, jak się wydaje, nieograniczone możliwości. Czemu więc miałbyś wybrać router zamiast przełącznika? Routery są raczej WAN-centryczne, podczas gdy przełączniki LAN-centryczne. Jeśli więc łączysz sieci T1, wybierzesz prawdopodobnie router. Do sieci Ethernet lepiej nada się switch.

Tabele routingu

Routing jest podstawowym procesem, wspólnym dla prawie każdego rodzaju stosowanych obecnie sieci. Mimo to nadal wielu inżynierów nie rozumie, jak działa routing. O ile proces uzyskiwania certyfikatu Cisco powinien Ci pomóc opanować konfigurację routera, o tyle w tej sekcji pokażę Ci, co powinieneś wiedzieć o routingu w prawdziwym świecie. Skupię się na fundamentach, bo tego właśnie wydaje się brakować wielu inżynierom — spędzamy mnóstwo czasu na studiowaniu najnowszych technologii, zapominając czasem o podstawowych zasadach, na których wszystko się opiera.

W routerach Cisco tabela routingu (inaczej tablica trasowania) zwana jest **bazą informacji routingu** — RIB (ang. *route information base*). Jeśli wpiszesz polecenie `show ip route`, otrzymasz sformatowany widok informacji z RIB.

Każdy protokół trasowania ma swoją własną tablicę informacji. Przykładowo EIGRP posiada tablicę topologii, a OSPF (ang. *Open Shortest Path First*) bazę danych OSPF. Każdy protokół sam decyduje też, które trasy będzie przechowywał w swojej bazie danych. Protokoły routingu używają indywidualnych metryk do określania najlepszych tras. Mogą to być bardzo różne metryki, a wartości metryczne zależą od protokołu, który posłużył do wyznaczenia trasy. W ten sposób ta sama ścieżka może mieć wartość metryczną 2 w przypadku protokołu RIP (ang. *Routing Information Protocol*), 200 w przypadku OSPF oraz 156160 dla EIGRP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Protokoły routingu oraz metryki opisane są bardziej szczegółowo w rozdziale 10.

Jeśli ta sama ścieżka wyznaczona została na podstawie informacji z dwóch źródeł w obrębie jednego protokołu routingu, to stosowana jest ta z lepszą metryką. Z kolei jeśli ta sama ścieżka wyznaczona została przez dwa protokoły na tym samym routerze, to zostanie wybrany protokół charakteryzujący się mniejszym dystansem administracyjnym. **Dystans administracyjny** (ang. *administrative distance*) to wartość przypisana każdemu protokołowi routingu, umożliwiająca routerowi priorytetyzację ścieżek wytyczonych przez kilka źródeł. Tabela 9.1 przedstawia zestawienie wartości dystansów administracyjnych dla różnych protokołów routingu.

Tabela 9.1. Dystans administracyjny poszczególnych protokołów routingu

Rodzaj trasy	Dystans administracyjny
Interfejs podłączony bezpośrednio	0
Trasa statyczna	1
Protokół EIGRP trasa skonsolidowana	5
Protokół External BGP	20
Protokół Internal EIGRP	90
Protokół IGRP	100
Protokół OSPF	110
Protokół IS-IS	115
Protokół RIP	120
Protokół EGP	140
Protokół ODR	160
Protokół External EIGRP	170
Protokół Internal BGP	200
Nieznany	255

Spanning Tree, omówiony w rozdziale 8., nie jest właściwie protokołem trasowania, gdyż nie dotyczy bezpośrednio transmisji danych. STP odpowiada za pętle i zapobiega ich powstawaniu z perspektywy fizycznej oraz warstwy drugiej modelu ISO/OSI. Innymi słowy, Spanning Tree zajmuje się raczej wybieraniem takich ścieżek w obrębie swojej domeny, które będą wolne od pętli, niż takich, po których mają być przesyłane dane.

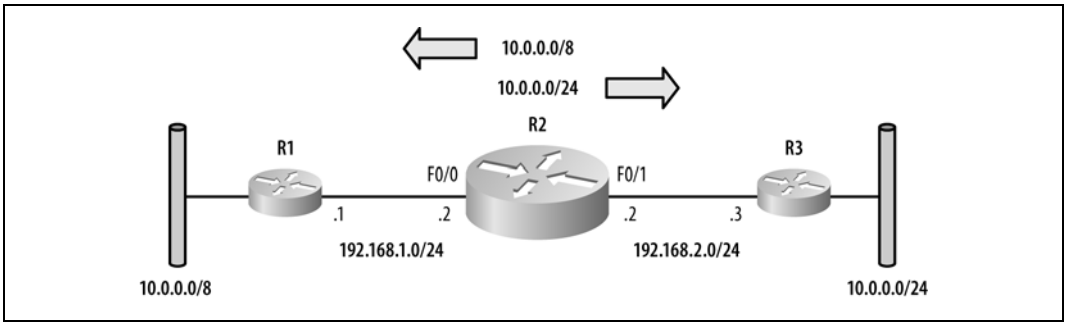
Router, odbierając pakiet, decyduje, czy ma on być przesłany do innej sieci. Jeśli tak, router sprawdza, czy w bazie RIB istnieje zapisana ścieżka do docelowej sieci. W przypadku odnalezienia zgodności pakiet jest dostosowywany i przesyłany właściwym interfejsem do miejsca przeznaczenia (więcej szczegółów na temat tego procesu znajdziesz w rozdziale 15.). Jeśli w bazie nie odnaleziono ścieżki, pakiet jest przekazywany do bramy domyślnej, o ile taka istnieje. Jeśli nie ma bramy domyślnej, to pakiet jest odrzucany.

Pierwotnie sieć docelową opisywano przez podanie adresu sieciowego oraz maski podsieci. Dzisiaj dla określenia sieci docelowej podaje się adres sieciowy i długość prefiksu. Adres sieciowy

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

to adres IP sieci, do której się odnosimy. Długość prefiksu to liczba bitów w masce podsieci, które mają wartość 1. Do określenia sieci stosuje się więc format **adres-sieciowy/długość-prefiksu**. Przykładowo sieć 10.0.0.0 z maską podsieci 255.0.0.0 może być opisana jako 10.0.0.0/8. W przypadku takiego formatu ścieżka zwana jest po prostu prefiksem. Mówi się, że sieć 10.0.0.0/24 ma dłuższy prefiks niż sieć 10.0.0.0/8. Im więcej bitów wykorzystanych jest do identyfikacji części sieciowej adresu, tym dłuższy będzie prefiks.

Baza RIB może zawierać wiele ścieżek do tej samej sieci. W przykładzie przedstawionym na rysunku 9.1 router R2 uczy się sieci 10.0.0.0 z dwóch źródeł: R1 rozgłasza trasę 10.0.0.0/8, a R3 trasę 10.0.0.0/24. Ponieważ długości prefiksów tych tras różnią się od siebie, są one traktowane jako różne trasy. W rezultacie obie zostają zapisane w tabeli routingu.



Rysunek 9.1. Ta sama sieć z różnymi długościami prefiksów

Oto trasy, które widzi R2:

```
      10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
D        10.0.0.0/8 [90/30720] via 192.168.1.1, 00:12:01, FastEthernet0/0
D        10.0.0.0/24 [90/30720] via 192.168.2.3, 00:12:01, FastEthernet0/1
```

Kiedy R2 otrzymuje pakiet, jego docelowy adres IP jest porównywany z danymi zawartymi w tabeli routingu. Jeśli docelowym adresem IP pakietu jest 10.0.0.1, to którą trasę wybierze router? Są dwie trasy w tabeli, które potencjalnie pasują: 10.0.0.0/8 oraz 10.0.0.0/24. Trasa z dłuższym prefiksem jest najbardziej pożądaną trasą. Zwana jest również **najbardziej szczegółową** trasą (ang. *most specific*). Tak więc, jeśli pakiet z adresem docelowym 10.0.0.1 dociera do R2, jest przekazywany do R3. Warto, byś uświadomił sobie odnośnie do tego przykładu, że mogą istnieć też odpowiednie adresy z zakresu 10.0.0.0/24 znajdujące się za R1, do których R2 nigdy nie będzie miał dostępu.

Technicznie rzecz biorąc, 10.0.0.0/8 jest siecią, a 10.0.0.0/24 podsiecią. Więcej informacji na ten temat znajdziesz w kolejnych sekcjach tego rozdziału.

Rodzaje tras

W tabeli routingu rozróżnia się sześć rodzajów tras:

Trasa hosta (ang. *Host route*)

Trasa hosta to trasa, która nie prowadzi do sieci, a do konkretnego hosta. Trasy hosta mają zawsze maskę podsieci 255.255.255.255 z prefiksem o długości /32.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Podsieć (ang. *Subnet*)
Podsieć jest częścią sieci nadrzędnej. Do określenia wielkości podsieci używana jest maska podsieci. 10.0.0.0/24 jest podsiecią.

Trasa skonsolidowana (ang. *Summary*)
Trasa skonsolidowana jest pojedynczą trasą odnoszącą się do grupy podsieci. 10.10.0.0/16 (255.255.0.0) może być trasą skonsolidowaną, przy założeniu, że istnieje podsieć z dłuższą maską (np. 10.10.10.0/24).

Sieć nadrzędna (ang. *Major network*)
Sieć nadrzędna to każda sieć klasowa z natywną maską. 10.0.0.0/8 (255.0.0.0) jest siecią nadrzędną.

Supersieć (ang. *Supernet*)
Supersieć jest pojedynczą trasą odnoszącą się do grupy sieci nadrzędnych. Dla przykładu 10.0.0.0/7 jest supersiecią odnoszącą się do 10.0.0.0/8 i 11.0.0.0/8.

Trasa domyślna (ang. *Default route*)
Trasa domyślna jest przedstawiana jako 0.0.0.0/0 (0.0.0.0). Ta trasa zwana jest również trasą ostatniego wyboru (ang. *the route of last resort*). Jest używana, kiedy żadna inna trasa nie odpowiada docelowemu adresowi IP pakietu.

Tabela routingu IP

Aby wyświetlić tablicę routingu IP, użyj polecenia `show ip route`:

```
R2#sho ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
                // C — podłączona, S — statyczna, I — IGRP, R — RIP, M — mobilna, B — BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
                // D — EIGRP, EX — EIGRP zewnętrzna, O — OSPF, IA — OSPF strefa wewnętrzna
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 // N1 — OSPF NSSA zewnętrzna
typ 1, N2 — OSPF NSSA zewnętrzna typ 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
                // E1 — OSPF zewnętrzna typ 1, E2 — OSPF zewnętrzna typ 2, E — EGP
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
                // I — IS-IS, su IS-IS skonsolidowana, L1 — IS-IS poziom-1, L2 IS-IS poziom-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
                // ia — IS-IS strefa wewnętrzna, * kandydatka na trasę domyślną, U — trasa statyczna na-użytkownika
       o - ODR, P - periodic downloaded static route // o — ODR, P — cyklicznie pobierana trasa statyczna
Gateway of last resort is 11.0.0.1 to network 0.0.0.0
     172.16.0.0/16 is variably subnetted, 6 subnets, 2 masks
D       172.16.200.0/23 is a summary, 00:56:18, Null0
C       172.16.200.0/24 is directly connected, Loopback2
C       172.16.201.0/24 is directly connected, Serial0/0
C       172.16.202.0/24 is directly connected, Loopback3
C       172.16.100.0/23 is directly connected, Loopback4
D       172.16.101.0/24 [90/2172416] via 11.0.0.1, 00:53:07, FastEthernet0/1
C    10.0.0.0/8 is directly connected, FastEthernet0/0
C    11.0.0.0/8 is directly connected, FastEthernet0/1
     192.168.1.0/32 is subnetted, 1 subnets
D       192.168.1.11 [90/156160] via 11.0.0.1, 00:00:03, FastEthernet0/1
S*   0.0.0.0/0 [1/0] via 11.0.0.1
D    10.0.0.0/7 is a summary, 00:54:40, Null0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Pierwszy blok informacji wyświetlany jest za każdym razem, gdy wpisujesz tę komendę. Dla zwięzłości usunę go z większości przykładów w tej książce. Fragment ten zawiera objaśnienie kodów umieszczanych w kolumnie po lewej stronie tabeli routingu.

Następna linijka wyświetla informacje o bramie domyślnej, jeśli taka istnieje:

```
Gateway of last resort is 11.0.0.1 to network 0.0.0.0 // Bramą ostatniego wyboru do sieci 0.0.0.0 jest 11.0.0.1
```

Jeśli bram jest więcej niż jedna, wszystkie zostaną wyświetlone. Jest to częsta sytuacja, kiedy brama domyślna jest pozyskana (wyuczona) z protokołu trasowania, który umożliwia jednakowy koszt podziału obciążenia. Jeśli dwa linki zapewniają dostęp do rozgłoszonej bramy domyślnej i obydwa posiadają taką samą metrykę, nie będą wymienione jako trasy domyślne. W takim przypadku pakiety zostaną równo rozdzielone pomiędzy te linki z wykorzystaniem algorytmu rozłożenia obciążenia na pakiet (ang. *per-packet load balancing*).

Jeśli żadna brama domyślna nie została skonfigurowana lub pozyskana, otrzymasz z kolei taki komunikat:

```
Gateway of last resort is not set // Brama ostatniego wyboru nie została ustawiona
```

Następny blok listingu zawiera właściwą tabelę routingu:

```
     172.16.0.0/16 is variably subnetted, 6 subnets, 2 masks
D       172.16.200.0/23 is a summary, 00:56:18, Null0
C       172.16.200.0/24 is directly connected, Loopback2
C       172.16.201.0/24 is directly connected, Serial0/0
C       172.16.202.0/24 is directly connected, Loopback3
C       172.16.100.0/23 is directly connected, Loopback4
D       172.16.101.0/24 [90/2172416] via 11.0.0.1, 00:53:07, FastEthernet0/1
C    10.0.0.0/8 is directly connected, FastEthernet0/0
C    11.0.0.0/8 is directly connected, FastEthernet0/1
     192.168.1.0/32 is subnetted, 1 subnets
D       192.168.1.11 [90/156160] via 11.0.0.1, 00:00:03, FastEthernet0/1
S*   0.0.0.0/0 [1/0] via 11.0.0.1
D    10.0.0.0/7 is a summary, 00:54:40, Null0
```

Przeanalizujmy pojedynczy wiersz tabeli routingu, abyś mógł się zorientować, co jest istotne:

```
D       172.16.101.0/24 [90/2172416] via 11.0.0.1, 00:53:07, FastEthernet0/1
```

Najpierw kod trasy. W tym przypadku jest to `D`, co oznacza, że trasa została wyuczona via EIGRP (możesz to sprawdzić w bloku kodów na listingu otrzymanym po wpisaniu polecenia `show ip route`).

Następnie sama trasa. W tym przykładzie trasa prowadzi do podsieci 172.16.101.0/24. Później są dwie wartości w nawiasie kwadratowym: pierwsza z nich to dystans administracyjny (zobacz tabela 9.1), a druga to metryka trasy. Metryka zależy od protokołu, z którego trasa została pozyskana (w tym przypadku `EIGRP`).

Kolejna informacja to następny skok (ang. *hop*), do którego router musi przesłać pakiet, aby osiągnąć podsieć docelową. W tym przypadku `via 11.0.0.1` wskazuje, że pakiet przeznaczony dla podsieci 172.16.101.0/24 powinien zostać przekazany do adresu IP 11.0.0.1. Na koniec informacja o wieku trasy (`00:53:07`). Po niej następuje oznaczenie interfejsu, przez który router wyśle dany pakiet (`FastEthernet0/1`).

Przygotowałem przykładowy router, więc tabela routingu będzie zawierała każdy z dostępnych rodzajów tras. Przypomnę, że wyróżniamy następujące rodzaje tras: **trasa hosta**, **podsieć**, **trasa skonsolidowana**, **sieć nadrzędna**, **supersieć** oraz **trasa domyślna**. Kolejne sekcje opisują je bardziej szczegółowo. Zaznaczę w tabelach wpisy dotyczące każdej z nich.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Trasa hosta

Trasa hosta to po prostu trasa z maską podsieci, w której wszystkie bity mają wartość 1 (255.255.255.255), czyli długość prefiksu wynosi /32. W tej przykładowej tablicy routingu trasa do 192.168.1.11 jest trasą hosta:

```
     172.16.0.0/16 is variably subnetted, 6 subnets, 2 masks
D       172.16.200.0/23 is a summary, 00:56:18, Null0
C       172.16.200.0/24 is directly connected, Loopback2
C       172.16.201.0/24 is directly connected, Serial0/0
C       172.16.202.0/24 is directly connected, Loopback3
C       172.16.100.0/23 is directly connected, Loopback4
D       172.16.101.0/24 [90/2172416] via 11.0.0.1, 00:53:07, FastEthernet0/1
C    10.0.0.0/8 is directly connected, FastEthernet0/0
C    11.0.0.0/8 is directly connected, FastEthernet0/1
        192.168.1.0/32 is subnetted, 1 subnets
D       192.168.1.11 [90/156160] via 11.0.0.1, 00:00:03, FastEthernet0/1
S*   0.0.0.0/0 [1/0] via 11.0.0.1
D    10.0.0.0/7 is a summary, 00:54:40, Null0
```

Zauważ, że ta trasa została pokazana jako część większej sieci (w tym przykładzie 192.168.1.0). Wiemy o tym, ponieważ informację o trasie hosta poprzedza wiersz wskazujący, jaka jest sieć nadrzędna. Router będzie próbował pokazać Ci, które klasowe (nadrzędne) sieci zawierają tę trasę. Jeśli router zna tylko jedną maskę podsieci, założy, że sieć została podzielona po równo tą maską. W danym przypadku router uznał, że sieć nadrzędna 192.160.1.0/24 została podzielona na równe podsieci, z których każda posiada maskę /32. Stąd sieć 192.168.1.0, mająca natywnie prefiks /24, została przedstawiona jako 192.168.1.0/32.

Podsieć

Podsieci są wymienione pod informacją o sieci nadrzędnej. W naszym przykładzie sieć nadrzędna 172.16.0.0/16 została podzielona na podsieci. De facto odbyło się to z wykorzystaniem rozwiązania VLSM — maski o zmiennej długości — (ang. *Variable Length Subnet Masks*), które umożliwia każdej podsieci posiadanie innej maski podsieci (z pewnymi ograniczeniami — więcej szczegółów znajdziesz w rozdziale 34.). Trasa znajdująca się na liście pomiędzy siecią nadrzędną a podsieciami nie jest zaznaczona wytłuszczoną czcionką, gdyż jest to trasa skonsolidowana — o niej powiem w następnej kolejności:

```
172.16.0.0/16 is variably subnetted, 6 subnets, 2 masks
D       172.16.200.0/23 is a summary, 00:56:18, Null0
C       172.16.200.0/24 is directly connected, Loopback2
C       172.16.201.0/24 is directly connected, Serial0/0
C       172.16.202.0/24 is directly connected, Loopback3
C       172.16.100.0/23 is directly connected, Loopback4
D       172.16.101.0/24 [90/2172416] via 11.0.0.1, 00:53:07, FastEthernet0/1
C    10.0.0.0/8 is directly connected, FastEthernet0/0
C    11.0.0.0/8 is directly connected, FastEthernet0/1
     192.168.1.0/32 is subnetted, 1 subnets
D       192.168.1.11 [90/156160] via 11.0.0.1, 00:00:03, FastEthernet0/1
S*   0.0.0.0/0 [1/0] via 11.0.0.1
D    10.0.0.0/7 is a summary, 00:54:40, Null0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Trasa skonsolidowana

Termin **skonsolidowana** (ang. *summary*) oznacza w tabelach routingu dowolną grupę tras. Z technicznego punktu widzenia według dokumentacji Cisco skonsolidowana może być grupa podsieci, podczas gdy grupa sieci nadrzędnych nazywana jest supersiecią. W tabeli routingu w obu tych przypadkach użyte jest określenie „skonsolidowana". Tak więc, mimo iż w przykładzie znaleźć można dwa takie wpisy, tylko jeden z nich faktycznie dotyczy trasy skonsolidowanej:

```
     172.16.0.0/16 is variably subnetted, 6 subnets, 2 masks
D       172.16.200.0/23 is a summary, 00:56:18, Null0
C       172.16.200.0/24 is directly connected, Loopback2
C       172.16.201.0/24 is directly connected, Serial0/0
C       172.16.202.0/24 is directly connected, Loopback3
C       172.16.100.0/23 is directly connected, Loopback4
D       172.16.101.0/24 [90/2172416] via 11.0.0.1, 00:53:07, FastEthernet0/1
C    10.0.0.0/8 is directly connected, FastEthernet0/0
C    11.0.0.0/8 is directly connected, FastEthernet0/1
     192.168.1.0/32 is subnetted, 1 subnets
D       192.168.1.11 [90/156160] via 11.0.0.1, 00:00:03, FastEthernet0/1
S*   0.0.0.0/0 [1/0] via 11.0.0.1
D    10.0.0.0/7 is a summary, 00:54:40, Null0
```

Ostatni wpis w tabeli, również oznaczony jako trasa skonsolidowana, jest właściwie grupą sieci nadrzędnych, czyli supersiecią.

Rozróżnienie między supersieciami a trasami skonsolidowanymi jest głównie przedmiotem sporów akademickich. W praktyce w obu przypadkach stosuje się rutynowo określenie trasa skonsolidowana lub trasa agregowana. W zależności od protokołu routingu grupy tras mogą być różnie nazywane, zarówno w przypadku podsieci, jak i sieci nadrzędnych — BGP stosuje termin „agregacja", podczas gdy OSPF używa określenia „konsolidacja".

Miejscem docelowym dla obu tras jest w tym przypadku `Null0`. Parametr ten oznacza, że pakiety przesyłane do tej sieci będą odrzucane. Trasy skonsolidowane prowadzą do `Null0`, ponieważ powstały w obrębie protokołu EIGRP na tym routerze.

Trasa `Null0` znajduje się tutaj dla celów protokołu routingu. Bardziej konkretne trasy muszą się również znajdować w tabeli routingu, ponieważ lokalny router wykorzystuje je do przekazywania pakietów. Konkretne trasy nie będą rozgłaszane w protokole routingu. Rozgłaszane będą jedynie trasy skonsolidowane. Możesz zobaczyć to na tym podłączonym routerze:

```
     172.16.0.0/16 is variably subnetted, 4 subnets, 2 masks
D       172.16.200.0/23 [90/156160] via 11.0.0.2, 04:30:21, FastEthernet0/1
D       172.16.202.0/24 [90/156160] via 11.0.0.2, 04:30:21, FastEthernet0/1
D       172.16.100.0/23 [90/156160] via 11.0.0.2, 04:30:21, FastEthernet0/1
C       172.16.101.0/24 is directly connected, Serial0/0
```

Na podłączonym routerze widoczna jest trasa skonsolidowana 172.16.200.0/23, natomiast trasy konkretne 172.16.200.0/24 i 172.16.201.0/24 nie są pokazane.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Sieć nadrzędna

Sieć nadrzędna to sieć w swojej natywnej postaci. Przykładowo sieć 10.0.0.0/8 posiada natywną maskę podsieci 255.0.0.0. Dlatego też jest siecią nadrzędną. Zastosowanie do 10.0.0.0 prefiksu maski dłuższego niż /8 zmienia trasę w podsieć. Z kolei użycie prefiksu krótszego niż /8 zmienia trasę w supersieć.

Sieci nadrzędne są pokazane w tabeli routingu:

```
      172.16.0.0/16 is variably subnetted, 6 subnets, 2 masks
D        172.16.200.0/23 is a summary, 00:56:18, Null0
C        172.16.200.0/24 is directly connected, Loopback2
C        172.16.201.0/24 is directly connected, Serial0/0
C        172.16.202.0/24 is directly connected, Loopback3
C        172.16.100.0/23 is directly connected, Loopback4
D        172.16.101.0/24 [90/2172416] via 11.0.0.1, 00:53:07, FastEthernet0/1
C     10.0.0.0/8 is directly connected, FastEthernet0/0
C     11.0.0.0/8 is directly connected, FastEthernet0/1
      192.168.1.0/32 is subnetted, 1 subnets
D        192.168.1.11 [90/156160] via 11.0.0.1, 00:00:03, FastEthernet0/1
S*    0.0.0.0/0 [1/0] via 11.0.0.1
D     10.0.0.0/7 is a summary, 00:54:40, Null0
```

172.16.0.0/16 również znajduje się na liście, ale jedynie jako odniesienie do grupy podsieci wymienionych poniżej. Wpis 172.16.0.0/16 nie jest trasą.

Supersieć (grupa sieci nadrzędnych)

Supersieć jest grupą sieci nadrzędnych. W tym przykładzie na liście znajduje się trasa do 10.0.0.0/7, która jest grupą sieci nadrzędnych 10.0.0.0/8 oraz 11.0.0.0/8:

```
      172.16.0.0/16 is variably subnetted, 6 subnets, 2 masks
D        172.16.200.0/23 is a summary, 00:56:18, Null0
C        172.16.200.0/24 is directly connected, Loopback2
C        172.16.201.0/24 is directly connected, Serial0/0
C        172.16.202.0/24 is directly connected, Loopback3
C        172.16.100.0/23 is directly connected, Loopback4
D        172.16.101.0/24 [90/2172416] via 11.0.0.1, 00:53:07, FastEthernet0/1
C     10.0.0.0/8 is directly connected, FastEthernet0/0
C     11.0.0.0/8 is directly connected, FastEthernet0/1
      192.168.1.0/32 is subnetted, 1 subnets
D        192.168.1.11 [90/156160] via 11.0.0.1, 00:00:03, FastEthernet0/1
S*    0.0.0.0/0 [1/0] via 11.0.0.1
D     10.0.0.0/7 is a summary, 00:54:40, Null0
```

Zauważ, że ponownie miejscem przeznaczenia trasy jest `Null0`. Z pewnością na podłączonym routerze zobaczysz tylko trasę skonsolidowaną, bez tras konkretnych:

```
D     10.0.0.0/7 [90/30720] via 11.0.0.2, 04:30:22, FastEthernet0/1
```

Trasa domyślna

Trasa domyślna, zwana też „trasą ostatniego wyboru", jest wyświetlona specjalnie nad tabelą routingu, aby była dobrze widoczna:

```
Gateway of last resort is 11.0.0.1 to network 0.0.0.0
      172.16.0.0/16 is variably subnetted, 6 subnets, 2 masks
D        172.16.200.0/23 is a summary, 00:56:18, Null0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
C       172.16.200.0/24 is directly connected, Loopback2
C       172.16.201.0/24 is directly connected, Serial0/0
C       172.16.202.0/24 is directly connected, Loopback3
C       172.16.100.0/23 is directly connected, Loopback4
D       172.16.101.0/24 [90/2172416] via 11.0.0.1, 00:53:07, FastEthernet0/1
C    10.0.0.0/8 is directly connected, FastEthernet0/0
C    11.0.0.0/8 is directly connected, FastEthernet0/1
     192.168.1.0/32 is subnetted, 1 subnets
D       192.168.1.11 [90/156160] via 11.0.0.1, 00:00:03, FastEthernet0/1
S*   0.0.0.0/0 [1/0] via 11.0.0.1
D    10.0.0.0/7 is a summary, 00:54:40, Null0
```

W tym przykładzie trasa domyślna jest trasą statyczną, na co wskazuje oznaczenie `S` w pierwszej kolumnie. Mogłaby być jednak równie dobrze pozyskana z protokołu routingu. Gwiazdka przy `S` informuje, że trasa jest kandydatką na trasę domyślną. Kandydatek może być więcej niż jedna, a wtedy będzie kilka wpisów oznaczonych gwiazdką. Podobnie tras domyślnych może być kilka, ale tylko jedna zostanie wyświetlona w pierwszej linii listingu.

Poniższy listing pochodzi z routera z dwoma aktywnymi bramami domyślnymi, ale tylko jedna wyświetlona jest nad tabelą:

```
Gateway of last resort is 10.0.0.1 to network 0.0.0.0
     20.0.0.0/24 is subnetted, 1 subnets
S       20.0.0.0 [1/0] via 10.0.0.1
     10.0.0.0/24 is subnetted, 3 subnets
C       10.0.0.0 is directly connected, FastEthernet0/0
C    192.168.1.0/24 is directly connected, FastEthernet0/1
S*   0.0.0.0/0 [1/0] via 10.0.0.1
                [1/0] via 10.0.0.2
```

Jeśli masz wątpliwości, przyjrzyj się wpisowi 0.0.0.0 w tabeli routingu, ponieważ zawiera on najbardziej dokładne informacje.

Virtual Routing and Forwarding

Przełączniki Nexus oraz nowsze wersje IOS obsługują funkcję, która nazywana jest instancjami VRF (ang. *Virtual Routing and Forwarding*). Każda instancja VRF jest samodzielną tablicą routingu w obrębie tego samego routera. Nexus 7000 może posiadać kilka instancji w ramach jednego urządzenia wirtualnego VDC (ang. *Virtual Device Context*). W przypadku przełącznika Nexus 5000, który na obecnym etapie nie obsługuje VDC ani routingu, wciąż możesz mieć kilka instancji VRF. W rzeczywistości sieć zarządzająca zawiera się domyślnie w zarządzającym VRF, podczas gdy pozostały ruch sieciowy znajduje się w domyślnym VRF. W IOS w wersji 15.x również możesz konfigurować VRF-y.

Na platformie Nexus istnieją dwa domyślnie ustanowione VRF-y: **zarządzający** (ang. *management*) oraz **domyślny** (ang. *default*):

```
NX-7K-1-Daisy# sho vrf
VRF-Name                    VRF-ID    State   Reason
default                          1    Up      --
management                       2    Up      --
```

Tworzenie nowego VRF jest proste. Wystarczy jedynie polecenie `vrf context` *`nazwa_vrf`*. Teraz utworzę dwa nowe VRF-y o nazwach Earth i Mars:

```
NX-7K-1-Daisy(config)# vrf context Earth
NX-7K-1-Daisy(config)# vrf context Mars
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Teraz możesz podejrzeć wszystkie instancje VRF, stosując komendę `show vrf`:

```
NX-7K-1-Daisy(config)# sho vrf
VRF-Name                    VRF-ID    State   Reason
Earth                            3    Up      --
Mars                             5    Up      --
default                          1    Up      --
management                       2    Up      --
```

Do VRF można przypisywać konkretne interfejsy. Pamiętaj, że w przełącznikach Nexus interfejsy są domyślnie ustawione w trybie rutowania (ang. *routed*), więc nie trzeba ich dodatkowo konfigurować. Aby dodać interfejs do VRF, wystarczy polecenie interfejsu `vrf member` *`nazwa_vrf`*:

```
NX-7K-1-Daisy(config)# int e3/25
NX-7K-1-Daisy(config-if)# vrf member Earth
NX-7K-1-Daisy(config-if)# ip address 10.0.0.1/24
```

Teraz przypiszę inny interfejs do drugiego VRF:

```
NX-7K-1-Daisy(config-if)# int e3/26
NX-7K-1-Daisy(config-if)# vrf member Mars
NX-7K-1-Daisy(config-if)# ip address 10.0.0.1/24
```

Zauważ, że mam na jednym routerze skonfigurowane dwa interfejsy z tym samym adresem IP. Jest to możliwe, ponieważ każdy z nich należy do innego VRF. Tablice routingu VRF Earth i VRF Mars są całkowicie od siebie oddzielone. To dosyć potężne narzędzie dodające kolejną warstwę wirtualizacji do platformy Nexus.

Instancje VRF są wyraźnie koncepcją warstwy trzeciej, nie możesz więc przypisywać VLAN-ów do VRF. Możesz jednak przypisywać do nich interfejsy VLAN-ów.

VRF-y mogą być frustrujące, jeśli nie jesteś do nich przyzwyczajony. W naszym przykładzie komenda `sho ip route` nie wyświetli skonfigurowanych interfejsów. Dzieje się tak, ponieważ domyślny VRF ma właśnie nazwę **domyślny** (ang. *default*). Skoro jest teraz kilka tablic routingu, aby zobaczyć tablicę wybranego VRF, musisz w poleceniu określić jego nazwę:

```
NX-7K-1-Daisy# sho ip route vrf Earth
IP Route Table for VRF "Earth"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
10.0.0.0/24, ubest/mbest: 1/0, attached
    *via 10.0.0.1, Eth3/25, [0/0], 02:49:03, direct
10.0.0.1/32, ubest/mbest: 1/0, attached
    *via 10.0.0.1, Eth3/25, [0/0], 02:49:03, local
```

W Nexusie większość komend odnoszących się do warstwy trzeciej obsługuje parametr VRF, tak jak w tym przykładzie:

```
NX-7K-1-Daisy# sho ip eigrp neighbors vrf Earth
IP-EIGRP neighbors for process 0 VRF Earth
IP-EIGRP neighbors for process 100 VRF Earth
H   Address                 Interface       Hold  Uptime  SRTT   RTO  Q  Seq
                                            (sec)         (ms)       Cnt Num
0   10.0.0.2                Eth3/25         11   00:00:50  3    200   0   2
```

W ten sposób działa nawet komenda `ping`:

```
NX-7K-1-Daisy# ping 10.0.0.2 vrf Earth
PING 10.0.0.2 (10.0.0.2): 56 data bytes
Request 0 timed out
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
64 bytes from 10.0.0.2: icmp_seq=1 ttl=254 time=0.884 ms
64 bytes from 10.0.0.2: icmp_seq=2 ttl=254 time=0.538 ms
64 bytes from 10.0.0.2: icmp_seq=3 ttl=254 time=0.597 ms
64 bytes from 10.0.0.2: icmp_seq=4 ttl=254 time=0.6 ms
```

Aby to nieco ułatwić, możesz przełączyć się na wybrany VRF, tak żeby wszystkie kolejne komendy dotyczyły właśnie tej instancji. Służy do tego polecenie `routing-context`, które nie jest poleceniem konfiguracyjnym:

```
NX-7K-1-Daisy# routing-context vrf Earth
NX-7K-1-Daisy%Earth#
```

Polecenie to modyfikuje nagłówek wiersza poleceń, dodając nazwę VRF po znaku %, żeby przypomnieć nam, gdzie jesteśmy. Teraz komenda `show ip route` (lub inna dowolna) będzie odnosić się do bieżącego kontekstu routingu, którym jest Earth:

```
NX-7K-1-Daisy%Earth# sho ip route
IP Route Table for VRF "Earth"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
10.0.0.0/24, ubest/mbest: 1/0, attached
    *via 10.0.0.1, Eth3/25, [0/0], 05:01:14, direct
10.0.0.1/32, ubest/mbest: 1/0, attached
    *via 10.0.0.1, Eth3/25, [0/0], 05:01:14, local
200.200.200.0/24, ubest/mbest: 1/0
    *via 10.0.0.2, Eth3/25, [90/130816], 00:21:50, eigrp-100, internal
```

Jeśli chcesz powrócić do domyślnego VRF, zastosuj polecenie `routing-context vrf default`. Nagłówek wiersza powróci do poprzedniej postaci:

```
NX-7K-1-Daisy%Earth# routing-context vrf default
NX-7K-1-Daisy#
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 10.

Protokoły routingu

Protokół routingu (ang. *routing protocol*) jest środkiem wymiany (między urządzeniami) informacji na temat stanu sieci. Informacje zbierane od innych urządzeń są wykorzystywane do wytyczania ścieżek dla przepływu pakietów do każdej sieci docelowej.

Protokoły routingu są zarówno protokołami, jak i aplikacjami. Same protokoły umiejscowione są w warstwie trzeciej modelu ISO/OSI, podczas gdy aplikacje odpowiadające za wytyczanie tras działają w warstwie siódmej. Istnieje wiele protokołów routingu, jednak obecnie w powszechnym użyciu jest tylko kilka z nich. Starsze protokoły są rzadko wykorzystywane, trzeba jednak mieć świadomość, że w wielu sieciach pracują jeszcze urządzenia, które obsługują tylko te rozwiązania. Niektóre firewalle i serwery mogą oferować ograniczony zakres protokołów routingu — najczęściej protokoły RIP (ang. *Routing Information Protocol*) i OSPF (ang. *Open Short Path First*) — ale dla uproszczenia będę nazywał wszystkie urządzenia pracujące w protokołach routingu **routerami**.

Dzięki protokołom routingu sieci są dynamiczne i mniej awaryjne. Gdyby wszystkie trasy w sieci były trasami statycznymi, jedyną dynamiczną formą routingu, jaką można by zastosować, byłaby **pływająca trasa statyczna** (ang. *floating static route*). Pływająca trasa statyczna staje się aktywna, kiedy inna trasa statyczna zostanie usunięta z tablicy routingu. Oto przykład:

```
ip route 0.0.0.0 0.0.0.0 192.168.1.1 1
ip route 0.0.0.0 0.0.0.0 10.0.0.1 2
```

Podstawowa trasa domyślna prowadzi do 192.168.1.1 i posiada metrykę 1. Druga trasa domyślna, posiadająca metrykę 2, prowadzi do 10.0.0.1.

W tablicy routingu umieszczane są trasy z najlepszą metryką, czyli w tym przypadku będzie to pierwsza z wymienionych tras. Jeśli sieć 192.168.1.0 stanie się niedostępna, wszystkie trasy prowadzące do niej zostaną usunięte z tablicy routingu. W tym momencie w tablicy tej zostanie umieszczona domyślna trasa 10.0.0.1 jako trasa z najlepszą obecnie metryką dla sieci 0.0.0.0/0.

Pływająca trasa statyczna umożliwia zmiany tras, jeśli awarii ulegnie bezpośrednio podłączony interfejs. Nie zapewni jednak niezawodności tras, gdy zawiedzie zdalne urządzenie lub łącze. Protokoły routingu **dynamicznego** (ang. *dynamic*) pozwalają zazwyczaj wszystkim routerom pracującym w danym protokole uczyć się o wszelkich awariach w sieci dzięki regularnej komunikacji między routerami.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Komunikacja między routerami

Routery muszą komunikować się ze sobą, aby uczyć się o stanie sieci. RIP, jeden z pierwszych protokołów routingu, wysyła aktualizacje dotyczące sieci, wykorzystując pakiety broadcastowe. Zdawało to egzamin w przypadku mniejszych sieci, ale kiedy sieci zaczęły być coraz większe, broadcasty te stały się problemem. Każdy host w sieci nasłuchuje broadcastów, a w przypadku RIP mogą one być całkiem spore.

Większość nowoczesnych protokołów routingu wykorzystuje do komunikacji w sieciach broadcastowych **pakiety multicast** (ang. *multicast packets*). Mają one konkretne adresy IP oraz odpowiadające im adresy MAC przypisane do określonej grupy odbiorców.

Ponieważ routing to proces raczej dynamiczny, routery już działające muszą być w stanie wykrywać nowe routery, aby dodawać ich informacje do tabel opisujących sieć. Przykładowo wszystkie routery EIGRP (ang. *Enhanced Internal Gateway Routing Protocol*) w obrębie tej samej domeny muszą mieć możliwość komunikowania się ze sobą. Definiowanie konkretnych sąsiadów nie jest wymagane w tym protokole, ponieważ są oni wykrywani dynamicznie.

Większość protokołów bramy wewnętrznej wykrywa sąsiadów dynamicznie. BGP (ang. *Border Gateway Protocol*) nie wykrywa sąsiadów. Zamiast tego BGP musi zostać skonfigurowany ręcznie, żeby komunikował się ze wszystkimi sąsiadami.

Organizacja IANA (ang. *Internet Assigned Numbers Authority*) publikuje zestawienie wykorzystanych adresów multicast na stronie *http://www.iana.org/assignments/multicast-addresses*. Poniżej najbardziej powszechne adresy multicast:

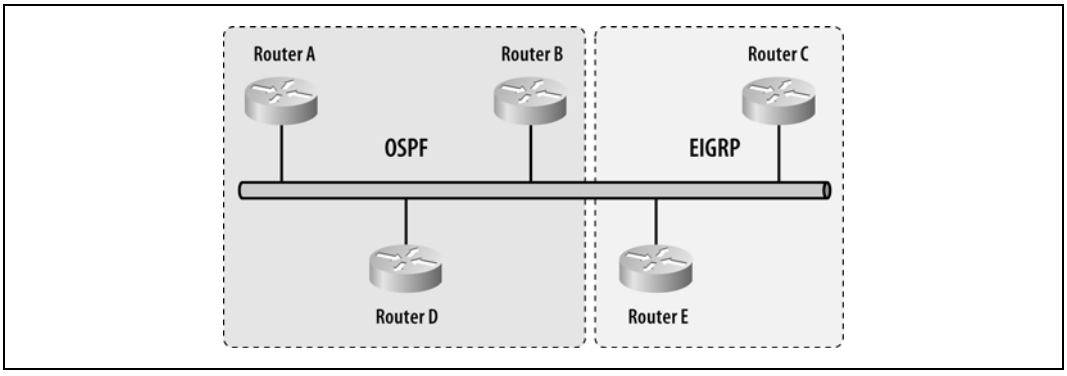
```
224.0.0.0  Adres bazowy (Zarezerwowany)                  [RFC 1112,JBP]
224.0.0.1  Wszystkie systemy w tej podsieci               [RFC 1112,JBP]
224.0.0.2  Wszystkie routery w tej podsieci                        [JBP]
224.0.0.4  Routery DVMRP                          [RFC 1075,JBP]
224.0.0.5  Wszystkie routery OSPFIGP             [RFC 2328,JXM1]
224.0.0.6  Routery wyznaczone OSPFIGP      [RFC 2328,JXM1]
224.0.0.9  Routery RIP2                           [RFC 1723,GSM11]
224.0.0.10 Routery IGRP                              [Farinacci]
224.0.0.12 Serwer DHCP / Relay Agent                    [RFC 1884]
224.0.0.18 VRRP                                         [RFC 3768, 5798]
224.0.0.102 HSRP                                         [Wilson]
```

Lista pokazuje, że routery IGRP (ang. *Internal Gateway Routing Protocol*), włączając w to routery Enhanced IGRP, będą nasłuchiwać pakietów wysłanych na adres 224.0.0.10.

Nie wszystkie protokoły routingu wykorzystują do komunikacji multicasty. Skoro BGP nie wykrywa sąsiadów, nie musi stosować multicastów, a zamiast tego używa pakietów unicast (ang. *unicast packets*). Również wiele innych protokołów routingu umożliwia statyczną konfigurację sąsiadów. Dzięki temu do określonych routerów wysyłane są informacje unicastowe zamiast multicastowych.

W pojedynczej sieci może być stosowany więcej niż jeden protokół routingu. Rysunek 10.1 przedstawia przykładową sieć Ethernet, w której pracuje pięć routerów. Trzy z nich działają w protokole OSPF, pozostałe dwa wykorzystują EIGRP. Routery EIGRP nie potrzebują otrzymywać aktualizacji OSPF i vice versa. Użycie multicastów zapewnia, że tylko routery działające w tym samym protokole będą się ze sobą komunikować oraz będą się nawzajem wykrywać.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 10.1. Kilka protokołów routingu w jednej sieci Ethernet

Sieć może zawierać kilka instancji tego samego protokołu routingu. Takie wydzielone obszary kontroli są nazywane **systemami autonomicznymi** (ang. *autonomous systems*) w przypadku protokołu EIGRP lub **domenami** (ang. *domains*), jeśli chodzi o OSPF (chociaż termin system autonomiczny jest często używany nieprawidłowo). W EIGRP każda instancja jest oznaczana numerem systemu autonomicznego ASN (ang. *autonomous system number*). Analogiczna konfiguracja OSPF jest bardziej skomplikowana i wykracza poza ramy tej książki. W routerach Cisco konfiguracja OSPF odbywa się z wykorzystaniem parametru **ID-procesu** (ang. *process-ID*). ID-procesu jest charakterystyczne lokalnie dla routera. W obrębie pojedynczego routera trasy pomiędzy wieloma procesami OSPF nie są automatycznie redystrybuowane. Jednak pomiędzy dwoma routerami OSPF, z których każdy ma inne ID-procesu, zachodzi przyleganie sąsiadów (zakładając, że spełnione są pozostałe niezbędne warunki). ID-procesu w OSPF nie odwołuje się do domeny.

Rysunek 10.2 przedstawia sieć z dwoma aktywnymi procesami EIGRP. Ponieważ pakiety multicast wysyłane przez router EIGRP są przeznaczone dla wszystkich pozostałych routerów EIGPR, będą one nasłuchiwać aktualizacji. Aktualizacje zawierają ID-systemu-autonomicznego (ang. *autonomous-system-ID*), tak więc routery mogą indywidualnie zdecydować, czy je zachować, czy odrzucić. Protokół RIP nie obsługuje koncepcji oddzielnych procesów, dlatego każdy router RIP będzie odbierał i przetwarzał aktualizacje od pozostałych routerów RIP w sieci.

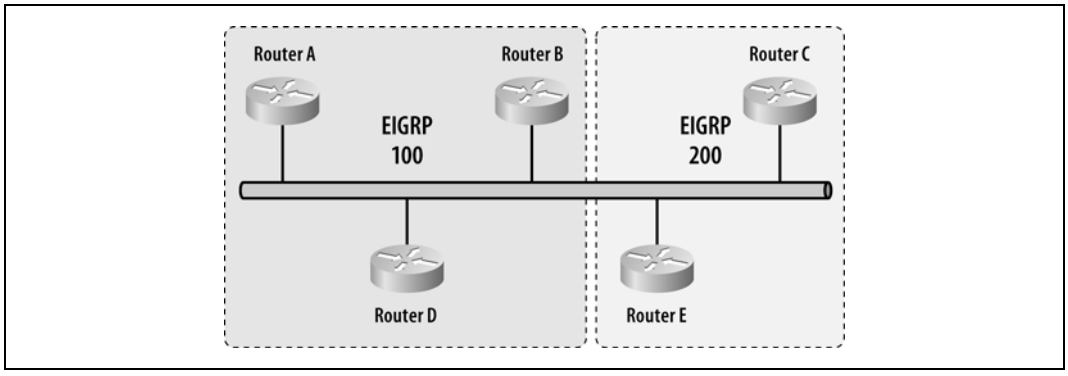
Kiedy w tej samej sieci są dwa systemy autonomiczne, informacje o trasach pozyskanych z obu z nich mogą (lub nie) być domyślnie współdzielone między procesami. Jedną z cech EIGRP jest automatyczna redystrybucja pomiędzy procesami. Aby trasy były dzielone z innymi protokołami, jeden z routerów musi uczestniczyć w obu procesach i być skonfigurowany do dzielenia się informacjami o trasach między nimi.

Procedura przekazywania informacji o trasach z jednego procesu lub protokołu routingu do drugiego procesu lub protokołu routingu nazywa się **redystrybucją** (ang. *redistribution*). Przykład redystrybucji dla kilku procesów routingu EIGRP pokazuje rysunek 10.3.

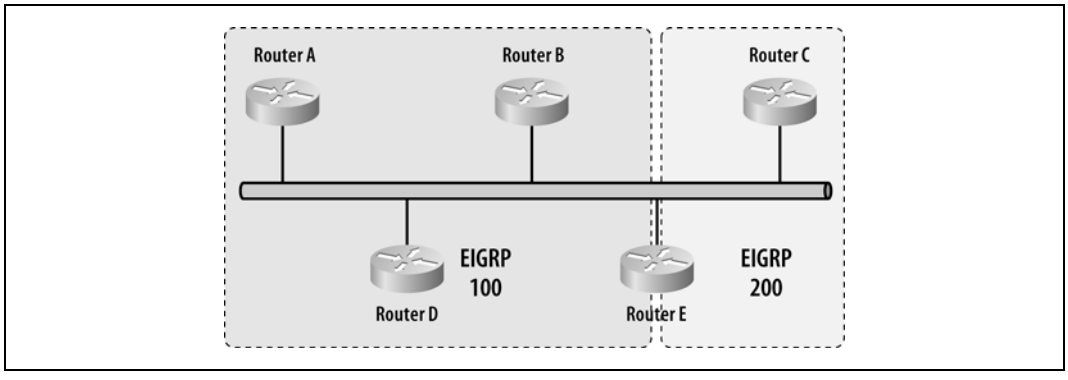
Rysunek 10.2 pokazuje dwa autonomiczne systemy EIGRP, które nie dzielą się wzajemnie informacjami o swoich trasach. Na rysunku 10.3 Router E jest skonfigurowany jako członek grup EIGRP 100 oraz EIGRP 200. Router ten może więc dystrybuować pomiędzy procesami informacje o trasach uzyskane w każdym AS.

Kiedy informacja o trasie pozyskana jest z systemu autonomicznego obsługiwanego przez dany protokół, taka trasa jest nazywana trasą **wewnętrzną** (ang. *internal*). Gdy informacje o trasie

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 10.2. Dwa procesy EIGRP w pojedynczej sieci



Rysunek 10.3. Redystrybucja protokołów routingu

pochodzą spoza protokołu routingu i są dystrybuowane w danym procesie, to jest to trasa **zewnętrzna** (ang. *external*). Trasy wewnętrzne są z reguły traktowane jako bardziej niezawodne niż trasy zewnętrzne. Decyzja podejmowana jest na podstawie metryki tras, zwanej **dystansem administracyjnym** (ang. *administrative distance*) — opisanej w dalszej części rozdziału. Wyjątkiem może być protokół BGP, który przedkłada trasy zewnętrzne nad wewnętrzne, oraz OSPF, który nie przydziela innej wartości dystansu administracyjnego trasom zewnętrznym w stosunku do wewnętrznych.

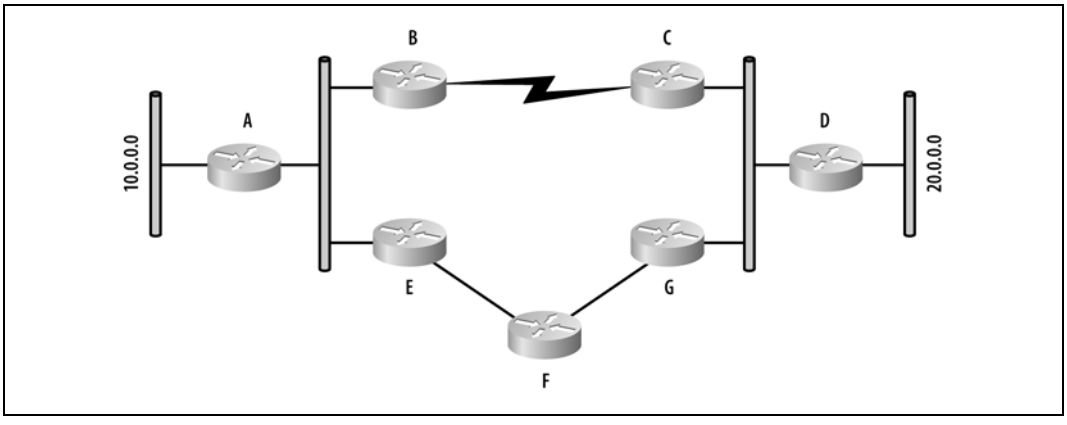
Metryki i rodzaje protokołów

Zadaniem protokołu routingu jest wyznaczanie najlepszej ścieżki do sieci docelowej. Taka najbardziej optymalna trasa wybierana jest na podstawie zestawu reguł charakterystycznego dla każdego protokołu. RIP decyduje na podstawie liczby przeskoków (routerów) między sieciami, podczas gdy OSPF oblicza koszt trasy, uwzględniając przepustowość wszystkich łączy w sieci. EIGRP stosuje domyślnie do określania trasy raportowaną przepustowość linków oraz opóźnienia, możesz jednak zdefiniować dodatkowe parametry. Każdy protokół przypisuje poszczególnym trasom pewną wartość zwaną **metryką** (ang. *metric*). Trasy z niższą metryką są bardziej pożądane.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Najłatwiejszą do zrozumienia formą metryki jest prawdopodobnie ta stosowana przez protokół RIP: liczba przeskoków (ang. *hop count*). Jest to liczba routerów pomiędzy routerem wyznaczającym ścieżkę a siecią docelową.

Rozważmy następujący przykład. Rysunek 10.4 przedstawia dwie sieci oznaczone jako 10.0.0.0 i 20.0.0.0. Router A znajduje dwie możliwe ścieżki do sieci 20.0.0.0: jedna prowadzi przez Router B, a druga przez Router E. W pierwszym przypadku trzeba przejść przez Routery B, C i D, co daje liczbę przeskoków 3 dla tej trasy. W drugim przypadku mamy po drodze Routery E, F, G i D, co w sumie daje 4 przeskoki.



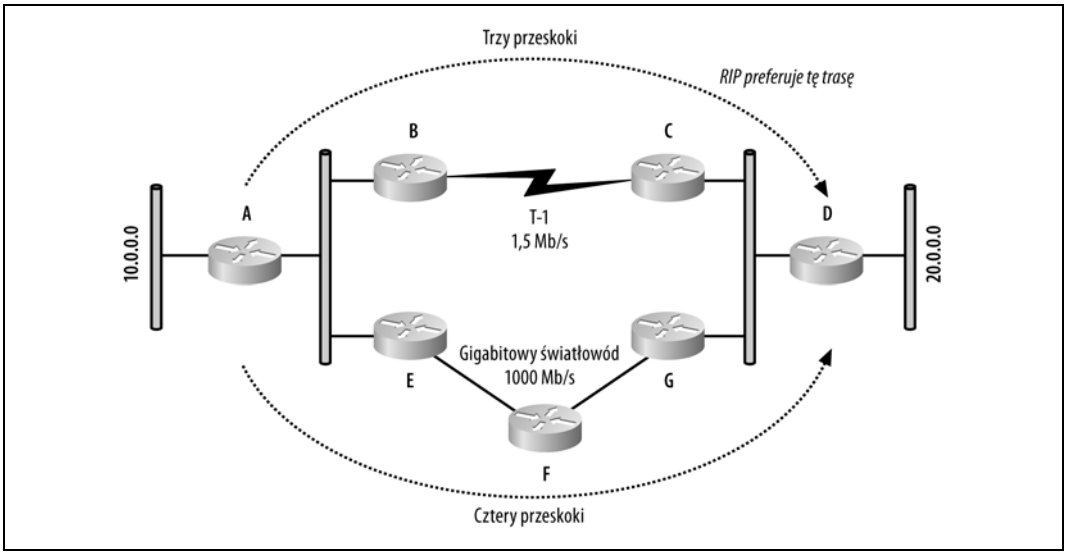
Rysunek 10.4. Przykład metryk w protokołach routingu

Wygrywa zawsze niższa metryka, więc Router A uzna trasę wiodącą przez Router B za najlepszą. Ten router zostanie dodany do tabeli routingu z metryką 3.

Stosowanie liczby przeskoków jako metryki ma swoje ograniczenia i może powodować wybranie mniej optymalnej trasy. Patrząc na rysunek 10.5, zauważysz, że link między Routerem B i C jest łączem w standardzie T1 pracującym z prędkością 1,54 Mb/s, podczas gdy linki między Routerami E, F i G są bezpośrednimi łączami światłowodowymi o prędkości 1 Gb/s. Oznacza to, że ścieżka prowadząca przez Routery E, F i G jest o wiele szybsza niż link między B i C, choć ten ma mniej przeskoków. RIP nie uwzględnia przepustowości wykorzystywanych łączy, a jedynie liczbę przeskoków.

Protokół taki jak RIP jest nazywany protokołem **wektora odległości** (ang. *distance-vector*), ponieważ przy określaniu trasy do sieci docelowej bazuje na dystansie między dwoma punktami. Protokoły wektora odległości mają jeszcze jedną wadę, zwaną **liczeniem do nieskończoności** (ang. *counting to infinity*). RIP i inne podobne protokoły wyznaczają górny limit dozwolonej liczby przeskoków dla osiągnięcia sieci docelowej. Jeśli trasa wymaga liczby przeskoków większej niż dopuszczalna, uznawana jest za nieosiągalną. Maksymalna liczba przeskoków dla protokołu RIP to 15, a 16 to już dystans nieosiągalny. Jak możesz sobie wyobrazić, nie jest to wystarczające w nowoczesnych sieciach, w których dość często trasa przechodzi przez więcej niż 16 routerów. Aktualne wersje RIP (zwane RIPv2 lub RIP2) zwiększają ten limit do 255 przeskoków, gdzie przy 256 dystans jest nieosiągalny. Ponieważ RIP nadal nie analizuje stanu i możliwości łączy przeskoków, w większości sieci stosuje się obecnie nowsze i bardziej niezawodne protokoły routingu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 10.5. RIP wykorzystuje liczbę przeskoków do określania najlepszych ścieżek

Z kolei protokoły routingu takie jak OSPF zwane są protokołami **stanu łącza** (ang. *link-state*) lub też protokołami opartymi na algorytmie Dijkstry. Bazują one na informacjach dotyczących stanu łączy pomiędzy źródłowym routerem a siecią docelową, a nie na liczeniu routerów między nimi.

W przypadku OSPF dodatkowym parametrem jest **koszt** (ang. *cost*) każdego linku. Koszt wyznacza się, dzieląc 100 000 000 przez przepustowość łącza podaną w bitach na sekundę (b/s). Oto kilka przykładów:

100 Mb/s (100 000 000 / 100 000 000 b/s) = 1

10 Mb/s (100 000 000 / 10 000 000 b/s) = 10

1,5 Mb/s (100 000 000 / 1 540 000 b/s) = 64 (w przybliżeniu)

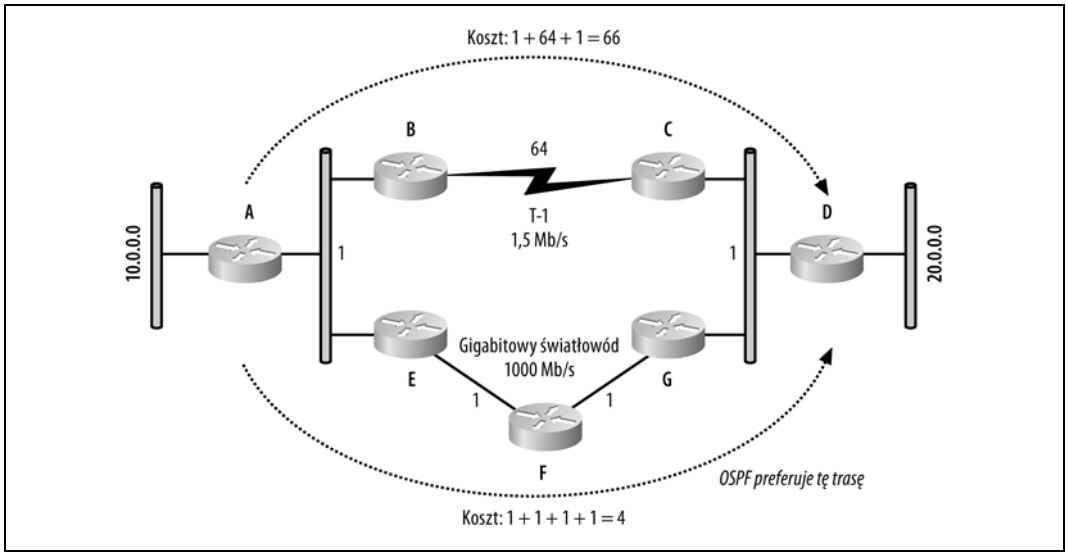
Rysunek 10.6 pokazuje tę samą sieć przedstawioną w poprzednim przykładzie. Tym razem protokół OSPF wyznacza najlepszą trasę do sieci docelowej na podstawie metryk przepustowości łączy. Metryka linku T1 to 64, podczas gdy metryka ścieżki gigabitowej wynosi 4. Ponieważ metryka łącza przechodzącego przez Routery E, F i G jest niższa niż metryka ścieżki przez Routery B i C, to właśnie ta trasa jest zapisywana w tabeli routingu.

EIGRP stosuje bardziej skomplikowaną formułę do obliczania kosztów. Określanie metryk może w tym przypadku opierać się na przepustowości, opóźnieniu, niezawodności, przepustowości efektywnej oraz parametrze MTU (ang. *maximum transmission unit*). EIGRP uznawany jest za protokół **hybrydowy** (ang. *hybrid*).

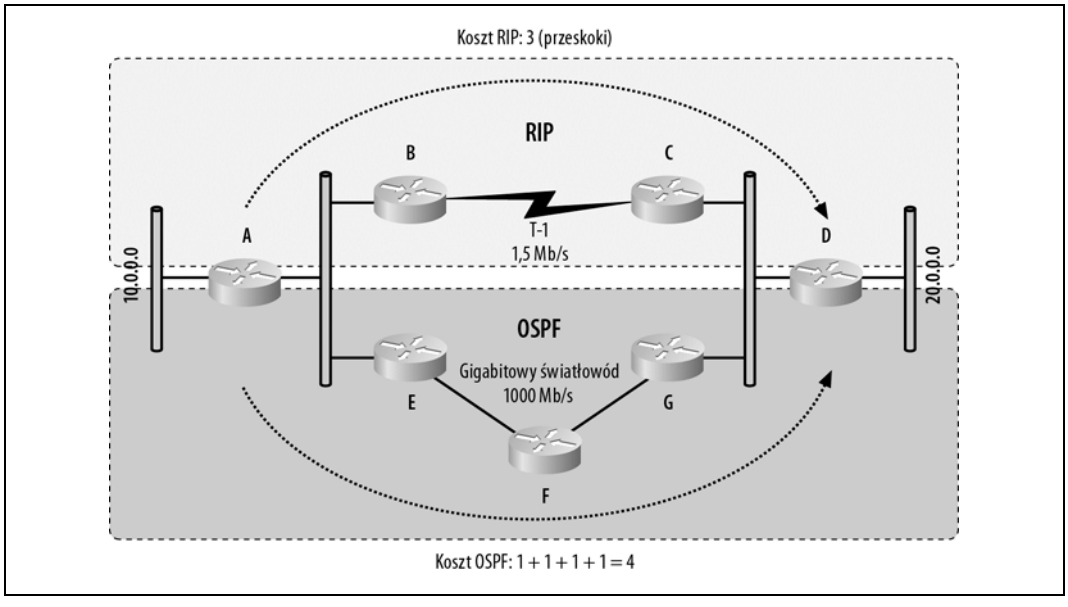
Dystans administracyjny

Sieci z reguły posiadają więcej niż jeden aktywny protokół routingu. W takiej sytuacji istnieje duże prawdopodobieństwo, że te same sieci będą rozgłaszane przez kilka protokołów routingu. Rysunek 10.7 pokazuje sieć, w której działają dwa takie protokoły: w segmencie obejmującym

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 10.6. OSPF uwzględnia przepustowość łączy przy wyborze najlepszych tras



Rysunek 10.7. Rywalizujące protokoły routingu

górną połowę schematu operuje RIP, a w dolnej połowie OSPF. Router A będzie otrzymywał informacje o trasach do sieci 20.0.0.0 zarówno z protokołu RIP, jak i OSPF. Trasa RIP posiada lepszą metrykę, ale jak już widziałeś, OSPF wykorzystuje bardziej odpowiednie parametry do określania tras. W jaki sposób jest więc wybierana najlepsza trasa?

Routery dokonują wyboru tras, opierając się na predefiniowanym zestawie reguł. Jednym z czynników decydujących o umieszczeniu trasy w tabeli routingu jest **dystans administracyjny** (ang. *administrative distance*). Dystans administracyjny to wartość przypisana każdemu

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

protokołowi routingu. W przypadku gdy dwa protokoły zgłaszają tę samą trasę, wybierany jest ten, który posiada niższy dystans administracyjny. Jego wersja trasy jest wtedy umieszczana w bazie RIP.

Dystans administracyjny różnych protokołów routingu przedstawia tabela 10.1.

Tabela 10.1. Dystans administracyjny protokołów routingu

Rodzaj trasy	Dystans administracyjny
Interfejs podłączony bezpośrednio	0
Trasa statyczna	1
EIGRP trasa skonsolidowana	5
Zewnętrzny BGP	20
Wewnętrzny EIGRP	90
IGRP	100
OSPF	110
IS-IS	115
RIP	120
EGP	140
ODR	160
Zewnętrzny EIGRP	170
Wewnętrzny BGP	200
Nieznany	255

Trasa statyczna do interfejsu podłączonego bezpośrednio ma dystans administracyjny 0 i jest to jedyna trasa, która może zastąpić normalną trasę statyczną. Trasa z dystansem administracyjnym 255 nie jest zaufana, nie zostanie więc zapisana w tabeli routingu.

Patrząc na tabelę 10.1, zauważysz, że RIP posiada dystans administracyjny 120, podczas gdy OSPF ma wskaźnik 110. Oznacza to, że mimo iż RIP na rysunku 10.7 ma lepszą metrykę, w tabeli routingu zostanie zapisana trasa zgłoszona przez protokół OSPF.

Wybrane protokoły routingu

Na temat każdego z protokołów routingu omówionych w tym rozdziale napisano już całe książki. Moim celem nie jest więc nauczenie Cię wszystkiego, co powinieneś wiedzieć o tych protokołach, ale raczej wprowadzenie w temat i zebranie informacji niezbędnych do ich uruchomienia. Opiszę też kilka poleceń powszechnie stosowanych przy rozwiązywaniu problemów z działaniem protokołów routingu.

Protokoły routingu dzielą się na rodzaje ze względu na ich przeznaczenie oraz sposób działania. Głównym rozróżnieniem jest podział na protokoły bramy wewnętrznej i protokoły bramy zewnętrznej.

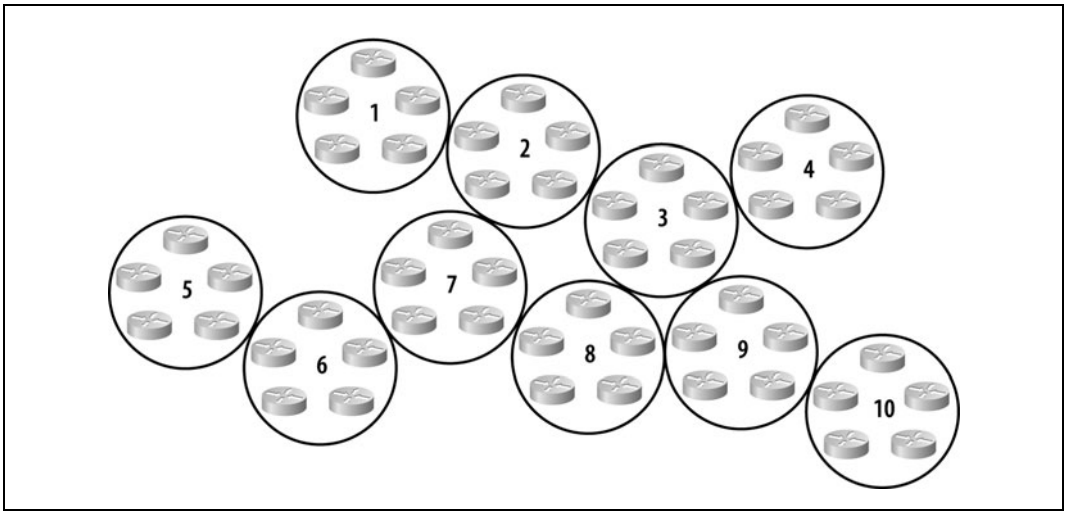
Protokół bramy wewnętrznej (ang. *internal gateway protocol*), czyli IGP, przeznaczony jest do utrzymania tras w obrębie **systemu autonomicznego** (ang. *autonomous system*). Systemem autonomicznym jest każda grupa urządzeń kontrolowanych przez jeden podmiot. Przykładem

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

takiej sytuacji może być przedsiębiorstwo lub szkoła, ale skala nie musi być tak duża — system autonomiczny może tworzyć piętro budynku lub departament przedsiębiorstwa. Do protokołów IGP można zaliczyć RIP, EIGRP oraz OSPF.

Protokół bramy zewnętrznej (ang. *external gateway protocol*), czyli EGP, przeznaczony jest do łączenia ze sobą systemów autonomicznych. Doskonałym przykładem implementacji EGP na dużą skalę jest internet. Systemy autonomiczne — grupy urządzeń kontrolowanych przez poszczególnych dostawców usług, szkoły, przedsiębiorstwa itp. — są samodzielne. Są one sterowane wewnętrznie przez IGP, a połączone ze sobą za pomocą EGP (w przypadku internetu BGP).

Rysunek 10.8 pokazuje, jak różne systemy autonomiczne mogą być połączone. Każdy okrąg przedstawia jeden system autonomiczny. Protokół IGP działający w każdym systemie autonomicznym nie jest powiązany z protokołem bramy zewnętrznej. EGP ma jedynie informację, że określona sieć należy do danego systemu autonomicznego. Załóżmy, że w obrębie ASN 1 jest sieć 1.0.0.0/8, w ASN 2 sieć 2.0.0.0/8, w ASN 3 sieć 3.0.0.0/8 itd. Jeśli urządzenie znajdujące się w ASN 1 chce osiągnąć sieć 10.0.0.0/8, musi wybrać ścieżkę wiodącą przez systemy autonomiczne 1, 2, 3, 9 i 10. Alternatywna droga wiodłaby przez systemy autonomiczne 1, 2, 7, 8, 9 i 10 lub nawet przez 1, 2, 7, 8, 3, 9 oraz 10. Podobnie jak w przypadku wektora odległości IGP, który liczy przeskoki, im mniejsza liczba systemów autonomicznych do pokonania po drodze, tym bardziej atrakcyjna trasa.



Rysunek 10.8. Połączenia między systemami autonomicznymi

Najważniejszą cechą do zapamiętania, jeśli chodzi o EGP, jest to, że protokół ten nie zwraca uwagi na liczbę routerów czy prędkość łączy. Jedyną rzeczą, jaką EGP uwzględnia, jest pokonanie jak najmniejszej liczby systemów autonomicznych w drodze do celu.

Zanim przejdziemy do dalszej części, zdefiniujmy kilka podstawowych pojęć dotyczących routingu:

Protokół routingu klasowego

Protokół routingu klasowego nie przewiduje wsparcia dla podsieci. Rozgłaszany jest zawsze naturalny stan sieci. Przykładowo sieć 10.0.0.0 będzie zawsze rozgłaszana z maską podsieci 255.0.0.0 (/8), bez względu na to, jaka maska podsieci jest aktualnie używana. Protokołami routingu klasowego są RIPv1 oraz IGRP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Protokół routingu bezklasowego

Protokół routingu bezklasowego umieszcza w rozgłoszeniach maski podsieci. Wszystkie nowoczesne protokoły są bezklasowe. Protokołami routingu bezklasowego są EIGRP oraz OSPF.

Poison reverse

Jeśli jeden z routerów chce przekazać drugiemu informację, że jakaś sieć nie jest dłużej opłacalna, może skorzystać z metody zwanej **route poisoning**, czyli zatruwanie tras. Jako przykład rozważmy RIPv1. Jak pamiętasz, metryka o wartości 16 jest uznawana za nieosiągalną. Router może wysłać aktualizację, w której nada sieci metrykę 16, **zatruwając** (ang. *poisoining*) w ten sposób wpis w tablicy routingu routera otrzymującego daną aktualizację. Kiedy router otrzyma zatrutą aktualizację, odsyła taką samą aktualizację do nadawcy. To odbicie route poisoning zwane jest **poison reverse**. Protokoły routingu wektora odległości (łącznie z hybrydowym EIGRP) stosują route poisoning, podczas gdy protokoły stanu łącza, jak OSPF, tego nie robią.

Split horizon

Split horizon to technika stosowana przez wiele protokołów routingu do eliminowania pętli routingu. Kiedy Split horizon jest włączony, trasy, które zostały pozyskane z protokołów routingu, nie są rozgłaszane przez ten sam interfejs, przez jaki się dostały. Zasada ta może być problematyczna w topologiach obwodów wirtualnych, takich jak Frame Relay czy ATM. Jeśli informacja o trasie została pozyskana na jednym ze stałych obwodów wirtualnych PVC (ang. *permanent virtual circuit*) w interfejsie Frame Relay, są szanse, że inny PVC będzie potrzebował aktualizacji, ale nigdy jej nie otrzyma. Dzieje się tak, dlatego że obydwa obwody PVC istnieją na tym samym interfejsie fizycznym. Subinterfejsy Frame Relay są często preferowaną metodą radzenia sobie z problemami Split horizon.

Konwergencja

Mówimy, że sieć jest **konwergentna** (ang. *convergend*), czyli zbieżna, kiedy wszystkie routery w tej sieci otrzymały i przetworzyły wszystkie aktualizacje. Zasadniczo ten stan można osiągnąć, kiedy sieć jest stabilna. Za każdym razem, gdy zmienia się status łącza, protokoły routingu muszą propagować tę zmianę zarówno poprzez aktualizacje czasowe, jak i aktualizacje wyzwalane. W przypadku aktualizacji czasowych zdarza się, że są one wysyłane, mimo iż nie ma potrzeby dokonywania żadnych zmian. Wtedy sieć jest zbieżna.

Jak już wcześniej wspomniałem, istnieje wiele protokołów routingu, jednak na szczęście tylko kilka z nich jest powszechnie stosowanych. Każdy z nich ma swoje cechy charakterystyczne. W kolejnych sekcjach postaram się omówić zasadnicze koncepcje dotyczące najbardziej popularnych protokołów i pokazać, jak skonfigurować te protokoły do najbardziej typowych zastosowań. Nie ma jednego słusznego sposobu konfiguracji protokołów routingu, jednak niektóre z nich są z pewnością lepsze niż pozostałe. Przy projektowaniu sieci pamiętaj, że prostota jest bardzo ważną zasadą, która pozwoli Ci zaoszczędzić wiele godzin ślęczenia nad rozwiązywaniem problemów.

RIP

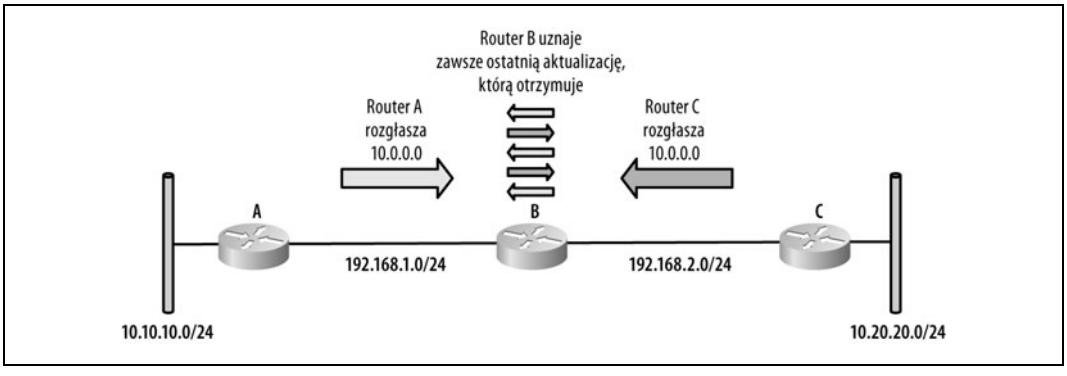
RIP jest najprostszym ze stosowanych obecnie protokołów routingu. O ile staram się mówić moim klientom, że **wszystko, co proste, jest dobre**, o tyle nie uważam, żeby RIP spełniał te warunki.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

RIP wysyła co około 30 sekund broadcasty do wszystkich routerów, które zna, bez względu na statusy pozostałych routerów w sieci. Ponieważ RIP wykorzystuje broadcasty, każdy host w sieci nasłuchuje aktualizacji, chociaż nie każdy może je przetworzyć. W bardziej rozległych sieciach aktualizacje mogą być całkiem spore i pochłaniać znaczną część pasma na kosztownych łączach WAN.

Kolejnym minusem protokołu RIP jest to, że nie stosuje on aktualizacji wyzwalanych. **Aktualizacja wyzwalana** (ang. *triggered update*) to taka, która wysyłana jest w przypadku zaistnienia zmian w sieci. **Aktualizacje niewyzwalane** (ang. *nontriggered updates*), czyli **czasowe** (ang. *timed*), są wysyłane regularnie według harmonogramu. W połączeniu z regułą, że aktualizacje wysyłane są co 30 sekund, doprowadza to do tego, że sieci wykorzystujące protokół RIP konwergują bardzo powoli. Powolna konwergencja jest nieakceptowana w większości nowoczesnych sieci wymagających usuwania awarii i konwergencji w kilka sekund. W sieci RIP, w której znajduje się zaledwie 5 routerów, konwergencja może zająć około 2 minut lub dłużej.

RIP jest protokołem klasowym, co oznacza, że maski podsieci nie są rozgłaszane. Jest to również ograniczenie nieakceptowalne w większości sieci. Rysunek 10.9 przedstawia jedną z kilku najbardziej powszechnych pułapek stosowania protokołów klasowych, takich jak RIP. Router A rozgłasza swoją bezpośrednio podłączoną sieć 10.10.10.0/24, jednak ponieważ RIP jest protokołem klasowym, sieć ta rozgłaszana jest bez maski podsieci. Z tego powodu router otrzymujący rozgłaszaną informację musi założyć, że w rozgłoszeniu zawarta jest cała sieć. W konsekwencji Router B otrzymuje od Routera A rozgłoszenie dla sieci 10.0.0.0 i wpisuje całą sieć 10.0.0.0/8 do swojej tabeli routingu. Do Routera C podłączona jest inna sieć 10: 10.20.20.0/24. Ponownie RIP rozgłasza sieć 10.0.0.0 z Routera C bez maski podsieci. Router B otrzymał w ten sposób kolejne rozgłoszenie dla 10.0.0.0/8. Sieć jest ta sama, protokół jest ten sam i liczba przeskoków jest ta sama. Kiedy otrzymywana jest nowsza aktualizacja dla trasy, która została już wpisana do tabeli routingu, to ta nowsza aktualizacja uznawana jest za bardziej wiarygodną i wprowadzana do tabeli routingu w miejsce poprzedniego wpisu. Oznacza to, że za każdym razem, kiedy Router B otrzymuje aktualizację od Routera A lub C, zmienia on swoją tabelę routingu, aby pokazać, że sieć 10.0.0.0 znajduje się za routerem, który wysłał daną aktualizację.



Rysunek 10.9. Problem projektu sieci wykorzystującej protokół klasowy RIP

Możesz się pokusić, aby stwierdzić, że sieci za Routerami A i C z rysunku 10.9 są różne. Jednak z perspektywy protokołu RIP byłbyś w błędzie. Technicznie rzecz biorąc, **sieci** (ang. *networks*) za Routerami A i C są takie same. Obie są częścią sieci 10.0.0.0/8. Routery są podłączone do różnych **podsieci** (ang. *subnets*) w obrębie tej samej sieci 10.0.0.0/8. Właśnie dlatego RIP ma problem związany z konstrukcją tej sieci.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jedynym rodzajem sieci, jaką RIP rozumie, jest sieć hosta. RIP może rozgłaszać trasę dla sieci /32 lub 255.255.255.255. Ponieważ RIP nie używa masek podsieci w swoich aktualizacjach, trasa jest rozpoznawana jako trasa hosta, wtedy gdy adres sieci jest inny niż adres normalnej sieci.

Protokoły routingu w IOS konfiguruje się poleceniem `router`. Nazwa protokołu routingu zawarta jest w poleceniu, co jednocześnie powoduje przejście w tryb konfiguracji routera (ang. *router configuration mode*):

```
Router-A (config)#router rip
Router-A (config-router)#
```

Jeśli chcesz zastosować RIPv1 w nowoczesnych routerach, które obsługują RIPv2, musisz to wyraźnie określić w konfiguracji routera. Domyślnie ustawiony jest RIPv2:

```
router rip
 version 1
```

W protokołach routingu domyślnie nie są zawarte żadne interfejsy. Oznacza to, że żadne aktualizacje nie będą do nich kierowane, a informacje otrzymane przez te interfejsy będą ignorowane.

Aby włączyć interfejsy w protokole routingu, musisz podać skonfigurowane sieci na interfejsach, które chcesz dołączyć. W tym celu użyj polecenia `network` w trybie konfiguracji routera:

```
Router-A (config)#router rip
Router-A (config-router)# network 10.10.10.0
```

W przypadku protokołu klasowego takiego jak RIP musisz być bardzo ostrożny. Przykładowo, jeśli będziesz chciał dołączyć sieć 10.0.0.0, dołączysz również każdy interfejs skonfigurowany z adresem IP 10.x.x.x, bez względu na maskę podsieci. RIP nie pozwala na załączanie maski podsieci lub maski odwrotnej w instrukcji `network`. Jeśli wpiszesz sieć inną niż 10.0.0.0, IOS i tak zmieni ten wpis w pełną sieć klasową.

Próba dokonania takiego wpisu skutkować będzie wyświetleniem poniższej informacji:

```
router rip
 network 10.0.0.0
```

Konfiguracja włączająca wszystkie interfejsy dla Routera A, tak jak pokazano na rysunku 10.10, wyglądać będzie następująco:

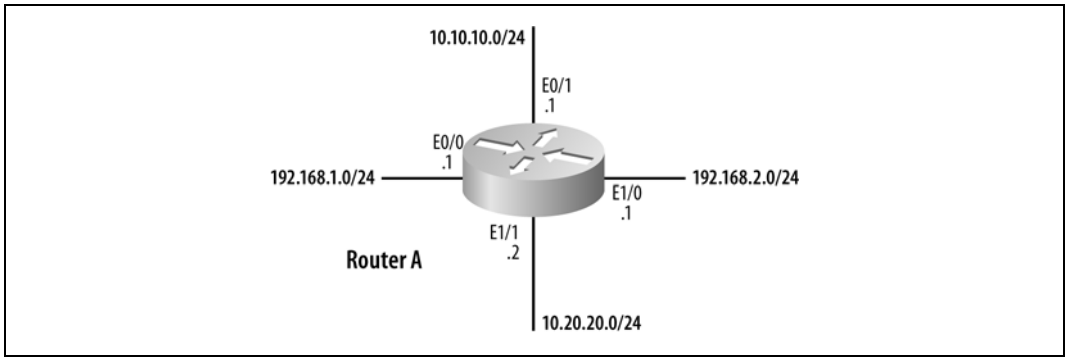
```
router rip
 version 1
 network 10.0.0.0
 network 192.168.1.0
 network 192.168.2.0
```

Sieci 10.10.10.0 oraz 10.20.20.0 konfiguruje się w jednym wierszu powyższego listingu. Jednak każda sieć 192.168 wymaga osobnej instrukcji `network`. Spowodowane jest to tym, że 192.x.x.x są sieciami klasy C, podczas gdy 10.x.x.x są sieciami klasy A.

Nie zawsze będziesz chciał dodać każdy interfejs, który jest zawarty w instrukcji `network`. Wracając do poprzedniego przykładu, możemy chcieć włączyć RIP na interfejsie E0/0, ale nie na interfejsie E1/1. Można to zrobić, stosując polecenie `passive-interface`, które usuwa interfejs z puli określonej komendą `network`:

```
router rip
 version 1
 passive-interface Ethernet1/1
 network 10.0.0.0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 10.10. Interfejsy sieciowe protokołu routingu

Polecenie `passive-interface` powoduje zatrzymanie wysyłania przez RIP aktualizacji na wybranym interfejsie. Router będzie jednak nadal otrzymywał i przetwarzał na tym interfejsie aktualizacje RIP.

Trasy uzyskane z protokołu RIP są identyfikowane w tabeli routingu przez dodanie litery `R` w pierwszej kolumnie. Poniższy przykład przedstawia sieć 172.16.0.0/16 uzyskaną z protokołu RIP. Aktualnie używaną siecią jest 172.16.100.0/24, ale ponieważ RIP jest protokołem klasowym, router przyjmie, że cała sieć 172.16.0.0/16 również się tu znajduje:

```
R3#sho ip route
[fragment usunięty]
Gateway of last resort is 192.168.1.2 to network 0.0.0.0
R    172.16.0.0/16 [120/1] via 10.10.10.4, 00:00:21, Ethernet0/0
     10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
C       10.10.10.0/24 is directly connected, Ethernet0/0
C       10.100.100.100/32 is directly connected, Loopback0
C    192.168.1.0/24 is directly connected, Ethernet1/0
S*   0.0.0.0/0 [254/0] via 192.168.1.2
```

W następnym przykładzie widzimy trasę hosta uzyskaną przez RIP:

```
R4#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     172.16.0.0/32 is subnetted, 1 subnets
C       172.16.1.1 is directly connected, Loopback0
     10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
C       10.10.10.0/24 is directly connected, Ethernet0/0
R       10.100.100.100/32 [120/1] via 10.10.10.3, 00:00:21, Ethernet0/0
C    192.168.1.0/24 is directly connected, Ethernet0/1
```

RIPv2

Protokół RIP został zaktualizowany w połowie lat 90. w celu uwzględnienia takich powszechnie stosowanych poprawek jak CIDR (ang. *Classless Internet Domain Routing*) oraz VLSM (ang. *Variable Length Subnet Masks*). Nowy protokół RIP w wersji 2 działa podobnie do RIP w wersji 1 pod tym względem, że nadal jedyną stosowaną metryką jest liczba przeskoków. Protokół ten posiada jednak kilka znaczących ulepszeń w stosunku do RIPv1:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

- RIPv2 obsługuje routing bezklasowy dzięki włączeniu masek podsieci do rozgłoszeń sieciowych.
- RIPv2 obsługuje aktualizacje wyzwalane.
- W RIPv2 aktualizacje wysyłane są z wykorzystaniem multicastowego adresu 224.0.0.9 zamiast broadcastów.
- Sąsiedzi mogą być skonfigurowani w RIPv2. W takim przypadku aktualizacje są wysyłane do sąsiada w postaci unicastów, co pozwala jeszcze bardziej zredukować natężenie ruchu w sieci.
- RIPv2 obsługuje uwierzytelnianie między routerami.

Mimo że RIPv2 obsługuje podsieci, nadal akceptuje jedynie adresy klasowe w poleceniu `network`. Bądź więc uważny przy określaniu, które sieci i interfejsy chcesz dodać. Skorzystaj z polecenia `passive-interface`, aby w razie konieczności ograniczyć zakres polecenia `network`.

RIPv2 jest protokołem bezklasowym i rozgłasza trasy z maskami podsieci, ale domyślnie sumuje trasy. Oznacza to, że jeśli trasa 10.10.10.0/24 jest podłączona do routera, to wciąż rozgłaszana będzie jako 10.0.0.0/8, tak jak w przypadku RIPv1. Pierwszą rzeczą, jaką powinieneś zrobić, konfigurując RIPv2, jest wyłączenie funkcji automatycznego sumowania (ang. *autosummarization*) poleceniem routera `no auto-summary`:

```
R3(config)#router rip
R3(config-router)# no auto-summary
```

W routerach Cisco tabela routingu nie dokonuje rozróżnienia pomiędzy RIPv1 oraz RIPv2. Oba protokoły identyfikowane są poprzez oznaczenie pojedynczą literą R w tabeli routingu. Routery Cisco stosują domyślnie RIPv1. Jeśli chcesz więc skorzystać z wersji drugiej tego protokołu, musisz ją najpierw włączyć. W przypadku korzystania z protokołu RIP prawdopodobnie będziesz chciał użyć właśnie wersji drugiej.

EIGRP

EIGRP jest protokołem bezklasowym i stanowi rozwinięcie protokołu IGRP obsługującego tylko sieci klasowe. EIGRP, podobnie jak IGRP, jest własnością firmy Cisco, co oznacza, że jedynie routery Cisco obsługują ten protokół. Jeśli zainstalujesz w sieci router Juniper czy Nortel, nie będzie mógł on komunikować się z routerami Cisco za pomocą EIGRP.

EIGRP jest bardzo popularnym protokołem routingu, ponieważ jest łatwy w konfiguracji i zarządzaniu. Przy minimalnej konfiguracji i niewielkich nakładach na projektowanie możesz uruchomić sprawnie działającą sieć EIGRP, która posłuży Twojej firmie przez długie lata.

Prostota konfiguracji EIGRP jest jednak również głównym powodem tego, że napotykam w praktyce wiele niewłaściwie działających sieci EIGRP. Załóżmy, że inżynier sieciowy buduje niewielką sieć dla swojej firmy. Z czasem sieć się rozrasta, a środowisko routingu robi się coraz bardziej skomplikowane. EIGRP radzi sobie z routingiem w sieci całkiem dobrze aż do czasu, gdy pewnego dnia coś zaczyna szwankować. Inżynier, który zaprojektował sieć, nie jest w stanie zorientować się, w czym tkwi problem. Wzywani są więc konsultanci, którzy całkowicie przebudowują całą sieć.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

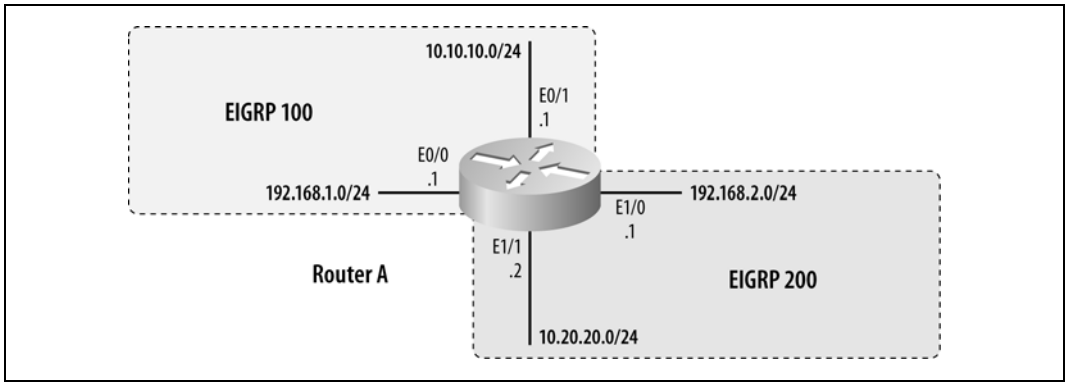
Nie znaczy to, że EIGRP jest złym protokołem routingu — uważam, że to bardzo silny protokół. Chodzi mi o to, że jest prawie zbyt łatwy w konfiguracji. Możesz wrzucić dwa routery EIGRP z minimalną konfiguracją w Ethernet LAN i będą się one ze sobą komunikować, wymieniając się informacjami o trasach. Możesz zrobić to samo z 10, 20 czy nawet 100 routerami, a nadal będą się one między sobą komunikować i wymieniać trasy. Nawet jeśli dodasz jeszcze 100 łączy szeregowych ze zdalnymi lokacjami, efekt będzie wciąż taki sam. Co prawda w tabeli routingu panować będzie spory bałagan, a routery będą nieustannie konwergować, niemniej pakiety będą płynąć. W końcu jednak domyślne ustawienia mogą przestać działać poprawnie, ponieważ konfiguracja domyślna nie jest przeznaczona dla rozległych sieci.

Jeśli jednak EIGRP jest poprawnie skonfigurowany, a sieć posiada dobrze zaprojektowany schemat adresacji IP, może on być doskonałym protokołem nawet w dużej sieci. W przypadku skonfigurowania wielu procesów EIGRP skaluje się bardzo dobrze.

EIGRP jest hybrydowym protokołem routingu łączącym w sobie cechy protokołów wektora odległości z cechami spotykanymi zazwyczaj w protokołach stanu łącza. EIGRP stosuje aktualizacje wyzwalane, są więc one wysyłane jedynie wtedy, gdy pojawiają się zmiany. Jako domyślne metryki wykorzystywane są przepustowość oraz opóźnienie i choć możesz dodać do tego równania jeszcze inne atrybuty, to najczęściej nie jest to dobry pomysł. EIGRP konwerguje bardzo szybko, nawet w dużych sieciach. Sieć, która konwergowałaby minutami w przypadku RIP, z EIGRP stanie się zbieżna w przeciągu kilku sekund.

Aby skonfigurować EIGRP, wejdź w tryb konfiguracji routera poleceniem `router eigrp` *`numer-systemu-autonomicznego`*. Numer systemu autonomicznego identyfikuje instancję EIGRP. Jeden router może posiadać wiele instancji EIGRP pracujących równolegle. Każda instancja ma własną bazę danych z trasami. Router będzie wybierał najlepszą trasę na podstawie takich kryteriów jak metryki, dystans administracyjny itd. Jest to zachowanie odmienne od tego, którym charakteryzuje się protokół RIP działający globalnie na całym routerze.

Rysunek 10.11 przedstawia router z dwoma aktywnymi instancjami EIGRP. Każda instancja oznaczana jest indywidualnym numerem systemu autonomicznego (ASN). Domyślnie informacje o trasach uzyskane w jednym procesie nie są dzielone z innymi procesami. Każdy proces to zasadniczo osobny protokół routingu. Aby informacje o trasach uzyskane przez jeden proces mogły być przekazywane innym procesom, router musi być skonfigurowany do redystrybucji. EIGRP będzie automatycznie redystrybuował trasy IGRP w obrębie tego samego ASN (redystrybucja opisana jest szczegółowo w rozdziale 11.).



Rysunek 10.11. Kilka instancji EIGRP

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jak w przypadku wszystkich protokołów IGP, możesz wyświetlić listę wszystkich interfejsów, które chciałbyś włączyć do protokołu, stosując polecenie `network`. EIGRP, podobnie jak RIP, automatycznie przekonwertuje bezklasową sieć na jej klasowy odpowiednik. Różnica polega na tym, że w EIGRP możesz dodać odwróconą maskę podsieci, aby wpis był bardziej szczegółowy. Poniższe komendy dodają do procesu EIGRP wszystkie interfejsy z adresami w sieci 10.0.0.0:

```
Router-A(config)#router eigrp 100
Router-A(config-router)# network 10.0.0.0
```

W przykładzie z rysunku 10.11 chcielibyśmy jednak dodać interfejs z siecią 10.10.10.0/24. Maska podsieci dla sieci /24 to 255.255.255.0, a odwrócona maska podsieci to 0.0.0.255. Odwrócone maski podsieci nazywane są również maskami wieloznacznymi (ang. *wildcard masks*) i zostały omówione w rozdziale 23. Aby dodać tylko ten jeden interfejs, użyjemy polecenia `network` w następujący sposób:

```
Router-A(config-router)#network 10.10.10.0 0.0.0.255
```

Po wykonaniu tego polecenia bieżąca konfiguracja będzie nadal zawierać nie całkiem precyzyjną informację o stanie sieci:

```
router eigrp 100
 network 10.10.10.0 0.0.0.255
 network 10.0.0.0
```

Obie komendy odniosą swój skutek. Musisz mieć to na uwadze i być ostrożny, gdyż może Cię to doprowadzić do stanu niekończącej się frustracji. W tym przykładzie interfejs E1/1 zostanie włączony do EIGRP 100, a nie to było naszym celem. Musimy usunąć nieprecyzyjny wpis, stosując zanegowanie komendy `network`:

```
router eigrp 100
 no network 10.0.0.0
```

Dobrą praktyką jest włączanie tylko tego określonego interfejsu, który chcesz dodać, w każdym z obsługujących go procesów routingu. Możesz to zrobić, podając na interfejsie adres IP z maską składającą się z samych zer. W naszym przykładzie będzie to komenda `network 10.10.10.1 0.0.0.0`. Zapobiegnie to niespodziankom związanym ze zmianą maski lub numeracji interfejsu. Proponowałbym więc następującą konfigurację EIGRP dla routera z rysunku 10.11:

```
router eigrp 100
 network 10.10.10.1 0.0.0.0
 network 192.168.1.1 0.0.0.0
!
router eigrp 200
 network 10.20.20.1 0.0.0.0
 network 192.168.2.1 0.0.0.0
```

EIGRP sumuje trasy, podobnie jak robi to RIP. Skoro jednak EIGRP jest protokołem bezklasowym, możemy wyłączyć tę funkcję poleceniem `no auto-summary`:

```
Router-A(config-router)#no auto-summary
```

W większości instancji lepiej jest wyłączyć tryb automatycznego sumowania, więc dobrze będzie, jeśli wejdzie Ci to w nawyk.

EIGRP działa poprzez wysyłanie pakietów hello przy wykorzystaniu multicastowego adresu IP 224.0.0.10 na skonfigurowanych interfejsach. Kiedy router, na którym pracuje EIGRP, otrzymuje pakiety hello, sprawdza informację o procesie zawartą w pakietach. Jeśli numer procesu

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

przesłany w pakiecie hello zgadza się z numerem procesu EIGRP działającego lokalnie, podejmowana jest próba nawiązania komunikacji. Kiedy komunikacja zostanie nawiązana, routery stają się **sąsiadami** (ang. *neighbors*).

W przeciwieństwie do RIP, który rozgłasza informacje o trasach do wszystkich urządzeń nasłuchujących, EIGRP wymienia trasy jedynie z sąsiadami. Kiedy między sąsiadami utworzona zostanie **relacja przylegania** (ang. *neighbor adjacency*), aktualizacje przesyłane są bezpośrednio do sąsiada z wykorzystaniem pakietów unicast.

Pożyteczną komendą w instalacjach EIGRP jest `eigrp log-neighbor-changes`. Komenda ta powoduje wyświetlenie komunikatu na konsoli/monitorze/w tabeli logów (w zależności od konfiguracji logowania) za każdym razem, kiedy stan relacji przylegania sąsiadów ulega zmianie:

```
1d11h: %DUAL-5-NBRCHANGE: IP-EIGRP 100: Neighbor 10.10.10.4 (Ethernet0/0) is up:
new adjacency // Sąsiad 10.10.10.4 (Ethernet0/0) jest włączony: nowa relacja przylegania
```

W większych sieciach może to być irytujące w trakcie rozwiązywania pojawiających się problemów, funkcję tę można jednak łatwo wyłączyć w razie potrzeby.

Aby sprawdzić status sąsiadów EIGRP na routerze, użyj polecenia `show ip eigrp neighbors`:

```
R3#sho ip eigrp neighbors
IP-EIGRP neighbors for process 100
H   Address                 Interface   Hold Uptime   SRTT   RTO  Q  Seq Type
                                        (sec)         (ms)       Cnt Num
1   10.10.10.5              Et0/0         14 00:00:19    4   200  0  1
0   10.10.10.4              Et0/0         13 00:02:35    8   200  0  3
```

Informacja uzyskana w ten sposób powinna być pierwszą rzeczą, jaką będziesz analizował w przypadku wystąpienia problemów z działaniem sieci, ponieważ bez relacji przylegania sąsiadów routery EIGRP nie będą wymieniać informacji o trasach.

Trasy uzyskane od wewnętrznego EIGRP mają dystans administracyjny 90 i są oznaczone pojedynczą literą `D` w pierwszej kolumnie tabeli routingu. Natomiast trasy uzyskane od zewnętrznego EIGRP mają dystans administracyjny 170 i są oznaczone literami `D EX` umieszczonymi na początku wpisu informacji o trasie:

```
R3#sho ip route
[fragment usunięty]
Gateway of last resort is 192.168.1.2 to network 0.0.0.0
     5.0.0.0/32 is subnetted, 1 subnets
D EX    5.5.5.5 [170/409600] via 10.10.10.5, 00:00:03, Ethernet0/0
     10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
C       10.10.10.0/24 is directly connected, Ethernet0/0
C       10.100.100.100/32 is directly connected, Loopback0
C    192.168.1.0/24 is directly connected, Ethernet1/0
D    192.168.3.0/24 [90/2195456] via 10.10.10.5, 00:08:42, Ethernet0/0
S*   0.0.0.0/0 [254/0] via 192.168.1.2
```

EIGRP przechowuje informacje w trzech bazach danych: bazie danych tras, bazie danych topologii oraz w bazie danych sąsiadów. Analiza bazy danych topologii może być ogromną pomocą w rozwiązywaniu problemów związanych z routingiem. Znajdziesz tam nie tylko informacje o tym, co EIGRP umieścił w tabeli routingu, ale również o tym, jakie są alternatywne możliwości wytyczenia tras:

```
R3#sho ip eigrp topology
IP-EIGRP Topology Table for AS(100)/ID(10.100.100.100)
Codes: P - Passive, A - Active, U - Update, Q - Query, R - Reply,
       r - reply Status, s - sia Status
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
P 5.5.5.5/32, 1 successors, FD is 409600
        via 10.10.10.5 (409600/128256), Ethernet0/0
P 10.10.10.0/24, 1 successors, FD is 281600
        via Connected, Ethernet0/0
P 192.168.3.0/24, 1 successors, FD is 2195456
        via 10.10.10.5 (2195456/2169856), Ethernet0/0
```

OSPF

W skrócie założeniem protokołu routingu OSPF jest stosowanie najkrótszej lub najszybszej trasy, jaka jest dostępna.

OSPF wybiera się w następujących przypadkach:

- W sieci pracują routery innych producentów niż Cisco.
- Sieć wymaga segmentacji na obszary lub strefy.
- Chcesz uniknąć stosowania protokołów routingu będących czyjąś własnością.

OSPF jest protokołem stanu łącza, a jako metrykę stosuje przepustowość łączy. Oblicza się ją, dzieląc 100 000 000 przez przepustowość danego linku podaną w bitach na sekundę. Tak więc link 100 Mb/s będzie miał metrykę lub „koszt" 1, link 10 Mb/s koszt 10, a link 1,5 Mb/s koszt 64. Łącze o prędkości 1 Gb/s również będzie miało koszt 1, gdyż wartość ta nie może być mniejsza niż 1. Koszty każdego linku ścieżki są sumowane w celu uzyskania metryki trasy.

W sieciach, w których występują łącza szybsze niż 100 Mb/s, formuła obliczania kosztów może być zmieniona poleceniem `auto-cost reference-bandwidth`. Domyślna przepustowość referencyjna wynosi 100. Innymi słowy, łącze o prędkości 100 Mb/s posiada domyślnie koszt 1. Aby ustawić koszt 1 dla łącza 1000 Mb/s, zmień referencyjną wartość przepustowości na 1000:

```
R3(config)#router ospf 100
R3(config-router)# auto-cost reference-bandwidth 1000
```

Jeśli chcesz zmienić wartość referencyjnej przepustowości, musisz zrobić to na każdym routerze komunikującym się w procesie OSPF. Jeśli tego nie zrobisz, sieć może zacząć działać niestabilnie, a routing może przebiegać nieprzewidywalnie.

OSPF klasyfikuje routery ze względu na funkcję, jaką pełnią w sieci. Oto typy routerów OSPF:

Router wewnętrzny (ang. *internal router*)
Router wewnętrzny znajduje się całkowicie w obrębie jednego obszaru OSPF w danym systemie autonomicznym.

Router granicy obszaru (ang. *area border router — ABR*)
Router ABR znajduje się w obrębie więcej niż jednego obszaru pojedynczego systemu autonomicznego OSPF.

Router granicy systemu autonomicznego (ang. *autonomous system border router — ASBR*)
ASBR połączony jest z wieloma systemami autonomicznymi OSPF lub z systemem autonomicznym OSPF oraz systemem autonomicznym innego protokołu routingu.

Routery szkieletowe (ang. *backbone routers*)
Routery szkieletowe znajdują się w obszarze zerowym (ang. *area zero*). Obszar zerowy traktowany jest jako szkielet w sieci OSPF.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Router wyznaczony (ang. *designated routers — DR*)
DR w sieciach rozgłoszeniowych to router wybierany do ponoszenia całego ciężaru przetwarzania OSPF. DR aktualizuje pozostałe routery w obszarze tras.

Zapasowy router wyznaczony (ang. *backup designated router — BDR*)
Jeśli następuje awaria DR, router BDR przejmuje jego funkcje.

W przeciwieństwie do innych protokołów routingu OSPF nie wysyła informacji o trasach, a raczej **ogłoszenia stanu łącza** (ang. *link state advertisements — LSA*). Każdy router OSPF określa, którą trasę wybrać, na podstawie wewnętrznej bazy danych tworzonej przez ogłoszenia LSA. Jest sześć typów LSA:

Typ 1, LSA routera (ang. *router LSA*)
LSA routera są wysyłane przez każdy router OSPF do każdego podłączonego obszaru. Ogłoszenia te opisują połączenia routera w obrębie obszaru.

Typ 2, LSA sieci (ang. *network LSA*)
LSA sieci są wysyłane przez routery DR i opisują routery podłączone do sieci, z której LSA zostało wysłane.

Typ 3, łączne LSA dla routerów ABR (ang. *summary LSA for ABR*)
Są one wysyłane przez routery ABR. Ogłoszenia te opisują trasy między obszarami w danej sieci. Są też stosowane do ogłaszania tras skonsolidowanych.

Typ 4, łączne LSA dla routerów ASBR (ang. *summary LSA for ASBR*)
Ogłoszenia te są wysyłane przez routery ASBR i ABR. Opisują linki do routerów ASBR.

Typ 5, zewnętrzne LSA systemu autonomicznego (ang. *autonomous system external (ASE) LSA*)
ASE LSA są wysyłane przez routery ASBR i ABR. Opisują sieci zewnętrzne względem systemu autonomicznego. Ogłoszenia te wysyłane są do wszystkich obszarów poza obszarami pnia (ang. *stub areas*).

Typ 7, LSA Not So Stubby Area (NSSA)
NSSA LSA są wysyłane przez routery ASBR obszaru NSSA. Ponieważ typ 5 LSA nie jest dopuszczalny w obszarze NSSA, wysyłany jest w zamian typ 7 LSA, który pozostaje w obszarze NSSA. LSA typu 7 są konwertowane na typ 5 przez routery ASR obszaru NSSA dla celów ogłaszania w szkielecie.

OSPF dzieli sieci na obszary. Obszarem rdzeniowym, z którym muszą łączyć się wszystkie inne obszary, jest obszar zwany zerowym. Jedną z dających się zauważyć zalet protokołu OSPF jest konieczność zaprojektowania sieci w taki sposób, aby posiadała ona rdzeń i obszary go otaczające. Możesz oczywiście zbudować sieć OSPF posiadającą jedynie obszar zerowy, ale taki projekt zazwyczaj nie skaluje się dobrze.

Są dwa główne rodzaje obszarów: **szkieletowe** (ang. *backbone*) i **nieszkieletowe** (ang. *nonbackbone*). Obszar zerowy jest obszarem szkieletowym. Pozostałe obszary zaliczają się do nieszkieletowych. Obszary nieszkieletowe dzielą się na poniższe typy:

Obszar normalny (ang. *normal area*)
Jest to obszar, który nie jest obszarem zerowym i nie jest skonfigurowany jak żaden z poniższych obszarów. Obszar normalny nie wymaga żadnej konkretnej konfiguracji.

Obszar pnia (ang. *stub area*)
Obszar OSPF, który nie dopuszcza ogłoszeń ASE LSA. Kiedy obszar jest skonfigurowany jako stub area, żadne trasy O E1 lub O E2 nie będą widoczne w tym obszarze.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Totally stubby area (TSA)
Obszar OSPF, który nie dopuszcza LSA typu 3, 4 lub 5, z wyjątkiem domyślnej trasy skonsolidowanej. TSA widzi jedynie trasę domyślną i trasy lokalne dla samych obszarów.

Not so stubby area (NSSA)
W obszarze NSSA nie są dopuszczalne ogłoszenia LSA typu 5. Dopuszczalny jest typ 7 LSA, który został skonwertowany na typ 5 przez routery ABR.

NSSA totally stub area
Obszar NSSA totally stub area jest kombinacją obszarów tottaly stubby oraz not so stubby. Obszar ten nie dopuszcza LSA typu 3, 4 lub 5, z wyjątkiem domyślnej trasy skonsolidowanej. Akceptowany jest typ 7 LSA, który został skonwertowany na typ 5 przez routery ABR.

W przypadku Ethernetu oraz innych sieci rozgłoszeniowych OSPF wybiera router, który zostaje routerem wyznaczonym, oraz kolejny router jako zapasowy router wyznaczony. Wyznaczanie tras w protokole OSPF może wymagać dużej mocy obliczeniowej, zwłaszcza w dynamicznej sieci. Jeśli główny ciężar obliczeniowy spoczywa na jednym routerze, sieć działa stabilniej i szybciej konwerguje. DR wyznacza najlepsze trasy, a następnie propaguje te informacje do swoich sąsiadów w sieci, którzy znajdują się w tym samym obszarze i procesie OSPF.

DR jest wybierany w OSPF dynamicznie i jest to relatywnie skomplikowany proces. Pierwszy krok związany jest z **priorytetem** (ang. *priority*) interfejsu OSPF routera. Domyślną wartością priorytetu jest 1. Jest to najniższa wartość, jaką może posiadać interfejs wybierany na DR. Wartość 0 wskazuje, że router nie może zostać wybrany routerem wyznaczonym w danej sieci. Im wyższa wartość priorytetu, tym większa szansa, że router zostanie wybrany. Priorytet interfejsu OSPF może być konfigurowany poleceniem interfejsu `ip ospf priority`. Dopuszczalne są wartości z zakresu 0 – 255.

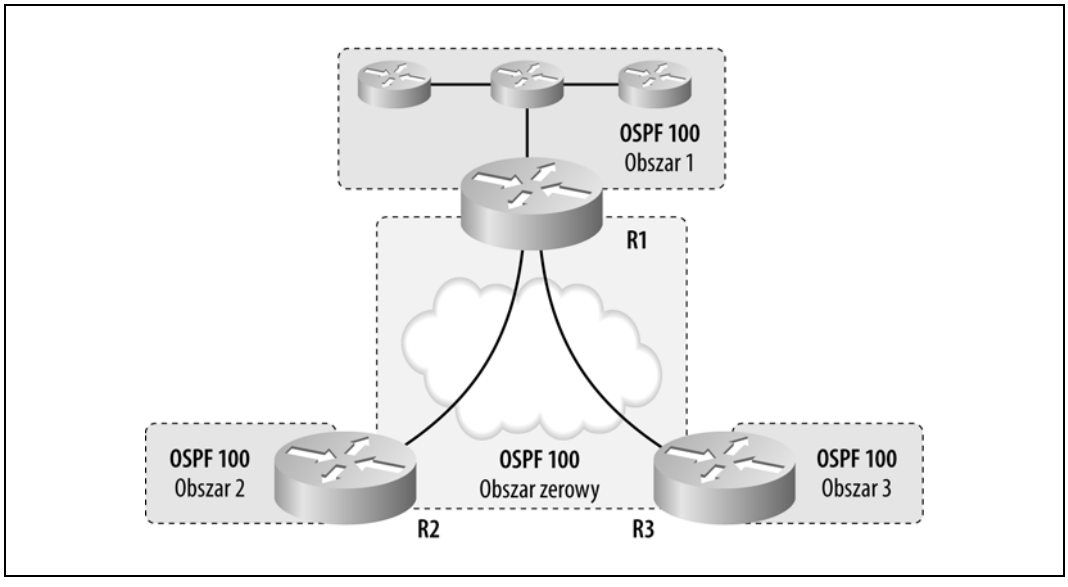
Najlepiej będzie, jeśli zaplanujesz, który router ma zostać DR, i ustawisz odpowiednio jego priorytet. Zazwyczaj oczywistym wyborem jest koncentrator, router rdzenia czy też po prostu najbardziej wydajny router w sieci. Router wyznaczony będzie wykonywał dużo większą pracę niż pozostałe routery w sieci, więc powinien to być router o największej mocy obliczeniowej. Jeśli w projekcie Twojej sieci znajduje się router zintegrowany z portami przełącznika, to właśnie on powinien zostać DR, gdyż będzie centralnym punktem topologii.

Jeśli priorytet interfejsu OSPF nie jest skonfigurowany, tworzy się sytuacja patowa. Aby z niej wybrnąć, router wykorzystuje parametr ID routera OSPF (ang. *OSPF router ID*). Każdy router posiada swój identyfikator OSPF. ID może być konfigurowane ręcznie poleceniem `router-id`. Jeśli ID nie zostało skonfigurowane, router przydzieli jego wartość, opierając się na adresie IP najwyższego interfejsu pętli zwrotnej (ang. *loopback*), jeśli ten jest skonfigurowany. Jeśli adres interfejsu pętli zwrotnej nie został skonfigurowany, identyfikatorem routera będzie najwyższy adres IP skonfigurowany na routerze. Jedynym sposobem na zmianę ID routera jest usunięcie i ponowne zainstalowanie konfiguracji protokołu OSPF lub ponowne uruchomienie routera. Bądź więc ostrożny i przewiduj z wyprzedzeniem przy planowaniu schematu IP swojej sieci.

Przy pierwszym wdrażaniu OSPF inżynierowie popełniają częsty błąd, zaniedbując podczas konfiguracji routerów priorytet oraz ID. Jeśli pozostawisz urządzenia samym sobie, OSPF prawdopodobnie dokona wyboru routerów DR i BDR niezgodnego z Twoim zamysłem.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Powszechnie przy projektowaniu sieci wykorzystującej protokół OSPF WAN umieszcza się w rdzeniu jako obszarze zerowym. Rysunek 10.12 przedstawia taką sieć. Zauważ, że wszystkie obszary oznaczone są tym samym numerem procesu OSPF. Każdy z obszarów graniczy z obszarem zerowym. Nie ma też żadnych ścieżek między obszarami poza tymi prowadzącymi przez obszar zerowy. Jest to właściwy schemat sieci OSPF.



Rysunek 10.12. Prosta sieć OSPF

Obszar zerowy nie musi być w pełni objęty siatką połączeń, jeśli stosujesz takie technologie jak Frame Relay. Jest tak po części dlatego, że OSPF zakłada istnienie różnych rodzajów sieci. OSPF wykrywa, że sieci obsługujące broadcasty działają odmiennie od sieci punkt-punkt, które posiadają tylko dwa aktywne adresy IP. Protokół OSPF wspiera następujące rodzaje sieci:

Punkt-punkt (ang. *point-to-point*)
Sieć typu punkt-punkt posiada jedynie dwa węzły. Typowym przykładem jest łącze szeregowe pomiędzy routerami, takie jak punkt-punkt T1. Ponieważ w sieciach punkt-punkt działają tylko dwa routery, nie ma w nich routera wyznaczonego. Jest to domyślny rodzaj sieci OSPF na interfejsach szeregowych z enkapsulacją PPP lub HDLC.

Punkt-wielopunkt (ang. *point-to-multipoint*)
W sieciach punkt-wielopunkt jeden router z przełącznikiem łączy się ze wszystkimi pozostałymi routerami. Każdy z pozostałych routerów łączy się tylko z tym routerem głównym. Zdalne routery są traktowane jak połączone obwodami wirtualnymi, chociaż wykorzystana jest tylko jedna sieć IP. Nie ma skonfigurowanych żadnych sąsiadów i nie wybiera się też routera DR. Obszar zerowy z rysunku 10.12 mógłby być skonfigurowany jako sieć OSPF punkt-wielopunkt.

Sieć rozgłoszeniowa (ang. *broadcast network*)
Siecią rozgłoszeniową jest Ethernet, Token Ring czy też FDDI. W tego typu sieci może występować dowolna liczba hostów, a każdy host może komunikować się bezpośrednio z każdym innym hostem. W sieci rozgłoszeniowej sąsiedzi muszą zostać skonfigurowani

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

lub automatycznie wykryci. Musi być też wybrany router DR. Sieć wykorzystuje multicasty do wysyłania pakietów typu hello pozwalających wykryć wszystkie routery OSPF. Jest to domyślny rodzaj sieci OSPF w przypadku Ethernetu i Token Ring.

Sieć nierozgłoszeniowa wielodostępowa (ang. *nonbroadcast multiaccess — NBMA*)
W nierozgłoszeniowych sieciach wielodostępowych wszystkie węzły mogą komunikować się ze sobą, ale nie współdzielą one jednego medium transmisyjnego. Przykładami takich sieci są Frame Relay, X.25 oraz SMDS (ang. *Switched Multimegabit Data Service*). Ponieważ sieci NBMA nie używają multicastów do wykrywania sąsiadów, musisz konfigurować ich ręcznie. Obszar zerowy pokazany na rysunku 10.12 mógłby być skonfigurowany jako sieć NBMA. Jest to domyślny rodzaj sieci na interfejsach szeregowych z enkapsulacją Frame Relay.

OSPF umożliwia włączanie interfejsów za pomocą komendy routera `network`. Jest to protokół bezklasowy, musisz więc zastosować odwrócone maski podsieci, aby ograniczyć dołączone interfejsy. W przeciwieństwie do EIGRP użycie maski odwróconej jest konieczne. Jeśli tego nie zrobisz, OSPF nie będzie zakładał, że jest to sieć klasowa, i wyświetli informację o błędzie:

```
R3(config-router)#network 10.10.10.0
% Incomplete command. // Komenda niekompletna
```

Oprócz odwróconej maski musisz określić także obszar, w którym rezyduje dana sieć.

```
R3(config-router)#network 10.10.10.0 0.0.0.255 area 0
```

Osobiście preferuję szczegółową konfigurację interfejsów, aby nie było żadnych niespodzianek. Robi się to za pomocą odwróconej maski 0.0.0.0:

```
R3(config-router)#network 10.10.10.1 0.0.0.0 area 0
```

Trasy OSPF oznaczone są literą `O` w pierwszej kolumnie tabeli routingu:

```
R3#sho ip route
[fragment usunięty]
Gateway of last resort is 192.168.1.2 to network 0.0.0.0
     192.192.192.0/30 is subnetted, 1 subnets
C       192.192.192.4 is directly connected, Serial0/0
     172.16.0.0/32 is subnetted, 1 subnets
O IA    172.16.1.1 [110/11] via 10.10.10.4, 00:00:09, Ethernet0/0
     10.0.0.0/24 is subnetted, 1 subnets
C       10.10.10.0 is directly connected, Ethernet0/0
C    192.168.1.0/24 is directly connected, Ethernet1/0
S*   0.0.0.0/0 [254/0] via 192.168.1.2
```

Różne rodzaje tras OSPF oznaczone są w tabeli routingu. Są tu więc: `O` (OSPF), `O IA` (OSPF interarea), `O N1` (OSPF NSSA external typ 1) oraz `O N2` (OSPF NSSA typ 2).

OSPF przechowuje informacje o swoich trasach w bazie danych podobnie jak EIGRP. Aby wyświetlić bazę danych, należy użyć polecenia `show ip ospf database`:

```
R3#sho ip ospf database
            OSPF Router with ID (192.192.192.5) (Process ID 100)
                Router Link States (Area 0)
Link ID         ADV Router      Age         Seq#       Checksum Link count
192.192.192.5   192.192.192.5   1769        0x8000002A 0x00C190 1
                Summary Net Link States (Area 0)
Link ID         ADV Router      Age         Seq#       Checksum
192.192.192.4   192.192.192.5   1769        0x8000002A 0x003415
                Router Link States (Area 1)
Link ID         ADV Router      Age         Seq#       Checksum Link count
192.192.192.5   192.192.192.5   1769        0x8000002A 0x00B046 1
                Summary Net Link States (Area 1)
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Link ID         ADV Router      Age         Seq#       Checksum
10.10.10.0      192.192.192.5   1769        0x8000002A 0x0002A2
   OSPF Router with ID (192.168.1.116) (Process ID 1)
```

Jeśli to wszystko wydaje Ci się niepotrzebnie skomplikowane, nie jesteś osamotniony. Złożoność protokołu OSPF jest jednym z powodów, dla którego większość osób wybiera EIGRP. Jeśli jednak pracujesz w środowisku, w którym znajdują się urządzenia różnych producentów, to EIGRP nie wchodzi w grę.

BGP

BGP znacząco różni się od pozostałych protokołów opisanych w tym rozdziale. Najbardziej oczywistą różnicą jest to, że BGP jest protokołem bramy zewnętrznej, podczas gdy wszystkie omawiane do tej pory protokoły były protokołami bramy wewnętrznej. BGP może być więc trudny do zrozumienia dla osób, które miały dotychczas do czynienia z protokołami wewnętrznymi, jak EIGRP czy OSPF, gdyż sama natura tych protokołów jest zupełnie odmienna. Ponieważ BGP jest rzadko spotykany w środowisku korporacyjnym, omówię go tutaj w skrócie.

W przypadku BGP nie masz do czynienia z przeskokami czy linkami, ale raczej z systemami autonomicznymi. Sieć w protokole BGP określana jest jako **prefiks** (ang. *prefix*). Prefiks jest ogłaszany z systemu autonomicznego. Następnie BGP propaguje tę informację na podłączone systemy autonomiczne, dopóki nie dotrze ona do wszystkich systemów.

Trasy w protokole BGP uznawane są za najbardziej odpowiednie, kiedy przechodzą przez możliwie najmniejszą liczbę systemów autonomicznych. Kiedy prefiks jest ogłaszany, numer systemu autonomicznego zostaje dołączony w formie prefiksu do informacji o **ścieżce** (ang. *path*) systemu autonomicznego. Ścieżka ta jest ekwiwalentem trasy w protokołach bramy wewnętrznej. Jeśli system autonomiczny uzyska informację o prefiksie, zdobędzie również informację o powiązanej z nim ścieżce. Gdy system autonomiczny rozgłasza ten prefiks do kolejnego systemu autonomicznego, ten również wstawia własny ASN do ścieżki. W trakcie rozgłaszania prefiksu do coraz większej liczby systemów autonomicznych ścieżka się wydłuża. Im krótsza ścieżka, tym lepsza.

Rysunek 10.13 przedstawia prosty przykład routingu BGP w akcji. Sieć 10.0.0.0/8 rezyduje w AS 105, który rozgłasza ten prefiks do AS 3 i AS 2. Ścieżka dla sieci 10.0.0.0/8 w AS 3 i AS 2 jest teraz opisana jako 10.0.0.0/8 AS105. AS 2 ogłasza z kolei prefiks do AS 1, dołączając własny ASN do ścieżki. W ten sposób AS 1 poznaje ścieżkę do 10.0.0.0/8 jako 10.0.0.0/8 AS2, AS105. W międzyczasie AS 3 rozgłasza prefiks do AS 100, który teraz zna ścieżkę do 10.0.0.0/8 jako 10.0.0.0/8 AS3, AS105.

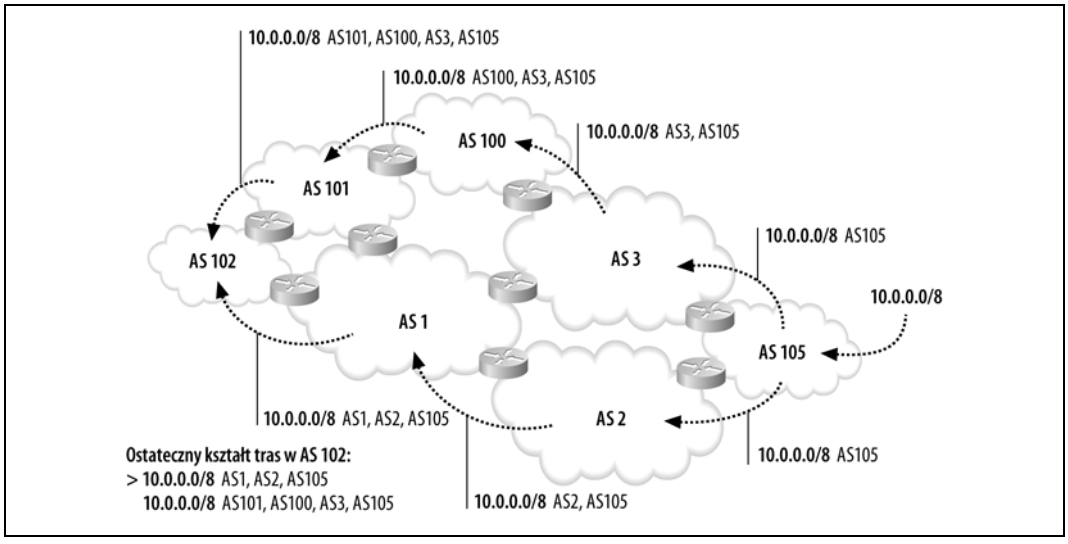
Na drugim końcu świata AS 102 otrzymuje dwie ścieżki:

```
> 10.0.0.0/8 AS1, AS2, AS105
  10.0.0.0/8 AS101, AS100, AS3, AS105
```

Oznaczenie > w pierwszej linii listingu wskazuje, że BGP uważa tę ścieżkę za preferowaną. Jest ona wyróżniona, gdyż jest najkrótszą ścieżką ze wszystkich znanych możliwości.

Tym, co czyni BGP tak trudnym do opanowania dla początkujących, jest mnogość parametrów, które można konfigurować. Ścieżka może posiadać wiele atrybutów, noszących takie nazwy jak: *local preference*, *weight*, *communities* czy *multiexit discriminator*. Aby jeszcze bardziej pogorszyć sprawę, wiele z tych atrybutów pełni bardzo podobną funkcję.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 10.13. Routing w protokole BGP

Protokół ten funkcjonuje ponadto inaczej niż pozostałe protokoły. Przykładowo instrukcja `network` stosowana w innych protokołach do włączania interfejsów w BGP służy do wyświetlenia listy tych sieci, które mogą być rozgłaszane.

BGP nie wykrywa sąsiadów — trzeba ich skonfigurować ręcznie. Na każdym z routerów może być tylko jeden system autonomiczny, który może jednak komunikować się z sąsiadami w innych systemach autonomicznych.

BGP jest protokołem routingu, na podstawie którego działa internet. Wielu głównych dostawców usług umożliwia anonimowe połączenie telnet z **serwerami tras** (ang. *route servers*), które działają jak routery Cisco. Poszukaj w internecie określenia „looking-glass routers", a z pewnością znajdziesz mnóstwo linków. Serwery tras są doskonałym sposobem, aby dowiedzieć się czegoś więcej o BGP, gdyż są częścią największej sieci na świecie i posiadają aktywne trasy do praktycznie każdej sieci publicznej na Ziemi. Jeśli nie jesteś zatrudniony u usługodawcy tier-1, to gdzie indziej znalazłbyś dostęp do routera BGP, który posiada 20 sąsiadów, 191 898 prefiksów oraz 3 666 117 ścieżek? Sam mam całkiem fajne laboratorium, ale z tym nie mogę konkurować! Oto listing z aktualnego serwera tras:

```
route-server>sho ip bgp summary
BGP router identifier 10.1.2.5, local AS number 65000
BGP table version is 208750, main routing table version 208750
191680 network entries using 19359680 bytes of memory
3641563 path entries using 174795024 bytes of memory
46514 BGP path attribute entries using 2605064 bytes of memory
42009 BGP AS-PATH entries using 1095100 bytes of memory
4 BGP community entries using 96 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 197854964 total bytes of memory
Dampening enabled. 2687 history paths, 420 dampened paths
191529 received paths for inbound soft reconfiguration
BGP activity 191898/218 prefixes, 3666117/24554 paths, scan interval 60 secs
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.0.0.2        4  7018       0       0        0    0    0 never    Idle (Admin)
12.0.1.63       4  7018   45038     188   208637    0    0 03:04:16        0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
12.123.1.236     4  7018   39405     189   208637    0    0 03:05:02   191504
12.123.5.240     4  7018   39735     189   208637    0    0 03:05:04   191504
12.123.9.241     4  7018   39343     189   208637    0    0 03:05:03   191528
12.123.13.241    4  7018   39617     188   208637    0    0 03:04:20   191529
12.123.17.244    4  7018   39747     188   208637    0    0 03:04:58   191505
12.123.21.243    4  7018   39441     188   208637    0    0 03:04:28   191528
12.123.25.245    4  7018   39789     189   208637    0    0 03:05:07   191504
12.123.29.249    4  7018   39602     188   208637    0    0 03:04:16   191505
12.123.33.249    4  7018   39541     188   208637    0    0 03:04:16   191528
12.123.37.250    4  7018   39699     188   208637    0    0 03:04:26   191529
12.123.41.250    4  7018   39463     188   208637    0    0 03:04:19   191529
12.123.45.252    4  7018   39386     188   208637    0    0 03:04:20   191505
12.123.133.124   4  7018   39720     188   208637    0    0 03:04:20   191528
12.123.134.124   4  7018   39729     188   208637    0    0 03:04:22   191529
12.123.137.124   4  7018   39480     188   208637    0    0 03:04:15   191528
12.123.139.124   4  7018   39807     188   208637    0    0 03:04:24   191528
12.123.142.124   4  7018   39748     188   208637    0    0 03:04:22   191505
12.123.145.124   4  7018   39655     188   208637    0    0 03:04:23   191529
```

Serwery tras mogą być często bardzo obciążone i dość powolne. Jeśli okaże się, że będziesz czekał dość długo na odpowiedź z serwera, spróbuj ponowić zapytanie lub poszukaj innego serwera tras.

Wybierz swoją ulubioną publiczną sieć IP (czyż nie każdy jakąś ma?) i sprawdź na „looking-glass router", jak wyglądają ścieżki. Jeśli nie masz żadnej ulubionej sieci, wybierz taką, o której możesz łatwo znaleźć informacje. Np. używaną przez *www.cisco.com* lub *www.oreilly.com*:

```
[bossman@myserver bossman]$nslookup www.oreilly.com
Server:  localhost
Address:  127.0.0.1
Name:    www.oreilly.com
Addresses:  208.201.239.36, 208.201.239.37
```

Kiedy znasz już adres, możesz przystąpić do przeszukiwania sieci:

```
route-server>sho ip bgp 208.201.239.0
BGP routing table entry for 208.201.224.0/19, version 157337
Paths: (19 available, best #15, table Default-IP-Routing-Table)
  Not advertised to any peer
  7018 701 7065, (received & used)
    12.123.137.124 from 12.123.137.124 (12.123.137.124)
      Origin IGP, localpref 100, valid, external, atomic-aggregate
      Community: 7018:5000
  7018 701 7065, (received & used)
    12.123.33.249 from 12.123.33.249 (12.123.33.249)
      Origin IGP, localpref 100, valid, external, atomic-aggregate
      Community: 7018:5000
  7018 701 7065, (received & used)
    12.123.29.249 from 12.123.29.249 (12.123.29.249)
      Origin IGP, localpref 100, valid, external, atomic-aggregate
      Community: 7018:5000
  7018 701 7065, (received & used)
    12.123.41.250 from 12.123.41.250 (12.123.41.250)
      Origin IGP, localpref 100, valid, external, atomic-aggregate
      Community: 7018:5000
  7018 701 7065, (received & used)
    12.123.1.236 from 12.123.1.236 (12.123.1.236)
      Origin IGP, localpref 100, valid, external, atomic-aggregate, best
      Community: 7018:5000
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

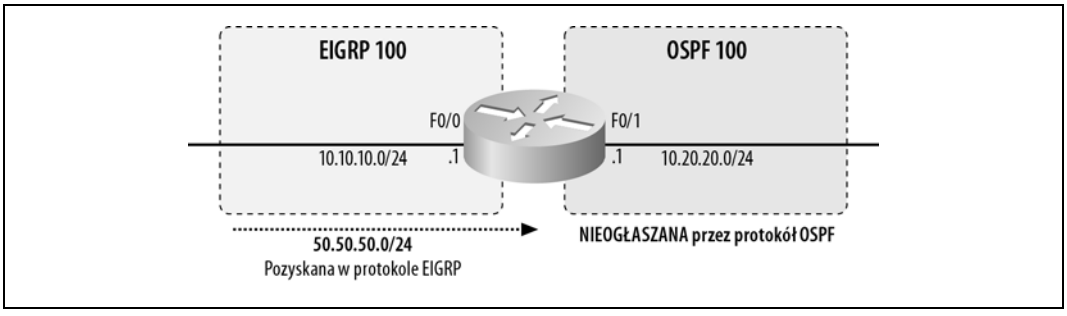
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 11.

Redystrybucja

Redystrybucja (ang. *redistribution*) jest procesem umieszczania w protokole routingu tras pochodzących spoza sfery tego protokołu. Załóżmy, że posiadasz router, na którym uruchomione są protokoły EIGRP oraz OSPF. Jeśli potrzebujesz, aby trasy uzyskane przez protokół EIGRP były ogłaszane w OSPF, musisz redystrybuować trasy EIGRP do OSPF. Innym przykładem jest redystrybucja tras statycznych oraz bezpośrednio podłączonych. Ponieważ trasy statyczne konfigurowane są ręcznie, a nie pozyskiwane dynamicznie, muszą być redystrybuowane w protokole routingu, jeśli chcesz, aby były ogłaszane.

Jak pokazuje rysunek 11.1, trasy uzyskane przez protokół EIGRP nie są automatycznie rozgłaszane na interfejsy OSPF. Aby umożliwić tłumaczenie informacji o trasach z jednego protokołu na drugi, musisz włączyć redystrybucję w obrębie tego protokołu, w którym chcesz, żeby pojawiły się te trasy.



Rysunek 11.1. Większość protokołów routingu domyślnie nie redystrybuuje tras

Jednym z głównych powodów, dla których protokoły nie redystrybuują tras automatycznie, jest to, że każdy z protokołów stosuje odmienne metryki. Przykładowo OSPF wytycza najlepsze trasy, opierając się na przepustowości łączy. Z kolei EIGRP stosuje domyślnie przepustowość oraz opóźnienie, tworząc zupełnie inną metrykę. Router mógłby z założenia uruchamiać redystrybucję i przypisywać standardowe metryki uzyskanym trasom. Jednak lepszym podejściem jest pozostawić administratorowi sieci decyzję, czy i w jaki sposób trasy mają być redystrybuowane.

W celu uruchomienia redystrybucji tras należy wykonać dwie czynności. Najpierw musisz skonfigurować metrykę. Umożliwi to protokołowi routingu przypisywanie uzyskiwanym trasom takich metryk, jakie dany protokół rozumie. Drugą czynnością jest wykonanie komendy

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

`redistribute`. Dokładna forma tej komendy różnić się będzie w zależności od protokołu i omówię ją dla każdego protokołu z osobna w kolejnych podrozdziałach.

Jednym z powodów redystrybucji tras może być istnienie firewalla, który musi uczestniczyć w procesie dynamicznego routingu, jednak nie może skorzystać z protokołu działającego aktualnie w sieci. Przykładowo wiele firewalli obsługuje RIP, ale nie obsługuje EIGRP. Aby umożliwić dynamiczny routing pomiędzy routerem EIGRP a firewallem RIP, musisz na routerze redystrybuować trasy między RIP i EIGRP.

Najlepszą zasadą przy wdrażaniu redystrybucji jest dążenie do prostoty. Łatwo się zagubić, kiedy trasy przesyłane są tam i z powrotem między protokołami routingu. Zachowanie przejrzystości projektu z pewnością pomoże w późniejszym zarządzaniu siecią. Jeśli redystrybucja nie będzie działała poprawnie, może to doprowadzić do powstawania całkiem interesujących komplikacji. Im prostszy projekt, tym łatwiej diagnozować problemy.

Redystrybucja polega na konwersji tras jednego protokołu na formę zrozumiałą dla innego protokołu. Osiąga się to przez przypisywanie nowych metryk trasom przechodzącym do nowego protokołu. Ponieważ trasy muszą zostać przypisane do metryk nowego protokołu, kluczem do zrozumienia redystrybucji jest zrozumienie, czym są metryki.

Jeśli protokół redystrybuuje trasy pochodzące z jakiegoś źródła, stają się one trasami **zewnętrznymi** (ang. *external*) w tym nowym protokole. Trasy mogą być redystrybuowane z ograniczonej liczby źródeł:

Trasy statyczne (ang. *static routes*)
: Trasy wprowadzone ręcznie w konfiguracji routera, na którym uruchomiona jest redystrybucja, mogą być redystrybuowane do protokołów routingu. Umieszczanie statycznych tras routera w dynamicznym protokole routingu może być skutecznym sposobem propagacji tych tras w całej sieci.

Trasy podłączone bezpośrednio (ang. *connected routes*)
: Trasy, które znalazły się w tabeli routingu poprzez bezpośrednio podłączone interfejsy na routerze wykonującym redystrybucję, mogą być również redystrybuowane do protokołów routingu. Jeśli trasa podłączona bezpośrednio jest redystrybuowana, sieć, od której pochodzi, będzie dodawana do protokołu routingu, ale interfejs skonfigurowany dla tej sieci nie będzie rozgłaszał ani nasłuchiwał rozgłoszeń tras. Możesz to wykorzystać jako alternatywę dla polecenia `network`, jeśli takie zachowanie jest pożądane.

Inne protokoły routingu
: Trasy mogą być pozyskiwane dynamicznie od innych protokołów routingu, które są aktywne na routerze wykonującym redystrybucję. Z jednego protokołu routingu można redystrybuować trasy do dowolnego innego protokołu routingu. Przykładem redystrybucji między protokołami routingu jest OSPF redystrybuujący do EIGRP.

Ten sam protokół routingu z innego systemu autonomicznego lub procesu
: Protokoły obsługujące systemy autonomiczne, takie jak EIGRP, OSPF czy BGP, mogą redystrybuować pomiędzy tymi systemami. Przykładem pojedynczego protokołu redystrybuującego pomiędzy systemami autonomicznymi może być EIGRP 100 redystrybuujący do EIGRP 200.

Bez względu na to, do jakiego protokołu odbywa się redystrybucja tras, jej źródłem mogą być zawsze tylko pozycje wymienione na tej liście. Do redystrybucji tras stosuje się komendę `redistribute` — z podaniem źródła trasy — w protokole otrzymującym trasy.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

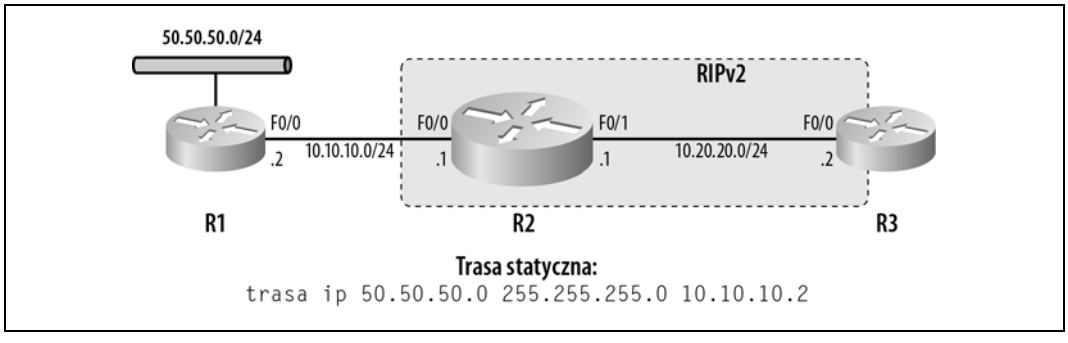
Redystrybucję konfiguruje się w protokole, dla którego trasy są przeznaczone, a nie w tym, z którego pochodzą. Dla protokołu będącego źródłem trasy nie jest wymagana żadna konfiguracja.

Redystrybucja w protokole RIP

Zaczniemy od protokołu RIP, gdyż stosuje on najprostszą metrykę, a dzięki temu również jego konfiguracja jest najłatwiejsza.

Najczęstszym problemem podczas konfiguracji protokołów routingu jest dodanie tras statycznych. Ponieważ z definicji trasy te są statyczne, to nie działają dynamicznie. Ale skoro zostały określone jako statyczne, po co w ogóle dodawać je do dynamicznego protokołu routingu?

Rysunek 11.2 przedstawia prostą sieć, w której wymagana jest redystrybucja trasy statycznej. R1 posiada interfejs bezpośrednio podłączony do sieci 50.50.50.0/24, ale nie ma uruchomionego protokołu routingu. R2 ma skonfigurowaną statyczną trasę do sieci 50.50.50.0/24 kierującą do R1. R2 i R3 komunikują się ze sobą, wykorzystując protokół RIPv2. W takim przypadku R3 nie może dostać się do sieci 50.50.50.0/24, ponieważ R2 nie ogłasza tej trasy.



Rysunek 11.2. Redystrybucja trasy statycznej do RIPv2

Aby RIP mógł ogłaszać trasę statyczną do R3, trasa musi być redystrybuowana w RIP. Oto pełna konfiguracja RIP dla R2:

```
router rip
 version 2
 redistribute static metric 1
 network 10.0.0.0
 no auto-summary
```

Zwróć uwagę na parametr `metric` w komendzie `redistribute`. Określa on metrykę RIP dla wszystkich tras statycznych dodanych do tego protokołu. Innym sposobem na osiągnięcie takiego rezultatu jest zastosowanie polecenia `default-metric`:

```
router rip
 version 2
 redistribute static
 network 10.0.0.0
 default-metric 3
 no auto-summary
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W tym przypadku domyślna metryka wynosi 3. Jeśli ustawisz domyślną metrykę tak, jak to pokazałem, nie będziesz już musiał określać tego parametru, stosując komendę `redistribute static`. Router domyślnie przyporządkuje domyślną metrykę, którą skonfigurowałeś, wszystkim redystrybuowanym trasom statycznym.

Możesz sprawdzić, jaka jest aktualnie domyślna metryka protokołu, wpisując polecenie `show ip protocols`:

```
R2#sho ip protocols
Routing Protocol is "rip"
  Sending updates every 30 seconds, next due in 2 seconds
  Invalid after 180 seconds, hold down 180, flushed after 240
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
  Default redistribution metric is 3 // Domyślna metryka redystrybucji wynosi 3
  Redistributing: static, rip
  Default version control: send version 2, receive version 2
  Automatic network summarization is not in effect
  Maximum path: 4
  Routing for Networks:
    10.0.0.0
  Routing Information Sources:
    Gateway         Distance      Last Update
  Distance: (default is 120)
```

Bądź ostrożny, konfigurując domyślne metryki, gdyż są one stosowane do wszystkich tras redystrybuowanych w procesie routingu, bez względu na ich źródło. Jeśli będziesz chciał teraz redystrybuować EIGRP w protokole RIP, metryka przydzielona w RIP będzie wynosiła 3, ponieważ taką wartość domyślną ustawiłeś. Możesz nadpisać domyślną metrykę, podając żądaną wartość metryki za każdym razem, kiedy wpisujesz komendę `redistribute`. Poniżej określiłem domyślną metrykę 5, ale przeprowadziłem konfigurację tak, aby trasy EIGRP miały metrykę 1, kiedy są redystrybuowane w RIP:

```
router rip
 version 2
 redistribute static
 redistribute eigrp 100 metric 1
 network 10.0.0.0
 default-metric 5
 no auto-summary
```

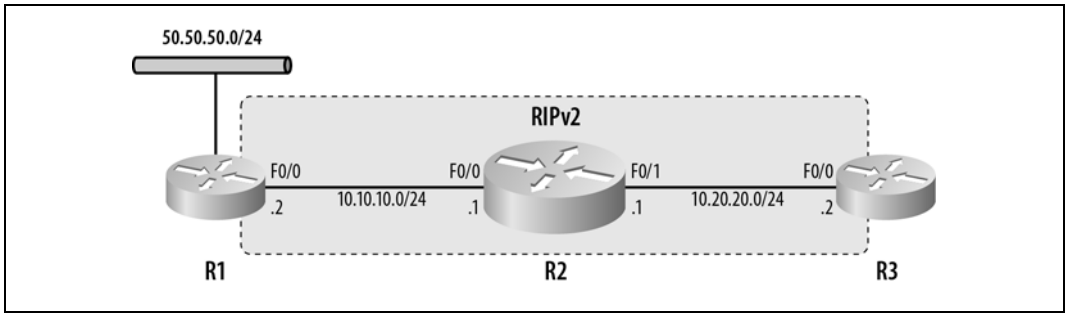
Oto tabela routingu na R3 po zakończeniu konfiguracji R2:

```
R3#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     192.192.192.0/30 is subnetted, 1 subnets
C       192.192.192.4 is directly connected, Serial0/0
     50.0.0.0/24 is subnetted, 1 subnets
R       50.50.50.0 [120/5] via 10.20.20.1, 00:00:07, Ethernet1/0
     10.0.0.0/24 is subnetted, 2 subnets
R       10.10.10.0 [120/1] via 10.20.20.1, 00:00:10, Ethernet1/0
C       10.20.20.0 is directly connected, Ethernet1/0
```

Trasa 50.50.50.0/24 znajduje się w tabeli routingu i posiada metrykę 5. Trasa poniżej jest wynikiem instrukcji `network 10.0.0.0` na R2. Nie powstała ona w procesie redystrybucji i ma normalną metrykę RIP.

Innym powszechnym problemem jest konieczność rozgłaszania sieci, które są podłączone do routera, ale nie są uwzględnione w procesie routingu. Rysunek 11.3 przedstawia sieć z trzema

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 11.3. Redystrybucja tras podłączonych w protokole RIP

routerami, z których każdy uczestniczy w protokole RIPv2. R1 ma podłączoną sieć, która nie jest uwzględniona w procesie RIP. Konfiguracja dla procesu RIP routera R1 jest następująca:

```
router rip
 version 2
 network 10.0.0.0
 no auto-summary
```

W sieci 50.50.50.0/24 nie ma żadnych routerów, więc włączanie RIP na tym interfejsie powodowałoby dodatkowe wysyłanie bezużytecznych broadcastów w tej sieci. Jednak R3 nadal potrzebuje możliwości dotarcia do tej sieci. Aby dodać 50.50.50.0/24 do rozgłoszeń wysyłanych przez R1, musisz redystrybuować podłączone bezpośrednio sieci w protokole RIP, korzystając z komendy `redistribute connected` na R1:

```
router rip
 version 2
 redisribute connected metric 1
 network 10.0.0.0
 no auto-summary
```

Wysyłanie bezużytecznych broadcastów może się wydawać trywialnym problemem, pamiętaj jednak, że RIP wysyła broadcasty zawierające całą tabelę routingu. Pojedynczy pakiet aktualizacyjny RIP może pomieścić jedynie 25 miejsc docelowych. W sieci, która posiada 200 tras, każda aktualizacja składać się będzie z czterech sporych pakietów broadcastowych. Każdy z tych pakietów musi być przetworzony przez wszystkie urządzenia w sieci. Daje to potencjalnie 12k danych co 30 sekund.

Jeśli nie jest to jeszcze wystarczający dowód dla Ciebie, rozważ, że aktualizacje RIP są klasyfikowane jako pakiety **kontrolne** (ang. *control*), co oznacza, że posiadają priorytet *IP precedence* 6 lub DSCP 48. Mają więc wyższy priorytet niż głosowe pakiety RTP (ang. *Real-Time Protocol*), które są sklasyfikowane jako pakiety typu **express forwarding** (*IP precedence* 5 lub DSCP 40). Mówiąc w skrócie, aktualizacje RIP z łatwością mogą wpływać na jakość usług głosowych w sieciach wykorzystujących VoIP.

Teraz R3 widzi 50.50.50.0/24 w swojej tabeli routingu, ponieważ została ona ogłoszona w sieci przez protokół RIP:

```
R3#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     192.192.192.0/30 is subnetted, 1 subnets
C       192.192.192.4 is directly connected, Serial0/0
     50.0.0.0/24 is subnetted, 1 subnets
R       50.50.50.0 [120/2] via 10.20.20.1, 00:00:05, Ethernet1/0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
     10.0.0.0/24 is subnetted, 2 subnets
R       10.10.10.0 [120/1] via 10.20.20.1, 00:00:24, Ethernet1/0
C       10.20.20.0 is directly connected, Ethernet1/0
     192.168.1.0/29 is subnetted, 1 subnets
R       192.168.1.0 [120/2] via 10.20.20.1, 00:00:24, Ethernet1/0
```

Redystrybucja w protokole EIGRP

Protokół EIGRP został zaprojektowany do automatycznej redystrybucji tras IGRP z tego samego ASN. Możesz jednak wyłączyć tę funkcję poleceniem `no redistribute igrp system-autonomiczny`:

```
router eigrp 100
 no redistribute igrp 100
```

W EIGRP trasy są redystrybuowane w ten sam sposób co w przypadku RIP. Może się to jedynie wydawać trudniejsze, ponieważ metryki w EIGRP są bardziej skomplikowane. RIP wykorzystuje jako metryki liczbę przeskoków, podczas gdy EIGRP stosuje kombinację wartości przepustowości oraz opóźnienia dla wszystkich linków danej ścieżki. W rzeczywistości EIGRP wykorzystuje więcej niż tylko te dwie miary, ale domyślnie pozostałe atrybuty nie są włączone. Przy uruchamianiu redystrybucji musisz je jednak skonfigurować, przyjrzymy się więc tym atrybutom.

Tak jak w przypadku RIP możesz użyć polecenia `default-metric`, aby skonfigurować metrykę dla wszystkich redystrybuowanych tras. Możesz też określać metrykę każdorazowo, dopisując ją w linii polecenia `redistribute`. Oto argumenty wymagane w poleceniu `default-metric` protokołu EIGRP oraz zakres ich dopuszczalnych wartości:

- przepustowość (ang. *bandwidth*) w kb/s: 1 – 4 294 967 295;
- opóźnienie (ang. *delay*) w jednostkach 10 mikrosekund: 0 – 4 294 967 295;
- niezawodność (ang. *reliability*), gdzie 255 to 100-procentowa niezawodność: 0 – 255;
- efektywna przepustowość, czyli obciążenie (ang. *loading*), gdzie 255 to 100-procentowe obciążenie: 1 – 255;
- MTU (ang. *Maximum Transmission Unit*): 1 – 65 535.

To, jakie wartości tych argumentów zastosujesz, będzie w dużym stopniu zależało od Twoich aktualnych potrzeb. Pamiętaj, że trasy redystrybuowane są trasami zewnętrznymi, będą więc posiadały wyższy dystans administracyjny niż wewnętrzne trasy EIGRP. Trasy zewnętrzne będą ogłaszane z dystansem administracyjnym 170.

Redystrybuowane trasy muszą wyglądać tak, jakby były linkami, gdyż taką formę akceptuje EIGRP. Jeśli chcesz, aby trasy redystrybuowane były postrzegane jako linki ethernetowe 100 Mb/s, domyślną metrykę musisz skonfigurować w następujący sposób:

```
R3(config-router)#default-metric 100000 10 255 1 1500
```

Nie zawsze jest oczywiste, jakie są właściwe wartości, które należy zastosować w tych poleceniach. Przykładowo przepustowość może być podawana w kb/s, a nie w b/s (łącze 100 Mb/s to 100000 kb/s). Odzwierciedla to sposób wyświetlania informacji o przepustowości dla komendy `show interface`:

```
R2#sho int f0/0 | include BW
  MTU 1500 bytes, BW 100000 Kbit, DLY 100 usec,
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przyjrzyj się też, w jaki sposób komenda `show interface` pokazuje informację o opóźnieniu na interfejsie:

```
R2#sho int f0/0 | include DLY
  MTU 1500 bytes, BW 100000 Kbit, DLY 100 usec,
```

Tutaj opóźnienie podane jest w mikrosekundach, jednak definiując opóźnienie w redystrybucji, musisz używać jednostek 10 mikrosekund. Tak więc, aby uzyskać opóźnienie 100 mikrosekund, należy ustawić jego wartość na 10.

Kiedy konfiguruję redystrybucję, zawsze ustawiam niezawodność 255, obciążenie 1 oraz MTU 1500. Zazwyczaj też określam trasy redystrybuowane jako linki 100 Mb/s, tak jak pokazałem poprzednio. Ma być jak najprościej. Mogą się trafić instancje, w których będziesz potrzebował zastosować inne parametry, jednak są to rzadkie przypadki.

Metoda definiowania domyślnej metryki oraz zastępowania jej jednorazowo inną metryką, opisana już w sekcji „Redystrybucja w protokole RIP" na stronie 156, jest również właściwa w przypadku EIGRP. Poniżej zdefiniowałem domyślną metrykę wskazującą łącze 100 Mb/s z opóźnieniem 100 mikrosekund. Skonfigurowałem jednocześnie redystrybucję procesu OSPF 100 z metryką wskazującą łącze 1000 Mb/s z opóźnieniem 50 mikrosekund:

```
router eigrp 100
 redistribute ospf 100 metric 1000000 5 255 1 1500
 network 10.0.0.0
default-metric 100000 10 255 1 1500
 no auto-summary
```

Redystrybuując trasy OSPF w innym protokole, możesz ograniczyć zakres typów redystrybuowanych tras. Przykładowo możesz redystrybuować jedynie trasy wewnętrzne OSPF, ignorując wszystkie trasy zewnętrzne tego protokołu. W tym celu należy użyć w komendzie `redistribute ospf` argumentu `match`:

```
R2(config-router)#redistribute ospf 100 match ?
  external       Redistribute OSPF external routes //Redystrybucja tras zewnętrznych OSPF
  internal       Redistribute OSPF internal routes //Redystrybucja tras wewnętrznych OSPF
  nssa-external  Redistribute OSPF NSSA external routes //Redystrybucja tras zewnętrznych OSPF
NSSA
```

Wybierając trasy zewnętrzne, możesz dodatkowo określić typ 1 lub typ 2 tras OSPF:

```
R2(config-router)#redistribute ospf 100 match external ?
  1              Redistribute external type 1 routes //Redystrybucja tras zewnętrznych typ 1
  2              Redistribute external type 2 routes //Redystrybucja tras zewnętrznych typ 2
  internal       Redistribute OSPF internal routes //Redystrybucja tras wewnętrznych OSPF
  metric         Metric for redistributed routes //Metryka dla tras redystrybuowanych
  nssa-external  Redistribute OSPF NSSA external routes //Redystrybucja tras zewnętrznych OSPF
NSSA
  route-map      Route map reference //Przywołanie mapy routingu
  <cr>
```

Możesz także łączyć różne typy tras w argumencie `match`. Poniżej przykładowa konfiguracja redystrybucji tras OSPF z procesu 100. Wybrałem trasy wewnętrzne oraz trasy zewnętrzne typu 2:

```
R2(config-router)#redistribute ospf 100 match internal external 2
```

Ponieważ nie podałem metryki, zastosowana zostanie metryka domyślna. Mogłem oczywiście podać metrykę do danej redystrybucji.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Redystrybucja RIP w protokole EIGRP odbywa się w ten sam sposób co redystrybucja OSPF. Nie ma jednak możliwości wyboru rodzaju tras, ponieważ RIP nie obsługuje tak wielu rodzajów tras co OSPF:

```
router eigrp 100
 redistribute rip metric 100000 100 255 1 1500
```

Redystrybucja w protokole OSPF

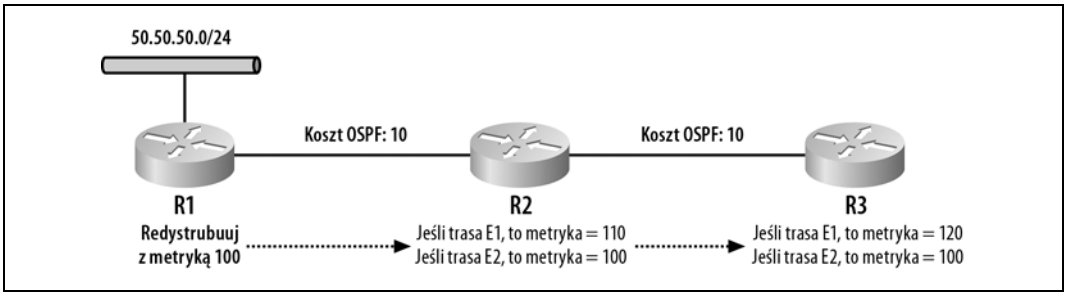
Redystrybucja tras w OSPF odbywa się w ten sam sposób co w przypadku pozostałych protokołów. Metryka stosowana w OSPF jest pochodną przepustowości łączy tworzących trasę. Poniżej konfiguracja łącza 10 Mb/s jako domyślnej metryki OSPF:

```
R3(config-router)#default-metric 10
```

Nie ma innych opcji. Metryka może przyjmować wartości: 1 – 16 777 214. Wartość 1 odpowiada łączu 100 Mb/s (przy ustawieniach domyślnych).

Jeśli nie zdefiniujesz domyślnej metryki lub nie określisz metryki w komendzie `redi➥stribute`, OSPF przypisze metrykę 20 wszystkim redystrybuowanym trasom. Wyjątkiem są trasy z protokołu BGP, które otrzymają metrykę 1.

Choć wszystkie trasy redystrybuowane są trasami zewnętrznymi, OSPF obsługuje dwa rodzaje tras zewnętrznych. Są to trasy zewnętrzne **typu 1** (ang. *type-1*) oraz **typu 2** (ang. *type-2*). W tabeli routingu trasy typu 1 oznaczone są jako `O E1`, a trasy typu 2 jako `O E2`. Trasy `E1` posiadają metrykę, którą ustawiono przy konfiguracji redystrybucji oraz dodatkowo metrykę wszystkich linków znajdujących się w obrębie obszaru autonomicznego OSPF. Trasy `E2` posiadają jedynie metrykę ustawioną przy konfiguracji redystrybucji. Rysunek 11.4 ilustruje, jak zmieniają się metryki OSPF w prostej sieci w zależności od użytego rodzaju trasy zewnętrznej.



Rysunek 11.4. Rodzaje tras zewnętrznych OSPF

OSPF do redystrybucji wybiera domyślnie trasy typu 2. Jaki typ tras powinieneś więc wybrać? Zależy to od Twoich aktualnych potrzeb. Zasadniczo w mniejszych sieciach (powiedzmy mniej niż 10 routerów) trasy `E2` mogą być łatwiejsze w utrzymaniu, ponieważ posiadają tę samą wartość w każdym miejscu systemu autonomicznego OSPF. Według mnie jednak trasy `E1` funkcjonują bardziej logicznie, gdyż zwiększają wartość przy każdym przeskoku, tak jak dzieje się to w przypadku większości metryk.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Aby wybrać rodzaj trasy, dodaj argument `metric-type` do komendy `redistribute`:

```
R3(config-router)#redistribute eigrp 100 metric-type ?
  1  Set OSPF External Type 1 metrics //Ustawia trasy zewnętrzne OSPF typ 1
  2  Set OSPF External Type 2 metrics //Ustawia trasy zewnętrzne OSPF typ 2
```

Rodzaj trasy i wartość metryki można zdefiniować w jednej linii polecenia:

```
R3(config-router)#redistribute eigrp 100 metric-type 2 metric 100
```

W praktyce za każdym razem, gdy konfigurujesz redystrybucję w protokole OSPF, powinieneś dodać również argument `subnets`:

```
R3(config-router)#redistribute eigrp 100 metric 100 subnets
```

Bez argumentu `subnets` OSPF będzie redystrybuował jedynie trasy, które nie zostały podzielone na podsieci. W czasach stosowania VLSM, gdzie praktycznie wszystkie sieci są podsieciami, rzadkością jest sytuacja, w której nie chcesz redystrybuować podsieci.

W najnowszych wersjach IOS, jeśli nie dodasz argumentu `subnets`, zostaniesz ostrzeżony o prawdopodobnym błędzie:

```
R3(config-router)#redistribute eigrp 100 metric 100
% Only classful networks will be redistributed //Redystrybuowane będą jedynie sieci klasowe
```

Redystrybucja wzajemna

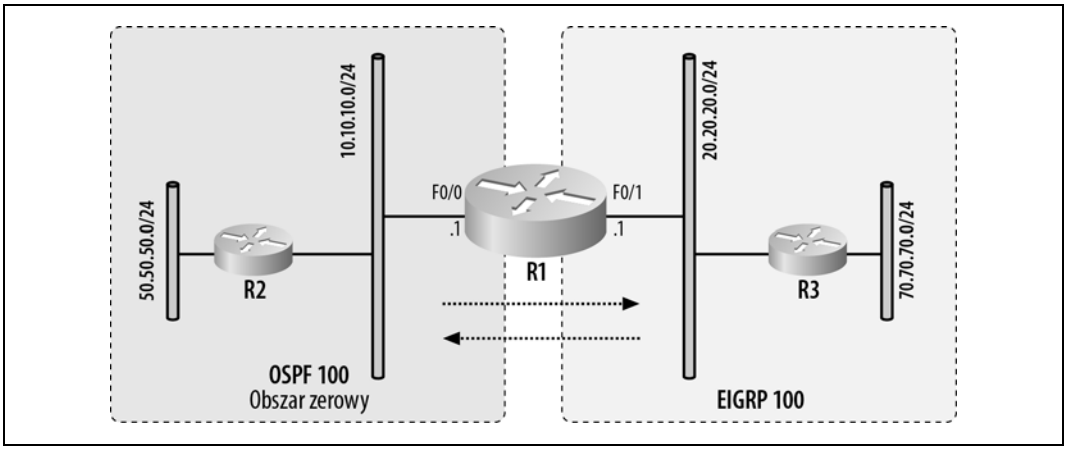
Termin **redystrybucja wzajemna** (ang. *mutual redistribution*) odnosi się do sytuacji, w której router redystrybuuje pomiędzy dwoma protokołami routingu w obu kierunkach, a nie tylko w jednym. Najczęściej redystrybuujemy trasy, ponieważ istnieje w sieci urządzenie lub podmiot, z którym chcemy się łączyć, a który nie obsługuje wybranego przez nas protokołu routingu. Musimy dzielić się trasami pomiędzy protokołami, które nie rozmawiają w tym samym języku, o ile można to tak ująć.

Rysunek 11.5 przedstawia sieć, w której wszystkie podsieci muszą mieć dostęp do siebie nawzajem. Problemem jest to, że sieć po lewej używa protokołu OSPF, a sieć po prawej korzysta z EIGRP. Aby host z sieci 50.50.50.0 znalazł trasę do hosta znajdującego się w sieci 70.70.70.0, trasy EIGRP muszą być redystrybuowane w OSPF. I odwrotnie, jeśli host z 70.70.70.0 będzie chciał skomunikować się z hostem w sieci 50.50.50.0, będzie to możliwe, tylko jeśli trasy OSPF będą redystrybuowane w EIGRP. Ponieważ obie te domeny łączy tylko jeden router, redystrybucja musi odbywać się w obu kierunkach.

Żeby uruchomić na pojedynczym routerze redystrybucję wzajemną, po prostu skonfiguruj każdy z protokołów do redystrybucji drugiego:

```
router eigrp 100
 redistribute ospf 100
 network 20.20.20.0 0.0.0.255
 default-metric 100000 100 255 1 1500
 no auto-summary
!
router ospf 100
 redistribute eigrp 100 metric 100 subnets
 network 10.10.10.0 0.0.0.255 area 0
 default-metric 10
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 11.5. Redystrybucja wzajemna

Konfiguracja jest prosta. Kierując się instrukcjami z poprzednich sekcji, należy najpierw zdefiniować domyślną metrykę dla każdego protokołu, a następnie uruchomić odpowiednio redystrybucję. W przypadku redystrybucji wzajemnej na pojedynczym routerze nie potrzeba żadnych dodatkowych konfiguracji.

Pętle redystrybucji

W przypadku kiedy w sieci redystrybuować będzie kilka routerów, może zacząć się robić interesująco. Trasy redystrybuowane z jednego protokołu do drugiego mogą być też redystrybuowane z powrotem do źródłowego protokołu. Rezultaty w takiej sytuacji bywają dość dziwne. Oryginalne metryki zostaną utracone, a trasa odziedziczy metryki skonfigurowane dla redystrybucji.

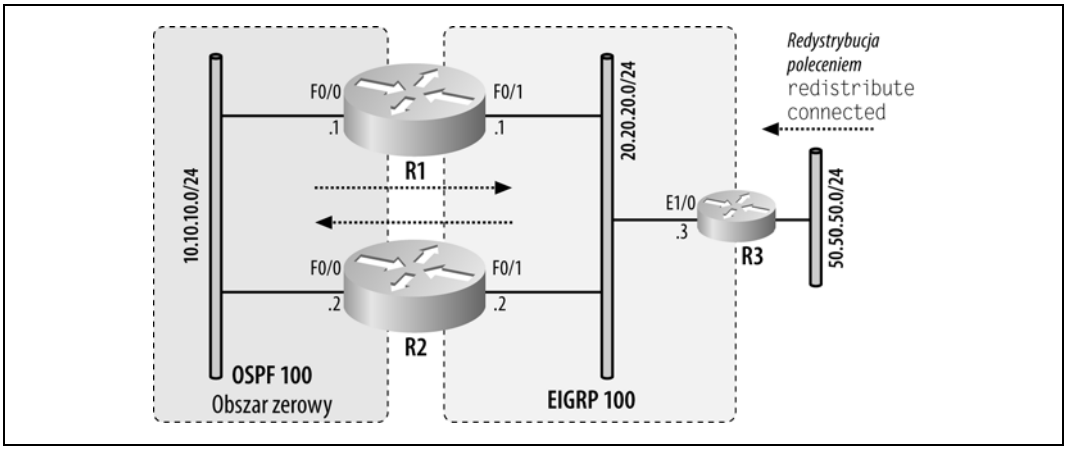
Rysunek 11.6 przedstawia sieć z trzema routerami. Do R3 podłączona jest sieć 50.50.50.0/24, która rozgłaszana jest w EIGRP 100 dzięki zastosowaniu komendy `redistribute connected`. R1 redystrybuuje trasy z OSPF do EIGRP (na rysunku od lewej do prawej), a R2 redystrybuuje trasy z EIGRP do OSPF (od prawej do lewej).

Sieć 50.50.50.0/24 jest ogłaszana z R3 do R1 i R2 przez EIGRP. Z kolei R2 redystrybuuje tę trasę do OSPF 100. W tym momencie R2 posiada wpis dotyczący 50.50.50.0/24 zarówno w bazie danych OSPF, jak i w tablicy topologii EIGRP. Ponieważ pierwotnie trasa była redystrybuowana do EIGRP, docierając do R2, posiada dystans administracyjny 170. R2 ogłasza trasę do R1 via OSPF, który ma dystans administracyjny 110. Pomimo że R1 uzyskał informacje o tej trasie od R3, który jest jej źródłem, wybierze trasę otrzymaną od R2, ponieważ ma ona bardziej atrakcyjny dystans administracyjny.

Poniżej znajdują się tabele routingu z każdego routera. Router R1 uzyskał trasę do 50.50.50.0/24 od R2 z protokołu OSPF:

```
R1#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     50.0.0.0/24 is subnetted, 1 subnets
O E2    50.50.50.0 [110/10] via 10.10.10.2, 00:16:28, FastEthernet0/0
     20.0.0.0/24 is subnetted, 1 subnets
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 11.6. Pętla redystrybucji

```
C       20.20.20.0 is directly connected, FastEthernet0/1
     10.0.0.0/24 is subnetted, 1 subnets
C       10.10.10.0 is directly connected, FastEthernet0/0
```

R2 uzyskał tę trasę od R3 jako trasę zewnętrzną EIGRP. Jest to trasa zewnętrzna, ponieważ była pierwotnie redystrybuowana w EIGRP na R3:

```
R2#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     50.0.0.0/24 is subnetted, 1 subnets
D EX    50.50.50.0 [170/156160] via 20.20.20.3, 00:17:30, FastEthernet0/1
     20.0.0.0/24 is subnetted, 1 subnets
C       20.20.20.0 is directly connected, FastEthernet0/1
     10.0.0.0/24 is subnetted, 1 subnets
C       10.10.10.0 is directly connected, FastEthernet0/0
```

R3 pokazuje tylko swoje dwie podłączone bezpośrednio trasy oraz sieć od strony OSPF, która była redystrybuowana w EIGRP na R2:

```
R3#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     50.0.0.0/24 is subnetted, 1 subnets
C       50.50.50.0 is directly connected, Loopback0
     20.0.0.0/24 is subnetted, 1 subnets
C       20.20.20.0 is directly connected, Ethernet1/0
     10.0.0.0/24 is subnetted, 1 subnets
D EX    10.10.10.0 [170/537600] via 20.20.20.1, 00:00:15, Ethernet1/0
```

Klucz do zrozumienia tego przykładu kryje się w tym, że EIGRP ma wyższy dystans administracyjny dla tras zewnętrznych (170) niż dla tras wewnętrznych (90). OSPF ma jeden dystans administracyjny dla wszystkich typów tras (110).

Tego typu problem pojawiający się w środowiskach produkcyjnych może być przyczyną niekończącego się bólu głowy. Łatwo popełnić taki błąd, jeśli nie masz doświadczenia z redystrybucją w bardziej złożonych sieciach. Symptomami występowania problemu mogą być np. trasy kierujące do miejsc, których się nie spodziewasz. Musisz wtedy przyjrzeć się schematowi sieci i, żeby wykryć przyczynę kłopotów, prześledzić trasy aż do ich źródła. W sieciach,

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

w których odbywa się redystrybucja pomiędzy różnymi systemami autonomicznymi wykorzystującymi ten sam protokół routingu, możesz napotkać trasy krążące między źródłami tam i z powrotem. Takie zachowanie może być spowodowane tym, że każdy system autonomiczny raportuje tę samą metrykę, a router aktualizuje swoje tabele routingu za każdym razem, gdy otrzymuje pakiety aktualizacyjne.

Rozwiązaniem problemu przedstawionego w analizowanym przez nas przykładzie jest przerwanie redystrybucji trasy podłączonej bezpośrednio i dodanie tego interfejsu do EIGRP poleceniem `network`. W ten sposób trasa ta stanie się trasą wewnętrzną z dystansem administracyjnym i będzie bardziej optymalna niż trasa OSPF z dystansem administracyjnym 110.

Ograniczanie redystrybucji

Projektując złożone sieci z wieloma punktami redystrybucji, musisz w jakiś sposób ograniczyć redystrybucję, aby zapobiec tworzeniu się pętli. Opiszę teraz metodę, którą preferuję, czyli znakowanie tras i filtrowanie za pomocą map routingu.

Znakowanie tras

Wiele protokołów routingu — np. EIGRP, OSPF oraz RIPv2 (ale nie RIPv1) — umożliwia znakowanie, czyli inaczej **tagowanie** (ang. *tag*) tras podczas redystrybucji. Tagi (znaczniki), które przypisuje się trasom, to nic innego jak liczby z zakresu 0 – 4 294 967 295 (niestety tagi nie mogą być alfanumeryczne). Znaczniki tras nie mają wpływu na działanie protokołów routingu. Są one po prostu polem, które przyjmuje określoną wartość, wykorzystywaną w innych celach.

Aby ustawić znacznik trasy przy redystrybucji w OSPF, dodaj argument `tag` *`nr-tagu`* do komendy `redistribute`:

```
R2(config-router)#redistribute eigrp 100 metric 10 subnets tag 2
```

Komenda uruchomi redystrybucję tras z EIGRP 100 do OSPF. Metryka OSPF będzie wynosić 10, a znacznik to 2. Jeśli chcesz sprawdzić tagi tras OSPF, użyj polecenia `show ip ospf database`. Redystrybuowane trasy będą trasami zewnętrznymi. Tagi tras podane są w ostatniej kolumnie tabeli:

```
R2#sho ip ospf dat
            OSPF Router with ID (10.10.10.2) (Process ID 100)
                Router Link States (Area 0)
Link ID         ADV Router      Age         Seq#       Checksum Link count
10.10.10.2      10.10.10.2      128         0x80000002 0x00F5BA 1
20.20.20.1      20.20.20.1      129         0x80000002 0x009DD9 1
                Net Link States (Area 0)
Link ID         ADV Router      Age         Seq#       Checksum
10.10.10.1      20.20.20.1      129         0x80000001 0x00B5CA
                Type-5 AS External Link States
Link ID         ADV Router      Age         Seq#       Checksum Tag
20.20.20.0      10.10.10.2      4           0x80000001 0x00D774 2
20.20.20.0      20.20.20.1      159         0x80000001 0x002DF9 0
50.50.50.0      10.10.10.2      4           0x80000001 0x009B56 2
```

Przypisywanie tagów w EIGRP wymaga użycia map routingu (ang. *route maps*). Na szczęście dla tych, którzy mają fobię na tym punkcie, przedstawiona poniżej metoda jest jednym z najłatwiejszych sposobów ich wykorzystania.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Mapy routingu są super! W końcu to powiedziałem. Mapy routingu stanowią całkiem potężne narzędzie. Jeśli się ich obawiasz, powinieneś spędzić nieco czasu, eksperymentując z nimi. Obawa ta wynika najczęściej z mylenia map routingu z listami dostępu (ang. *access lists*) oraz z braku wiedzy na temat ich wzajemnego oddziaływania. Warto poświęcić czas na zgłębienie tematu związanego z mapami routingu. Mogą Cię kiedyś wyciągnąć z technicznego pata (jakim są np. pętle redystrybucji), kiedy wszystkie inne opcje się wyczerpią. Więcej szczegółów na ten temat znajdziesz w rozdziale 14.

Aby dodać znacznik do trasy redystrybuowanej w EIGRP, musisz najpierw stworzyć mapę routingu, a następnie przywołać ją w komendzie `redistribute`. Mapy routingu w swojej najprostszej formie składają się z linii poleceń zawierającej nazwę mapy, instrukcję `permit` (zezwól) lub `deny` (zabroń) oraz numer. Po tym następuje opis akcji, która ma być podjęta. Oto prosta mapa routingu:

```
route-map TAG-EIGRP permit 10
 set tag 3
```

Pierwszy wiersz zawiera nazwę mapy routingu, argument `permit` oraz numer 10. Jest to numer domyślny. Numery używane są do porządkowania kolejności wpisów w mapie routingu, więc jeśli jest tylko jeden wpis, to numer nie ma właściwie znaczenia. Argument `permit` oznacza zezwolenie na wykonanie akcji określonej w wierszu poniżej. W tym przypadku dla naszej mapy routingu ma być wykonana czynność: „ustaw tag o wartości 3”.

Kiedy już stworzyłeś mapę routingu TAG-EIGRP, możesz ją wywołać poleceniem `redistribute` z argumentem `route-map`, po którym wpisujesz nazwę mapy:

```
R2(config-router)#redistribute connected route-map TAG-EIGRP
```

W ten sposób skonfigurowałeś redystrybucję w EIGRP tras podłączonych bezpośrednio, z użyciem domyślnej metryki i z zastosowaniem znacznika wpisanego w mapie routingu `TAG-EIGRP`.

Jeśli chcesz sprawdzić, czy Twój tag został zastosowany, poszukaj w tabeli topologii EIGRP trasy, która powinna być oznakowana. Aby to zilustrować, zastosowałem naszą mapę routingu na routerze R3, który znajduje się w sieci przedstawionej na rysunku 11.6. Tak wygląda tablica topologii EIGRP w R2 dla trasy 50.50.50.0/24:

```
R2#sho ip eigrp top
IP-EIGRP Topology Table for AS(100)/ID(10.10.10.2)
Codes: P - Passive, A - Active, U - Update, Q - Query, R - Reply,
       r - reply Status, s - sia Status
P 10.10.10.0/24, 1 successors, FD is 28160
         via Redistributed (281600/0)
P 20.20.20.0/24, 1 successors, FD is 28160
         via Connected, FastEthernet0/1
P 50.50.50.0/24, 1 successors, FD is 156160, tag is 3
         via 20.20.20.3 (156160/128256), FastEthernet0/1
```

Teraz szczegóły dla trasy 50.50.50.0/24 na routerze R3. Źródłem jest **Rconnected**, co oznacza, że trasa została uzyskana poprzez użycie komendy `redistributed connected`:

```
R3#sho ip eigrp top 50.50.50.0/24
IP-EIGRP (AS 100): Topology entry for 50.50.50.0/24
  State is Passive, Query origin flag is 1, 1 Successor(s), FD is 128256
  Routing Descriptor Blocks:
  0.0.0.0, from Rconnected, Send flag is 0x0
      Composite metric is (128256/0), Route is External
      Vector metric:
        Minimum bandwidth is 10000000 Kbit
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
      Total delay is 5000 microseconds
      Reliability is 255/255
      Load is 1/255
      Minimum MTU is 1514
      Hop count is 0
    External data:
      Originating router is 50.50.50.1 (this system)
      AS number of route is 0
      External protocol is Connected, external metric is 0
      Administrator tag is 3 (0x00000003)
```

Ostatnia linijka informuje, że **znacznik administratora** (ang. *administrator tag*) wynosi 3. Wskazuje to, że trasy redystrybuowane w EIGRP (w szczególności redystrybuowane trasy podłączone bezpośrednio) zostały oznaczone tagiem 3.

EIGRP nie wykorzystuje tej informacji do podejmowania jakichkolwiek działań — jedynie ją przechowuje. Do czego może więc posłużyć Ci tagowanie? Podobnie jak definiujesz znacznik do zastosowania przy redystrybucji tras w protokole routingu, możesz również badać trasy pod kątem znacznika oraz zezwalać na redystrybucję lub wstrzymywać ją, opierając się na znaczniku. Możesz nazwać mnie frajerem, ale uważam, że jest to całkiem fajne.

Aby sprawdzić, jaki jest tag przychodzącej trasy, po raz kolejny musisz skorzystać z mapy routingu. Musisz tak postąpić w przypadku wszystkich protokołów routingu łącznie z OSPF.

Powracając do przykładu z rysunku 11.6, załóżmy, że ustawiliśmy tag 3 dla podłączonej bezpośrednio trasy 50.50.50.0/24 na R3. Gdybyśmy mogli powstrzymać rozgłaszanie tej trasy w OSPF, rozwiązałoby to problem. R1 nigdy by nie uzyskał niewłaściwej informacji o tej trasie.

Jeśli chcesz redystrybuować w OSPF na R2, musisz wskazać routerowi mapę routingu. Nie definiujesz już tagów poleceniem `redistribute`, ponieważ określiłeś je w mapie routingu. Jeśli szukam tagów, wykorzystując mapę routingu, to z zasady konfiguruję je również tutaj. Łatwiej mi ogarnąć wszystko, jeśli mam to w jednym miejscu. Spotkałem się z tym, że sprawdzanie poprzez mapę routingu tagu zdefiniowanego argumentem `tag` nie zawsze działało poprawnie. Poniżej skonfiguruję router do redystrybucji tras EIGRP 100 w protokole OSPF, przypisując im metrykę 10 oraz aplikując schemat logiczny zawarty w mapie routingu `No-EIGRP-Tag3`:

```
router ospf 100
 redistribute eigrp 100 metric 10 subnets route-map No-EIGRP-Tag3
```

Tę mapę routingu skonfigurowałem następująco:

```
route-map No-EIGRP-Tag3 deny 10
 match tag 3
!
route-map No-EIGRP-Tag3 permit 20
 set tag 2
```

Przykład ten jest nieco bardziej skomplikowany niż poprzednia mapa routingu, wciąż jednak dosyć prosty. Pierwsza linijka to wpis `deny` (zabroń). Potem następuje instrukcja, która mówi: „dopasuj wszystko ze znacznikiem 3”. Argument `match` (dopasuj) pojawiający się po `deny` może być dezorientujący, jednak im więcej czasu spędzisz, ćwicząc z mapami routingu, tym więcej odnajdziesz w tym sensu. Następny wpis definiuje warunek dopuszczenia (`permit`) i ustawia dla tras znacznik 2. Biorąc pod uwagę całość, mapa routingu zawiera instrukcje: „dopasuj wszystko ze znacznikiem 3 i odrzuć to”, a następnie „dopasuj wszystko, co pozostało, i przypisz znacznik 2”.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli spojrzysz teraz na bazę danych OSPF routera R2, zauważysz, że zniknęła trasa dla 50.50.50.0/24. Pozostała jednak trasa do 20.20.20.0/24 uzyskana z R1, ponieważ nie była oznaczona tagiem 3:

```
R2#sho ip ospf database
            OSPF Router with ID (10.10.10.2) (Process ID 100)
                Router Link States (Area 0)
Link ID         ADV Router      Age         Seq#       Checksum Link count
10.10.10.2      10.10.10.2      769         0x80000002 0x00F5BA 1
20.20.20.1      20.20.20.1      770         0x80000002 0x009DD9 1
                Net Link States (Area 0)
Link ID         ADV Router      Age         Seq#       Checksum
10.10.10.1      20.20.20.1      771         0x80000001 0x00B5CA
                Type-5 AS External Link States
Link ID         ADV Router      Age         Seq#       Checksum Tag
20.20.20.0      10.10.10.2      224         0x80000001 0x00D774 2
20.20.20.0      20.20.20.1      800         0x80000001 0x002DF9 0
```

Trasa ta nadal istnieje w tabeli topologii EIGRP na routerze R2. Nie była po prostu redystrybuowana w OSPF z uwagi na naszą mapę routingu.

W tabeli routingu na R1 możesz zobaczyć, że 50.50.50.0/24 kieruje teraz do R3 w taki sposób, w jaki chcieliśmy:

```
R1#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     50.0.0.0/24 is subnetted, 1 subnets
D EX    50.50.50.0 [170/156160] via 20.20.20.3, 00:00:16, FastEthernet0/1
     20.0.0.0/24 is subnetted, 1 subnets
C       20.20.20.0 is directly connected, FastEthernet0/1
     10.0.0.0/24 is subnetted, 1 subnets
C       10.10.10.0 is directly connected, FastEthernet0/0
```

Przykład z życia

Dzisiejsze sieci często projektowane są z założeniem zapewnienia gotowości operacyjnej do bycia strukturą podstawową. Jeśli starasz się zaprojektować sieć, w której zostaną wyeliminowane wszystkie punkty awarii, realny wydaje się scenariusz przedstawiony na rysunku 11.7.

Mamy dwa routery, na których skonfigurowano redystrybucję wzajemną (oba redystrybuują EIGRP w OSPF oraz OSPF w EIGRP). Widziałeś już, co może się stać, jeśli każdy z nich redystrybuuje tylko w jednym kierunku. Prawdopodobieństwo powstania chaosu wywołanego przez routery jest tutaj dość duże, jednak taki schemat sieci jest bardzo popularny z powodów, o których już pisałem.

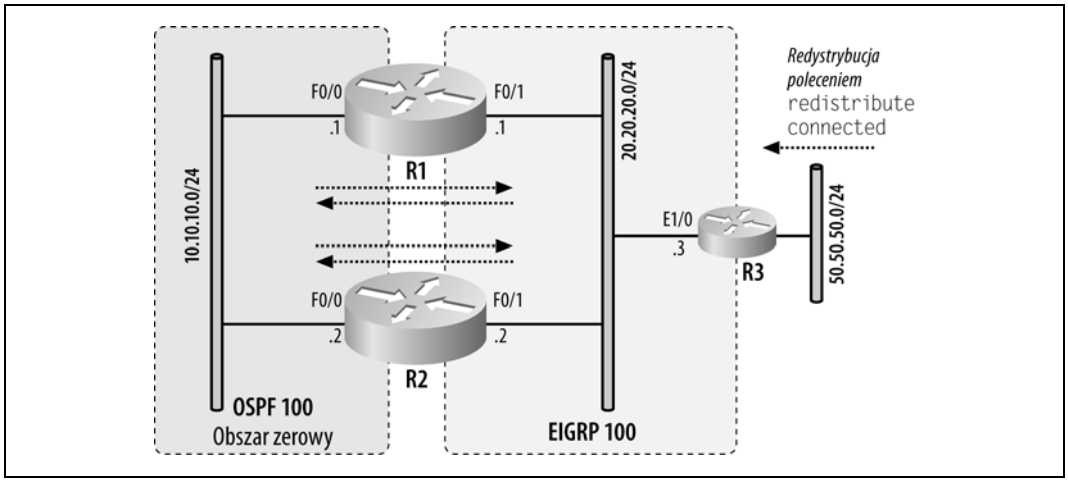
Aby schemat ten zadziałał poprawnie, musimy znów wykorzystać znakowanie tras. Jednak tym razem posłużymy się sprytem (tego nigdy za wiele).

Koncepcja jest prosta: trasy wysyłane z jednego protokołu do drugiego nie będą rozgłaszane z powrotem w protokole źródłowym.

Oto konfiguracja dla R1 z wykorzystaniem dystansu administracyjnego jako znacznika:

```
router eigrp 100
 redistribute ospf 100 route-map OSPF-to-EIGRP
 network 20.20.20.0 0.0.0.255
 default-metric 100000 100 255 1 1500
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 11.7. Dwa routery ze skonfigurowaną redystrybucją wzajemną

Staram się tagować moje trasy numerem routera, na którym pracuję. Nie zawsze jest to możliwe, szczególnie jeśli nazwałeś swój router pomysłowo, np. Boston-PoP-Router lub Michelle. Inną opcją jest tagowanie tras numerem dystansu administracyjnego protokołu routingu, z którego pochodzą — 90 dla EIGRP itd. Można również wykorzystać numery systemów autonomicznych. Cokolwiek wybierzesz, upewnij się, że jest to w miarę możliwości oczywiste. Zawsze staraj się też prowadzić dokumentację swoich dokonań, aby inni mieli szansę zrozumienia Twojego geniuszu.

```
 no auto-summary
!
router ospf 100
 log-adjacency-changes
 redistribute eigrp 100 subnets route-map EIGRP-to-OSPF
 network 10.10.10.0 0.0.0.255 area 0
 default-metric 100
!
route-map EIGRP-to-OSPF deny 10
 match tag 110
!
route-map EIGRP-to-OSPF permit 20
 set tag 90
!
route-map OSPF-to-EIGRP deny 10
 match tag 90
!
route-map OSPF-to-EIGRP permit 20
 set tag 110
```

Mapy routingu dla R2 są takie same, ponieważ zostały zastosowane identyczne reguły. Sprawdzona musi być każda trasa redystrybuowana z innych routerów.

Trasy redystrybuowane z OSPF do EIGRP zostaną oznaczone tagiem 110. Z kolei trasy redystrybuowane z EIGRP do OSPF otrzymają tag 90. Jeśli w OSPF znajdziemy trasę ze znacznikiem 90, będziemy wiedzieć, że jej źródłem jest EIGRP, ponieważ taką regułę znakowania przyjęliśmy. Przy próbie ponownej redystrybucji tej trasy do źródła, czyli z OSPF do EIGRP na drugim rou-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

terze, mapa routingu ją zablokuje. W ten sposób trasy nie będą redystrybuowane do EIGRP, jeśli pochodzą z EIGRP. Analogicznie będzie z dystrybucją do OSPF tras, których źródłem jest właśnie OSPF.

O ile formuła ta eliminuje pętle routingu, to nie rozwiązuje problemu sieci 50.50.50.0/24 routera R3, która jest rozgłaszana poprzez niewłaściwy protokół. Jeśli chodzi o tę sieć, to w tym momencie R2 kieruje nas z powrotem do OSPF:

```
R2#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     50.0.0.0/24 is subnetted, 1 subnets
O E2    50.50.50.0 [110/100] via 10.10.10.1, 00:11:50, FastEthernet0/0
     20.0.0.0/24 is subnetted, 1 subnets
C       20.20.20.0 is directly connected, FastEthernet0/1
     10.0.0.0/24 is subnetted, 1 subnets
C       10.10.10.0 is directly connected, FastEthernet0/0
```

Każda sieć jest inna i zawsze będą pojawiać się wyzwania, którym trzeba stawić czoła. W tym przypadku prawdopodobnie lepiej wyszlibyśmy na zastosowaniu poprzedniego schematu, w którym routery redystrybuowały tylko w jednym kierunku, mimo że schemat ten nie jest zbyt elastyczny.

Sposobem rozwiązania problemu i skonfigurowania wielu punktów redystrybucji wzajemnej jest połączenie tych dwóch scenariuszy. Dzięki mapom routingu możemy dopasować wiele tagów, więc poza wstrzymywaniem tras redystrybuowanych uprzednio w EIGRP na R2 możemy również dopasowywać do tagu 3 zdefiniowanego na R3:

```
route-map EIGRP-to-OSPF deny 10
 match tag 110 3
!
route-map EIGRP-to-OSPF permit 20
 set tag 90
```

Linia `match tag 110 3` oznacza: „dopasuj do tagu 110 lub 3". Teraz R2 posiada właściwe trasy:

```
R2#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     50.0.0.0/24 is subnetted, 1 subnets
D EX    50.50.50.0 [170/156160] via 20.20.20.3, 00:00:01, FastEthernet0/1
     20.0.0.0/24 is subnetted, 1 subnets
C       20.20.20.0 is directly connected, FastEthernet0/1
     10.0.0.0/24 is subnetted, 1 subnets
C       10.10.10.0 is directly connected, FastEthernet0/0
```

Metoda alternatywna

Przedstawię teraz kolejną metodę, która podoba mi się ze względu na swoją elegancję: ponieważ trasy redystrybuowane są trasami zewnętrznymi, należy zezwolić na redystrybucję jedynie tras wewnętrznych. Trasy, które przeszły już raz proces redystrybucji i stały się zewnętrzne, nie będą ponownie redystrybuowane.

Redystrybucja tras OSPF do innego protokołu jest prosta. Argument `match` w poleceniu `redistribute` pozwala dopasować typ trasy:

```
router eigrp 100
 redistribute ospf 100 match internal
 network 20.20.20.0 0.0.0.255
 default-metric 100000 100 255 1 1500
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
 no auto-summary
```

Przy redystrybucji pozostałych protokołów musisz uciec się do map routingu:

```
router ospf 100
 redistribute eigrp 100 route-map Only-Internal subnets
 network 10.10.10.0 0.0.0.255 area 0
 default-metric 100
!
route-map Only-Internal permit 10
 match route-type internal
```

Metoda ta rozwiązuje obydwa nasze problemy. Ponieważ trasa 50.50.50.0/24 jest trasą zewnętrzną z uwagi na pierwotną redystrybucję w R3, nie będzie ogłaszana w protokole OSPF. To, co przedtem wymagało wielu linii kodu oraz kilku map routingu, zostało teraz osiągnięte za pomocą jednego argumentu i pojedynczej mapy routingu składającej się z dwóch linijek. Prostota jest dobra.

Oto końcowa tabela routingu z routera R2:

```
R2#sho ip route
[fragment usunięty]
Gateway of last resort is not set
     50.0.0.0/24 is subnetted, 1 subnets
D EX    50.50.50.0 [170/156160] via 20.20.20.3, 00:13:30, FastEthernet0/1
     20.0.0.0/24 is subnetted, 1 subnets
C       20.20.20.0 is directly connected, FastEthernet0/1
     10.0.0.0/24 is subnetted, 1 subnets
C       10.10.10.0 is directly connected, FastEthernet0/0
```

Tak długo, jak długo uda Ci się zachować prostotę, tagowanie i filtrowanie tras redystrybuowanych będzie łatwe. Im bardziej skomplikowana jest struktura, tym trudniej zapewnić właściwe działanie wszystkich redystrybuowanych sieci. Dodatkowo w złożonej sieci możesz nie mieć możliwości zastosowania tego rozwiązania, ponieważ mogą w niej istnieć uzasadnione trasy zewnętrzne, które muszą być redystrybuowane.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 12.

Tunele

Tunel (ang. *tunnel*) jest metodą połączenia, w której urządzenie lokalne komunikuje się z urządzeniem odległym, tak jakby to drugie również było lokalne. Istnieje wiele rodzajów tuneli. Przykładem mogą być wirtualne sieci prywatne, czyli VPN-y (ang. *Virtual Private Networks*), lub tunele tworzone przez protokół GRE (ang. *Generic Routing Encapsulation*). Formą tunelu jest również SSH (ang. *Secure Shell*), różni się on jednak nieco od pozostałych dwóch. Przyjrzyjmy się tym trzem typom tuneli:

GRE

Tunele GRE umożliwiają podłączenie odległych sieci tak, że sprawiają wrażenie podłączonych lokalnie. GRE nie zapewnia szyfrowania, umożliwia jednak przekazywanie pakietów broadcast i multicast. Jeśli chcesz, aby protokół routingu tworzył relację przylegania z sąsiadami lub wymieniał trasy poprzez tunel, prawdopodobnie musisz skonfigurować właśnie GRE. Tunele GRE tworzone są często wewnątrz tuneli VPN, żeby wykorzystać zalety szyfrowania. Protokół GRE został opisany w publikacjach RFC 1701 oraz 2784.

VPN

Tunele VPN również pozwalają na podłączenie sieci odległych tak, by wyglądały na podłączone lokalnie. VPN szyfruje informacje przed wysłaniem ich w sieć, zazwyczaj jednak nie przekazuje pakietów broadcast i multicast. W konsekwencji często tworzy się wewnątrz tuneli VPN tunele GRE, aby umożliwić działanie protokołów routingu. Zazwyczaj VPN-y są wykorzystywane jako zdalny dostęp do zabezpieczonych sieci.

Rozróżnia się dwa podstawowe typy tuneli VPN: **point-to-point** (punkt-punkt) oraz **remote access** (zdalnego dostępu). VPN-y punkt-punkt oferują łączność pomiędzy dwoma odległymi routerami, tworząc wirtualne łącze. Z kolei VPN-y zdalnego dostępu są tunelami pojedynczego użytkownika tworzącymi łącze między użytkownikiem a routerem, firewallem lub **koncentratorem VPN** (ang. *VPN concentrator*). Ten ostatni jest wyspecjalizowanym urządzeniem służącym tylko do tworzenia VPN-ów.

VPN-y zdalnego dostępu wymagają z reguły zainstalowania na komputerze osobistym klienckiego oprogramowania VPN. Klient komunikuje się z urządzeniem VPN, aby nawiązać osobiste łącze wirtualne.

SSH

SSH jest aplikacją typu klient/serwer, która umożliwia bezpieczne połączenie z serwerami. W praktyce używa się jej w podobny sposób jak telnetu. Przewagą SSH nad telnetem jest szyfrowanie wszystkich danych przed wysyłką. SSH nie zostało zaprojektowane jako

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

tunel w takim sensie, w jakim za tunel uważa się VPN czy GRE. Mimo to może być wykorzystywane do łączenia się z dodatkowymi urządzeniami zdalnymi poza tym urządzeniem, do którego już się podłączyłeś. O ile SSH nie ma bezpośredniego zastosowania na routerach Cisco, o tyle koncepcja jest zbliżona do tuneli VPN i GRE, warto więc o niej wspomnieć. Osobiście stosuję właśnie SSH zamiast VPN-u do łączenia się z moją siecią domową.

Tunele mogą szyfrować dane tak, żeby tylko druga strona mogła je odczytać — co ma miejsce w przypadku SSH. Mogą też wywoływać wrażenie, że sieć odległa podłączona jest lokalnie, jak w przypadku GRE. I w końcu mogą spełniać obie te funkcje, tak jak robi to VPN.

W tym rozdziale jako przykład wykorzystam tunele GRE, gdyż są one najprostsze w konfiguracji i najłatwiejsze do zrozumienia. Tunele GRE są jedynie środkiem służącym do podłączenia sieci odległych tak, jakby były podłączone lokalnie. Uruchamiają one zdalny interfejs na drugim routerze, co sprawia wrażenie, że jest on bezpośrednio podłączony do routera lokalnego, mimo iż między nimi może się znajdować wiele innych routerów i sieci. GRE nie szyfruje danych.

Tunele GRE

Aby utworzyć tunel GRE, musisz skonfigurować wirtualne interfejsy na routerach na obu końcach łącza, a tunel musi zaczynać się i kończyć na istniejących rutowalnych adresach IP. Tunel nie jest łączem fizycznym, tylko logicznym, i jako taki musi opierać swoje istnienie na utworzonych już trasach. Tunel zachowuje się jak łącze fizyczne pod tym względem, że wymaga adresu IP po każdej stronie linku. Adresy te będą adresami IP interfejsu tunelu. Ponieważ link jest wirtualny, musi on dodatkowo mieć informacje, gdzie się zaczyna i gdzie kończy. Źródłem i miejscem docelowym powinny być adresy IP na routerach po obu stronach tunelu.

Najlepszym sposobem zagwarantowania dostępności punktu źródłowego i punktu docelowego tunelu jest zastosowanie jako celu interfejsów pętli zwrotnej (ang. *loopback interfaces*) na każdym końcu. W takiej sytuacji, nawet jeśli istnieje wiele ścieżek do routera, punkty źródłowy i docelowy nie będą zależne od pojedynczego łącza, ale raczej od logicznego interfejsu na samym routerze.

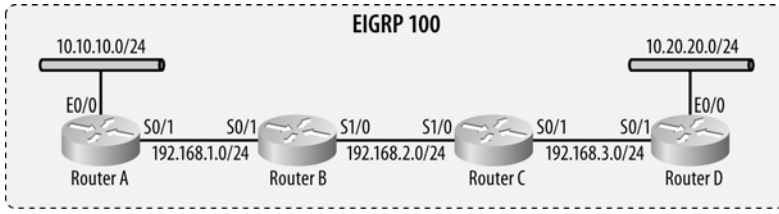
Interfejsy pętli zwrotnej różnią się od adresów IP pętli zwrotnej. Do interfejsu pętli zwrotnej może być przypisany każdy prawidłowy adres IP. Interfejs pętli zwrotnej na routerze jest jego interfejsem logicznym, który zawsze jest włączony. Może on być dodawany w protokołach routingu lub innych funkcjach jak każdy inny interfejs, z tym wyjątkiem, że interfejs pętli zwrotnej ma zawsze status *up/up* (interfejs jest włączony, a protokół linii jest aktywny). Na routerze możesz skonfigurować wiele interfejsów pętli zwrotnej.

Rysunek 12.1 przedstawia przykład prostej sieci, w której zbudujemy tunel GRE. Podłączone są tu cztery routery. Wszystkie pracują w protokole EIGRP, a trasy podłączone bezpośrednio są redystrybuowane w tabeli routingu. Nasz przykład zilustruje, jak ścieżka z Routera A do sieci 10.20.20.0/24 zmieni się, kiedy uruchomimy tunel GRE.

Tabela routingu na Routerze A, znajdującym się w tej sieci, wygląda następująco:

```
Router-A#sho ip route
Gateway of last resort is not set
     10.0.0.0/24 is subnetted, 2 subnets
D       10.20.20.0 [90/3196416] via 192.168.1.2, 03:39:06, Serial0/1
C       10.10.10.0 is directly connected, Ethernet0/1
```

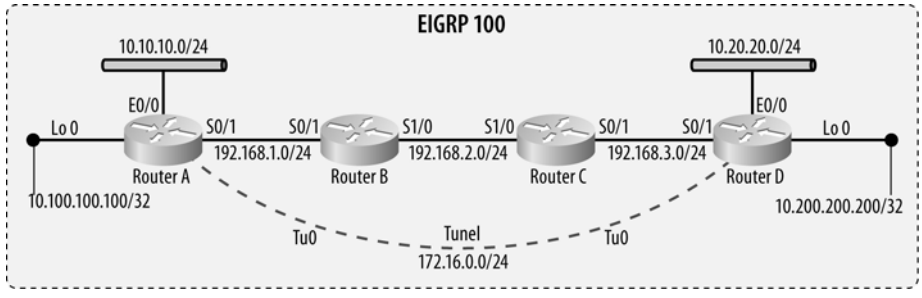

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 12.1. Prosta sieć

```
C    192.168.1.0/24 is directly connected, Serial0/1
D    192.168.2.0/24 [90/2681856] via 192.168.1.2, 03:39:06, Serial0/1
D    192.168.3.0/24 [90/3193856] via 192.168.1.2, 03:39:06, Serial0/1
```

Wszystkie trasy, z wyjątkiem tras podłączonych bezpośrednio, są dostępne przez Serial0/1. Dodamy teraz tunel pomiędzy Routerem A oraz Routerem D. Ponieważ preferujemy linkowanie tuneli do interfejsów pętli zwrotnej, dodamy po jednym interfejsie na każdym routerze. Rysunek 12.2 pokazuje sieć, którą stworzymy.



Rysunek 12.2. Prosta sieć z tunelem

Najpierw dodamy interfejs pętli zwrotnej na Routerze A:

```
Router-A#conf t
Enter configuration commands, one per line. End with CNTL/Z.
Router-A(config)# int lo 0
Router-A(config-if)# ip address 10.100.100.100 255.255.255.255
```

Teraz interfejs pętli zwrotnej na Routerze D:

```
Router-D#conf t
Enter configuration commands, one per line. End with CNTL/Z.
Router-D(config)# int lo 0
Router-D(config-if)# ip address 10.200.200.200 255.255.255.255
```

Ponieważ redystrybuujemy interfejsy podłączone bezpośrednio w EIGRP, obydwa są teraz widoczne w tabeli routingu Routera A:

```
Router-A#sho ip route
Gateway of last resort is not set
     10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
D       10.20.20.0/24 [90/3196416] via 192.168.1.2, 03:50:27, Serial0/1
C       10.10.10.0/24 is directly connected, Ethernet0/1
C       10.100.100.100/32 is directly connected, Loopback0
D EX    10.200.200.200/32 [170/3321856] via 192.168.1.2, 00:00:52, Serial0/1
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
C    192.168.1.0/24 is directly connected, Serial0/1
D    192.168.2.0/24 [90/2681856] via 192.168.1.2, 03:50:27, Serial0/1
D    192.168.3.0/24 [90/3193856] via 192.168.1.2, 03:50:27, Serial0/1
```

Skoro adresy pętli zwrotnej są już widoczne w tabeli routingu, pora utworzyć tunel. Jest to prosty proces. Zaczniemy od skonfigurowania wirtualnych interfejsów dla każdej strony tunelu. Interfejsy tunelu są ponumerowane jak wszystkie interfejsy w IOS. Początkowym interfejsem jest tunel 0:

```
Router-A(config)#int tunnel ?
  <0-2147483647>  Tunnel interface number //Numer interfejsu tunelu
Router-A(config)#int tunnel 0
Router-A(config-if)#
23:23:39: %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel0, changed state to
down //Protokół linii na interfejsie Tunel0, zmiana statusu na nieaktywny
```

Tunele muszą posiadać istniejące, rutowalne adresy IP jako punkty początkowy i końcowy. Muszą one być skonfigurowane na obu routerach. **Źródłem** (ang. *source*) tunelu jest strona lokalna, a **miejscem docelowym** (ang. *destination*) strona zdalna (z perspektywy routera, który konfigurujesz):

```
Router-A(config-if)#ip address 172.16.0.1 255.255.255.0
Router-A(config-if)# tunnel source loopback 0 //Źródło tunelu loopback 0
Router-A(config-if)# tunnel destination 10.200.200.200 //Miejsce docelowe tunelu 10.200.200.200
23:25:15: %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel0, changed state to up
                                        //Protokół linii na interfejsie Tunel0, zmiana statusu na aktywny
```

Gdy tylko dodamy docelowy adres IP do tunelu, interfejs tunelu na Routerze A stanie się aktywny:

```
Router-A#sho int tu0
Tunnel0 is up, line protocol is up
  Hardware is Tunnel
  Internet address is 172.16.0.1/24
  MTU 1514 bytes, BW 9 Kbit, DLY 500000 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation TUNNEL, loopback not set
Keepalive not set
 Tunnel source 10.100.100.100 (Loopback0), destination 10.200.200.200
 Tunnel protocol/transport GRE/IP, key disabled, sequencing disabled
 Checksumming of packets disabled, fast tunneling enabled
 Last input never, output never, output hang never
 Last clearing of "show interface" counters never
 Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
 Queueing strategy: fifo
 Output queue :0/0 (size/max)
 5 minute input rate 0 bits/sec, 0 packets/sec
 5 minute output rate 0 bits/sec, 0 packets/sec
    0 packets input, 0 bytes, 0 no buffer
    Received 0 broadcasts, 0 runts, 0 giants, 0 throttles
    0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored, 0 abort
    0 packets output, 0 bytes, 0 underruns
    0 output errors, 0 collisions, 0 interface resets
    0 output buffer failures, 0 output buffers swapped out
```

Zauważ, że chociaż Router A raportuje włączony interfejs, nie można przesyłać danych tym łączem, ponieważ Router D nie posiada jeszcze interfejsu tunelu. Musisz na to uważać, gdyż możesz się łatwo pomylić, jeśli będziesz działał pod presją. W tablicy routingu Routera A jest nawet aktywna sieć tunelu, czyli 176.16.0.0/24 (w tym momencie nie znajdziesz jej jeszcze na Routerze D):

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Router-A#sho ip route
Gateway of last resort is not set
     172.16.0.0/24 is subnetted, 1 subnets
C       172.16.0.0 is directly connected, Tunnel0
     10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
D       10.20.20.0/24 [90/3196416] via 192.168.1.2, 04:25:38, Serial0/1
C       10.10.10.0/24 is directly connected, Ethernet0/1
C       10.100.100.100/32 is directly connected, Loopback0
D EX    10.200.200.200/32 [170/3321856] via 192.168.1.2, 00:36:03, Serial0/1
C    192.168.1.0/24 is directly connected, Serial0/1
D    192.168.2.0/24 [90/2681856] via 192.168.1.2, 04:25:39, Serial0/1
D    192.168.3.0/24 [90/3193856] via 192.168.1.2, 04:25:39, Serial0/1
```

Aby skonfigurować zakończenie tunelu na Routerze D, musimy dodać tu interfejs tunelu. Użyjemy tego samego polecenia co w przypadku Routera A, ale odwrócimy źródło i miejsce docelowe:

```
Router-D(config)#int tu 0
23:45:13: %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel0, changed state to
down
Router-D(config-if)# ip address 172.16.0.2 255.255.255.0
Router-D(config-if)# tunnel source lo 0
Router-D(config-if)# tunnel destination 10.100.100.100
Router-D(config-if)#
23:47:06: %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel0, changed state to up
```

Mamy teraz aktywny link pomiędzy tymi routerami, który wygląda jak podłączony bezpośrednio. Mimo to sieć Ethernet na Routerze D jest nadal pozyskiwana jedynie przez łącze szeregowe, tak jak dotychczas, a nie przez tunel. Skoro tunel jest teoretycznie podłączony bezpośrednio do obu routerów, to czemu nie jest preferowaną ścieżką?

Przywołując nasze doświadczenie, możemy skorzystać z polecenia `show ip eigrp topology`, aby dowiedzieć się, co EIGRP „wie" na temat ścieżek:

```
Router-A#sho ip eigrp top
5d18h: %SYS-5-CONFIG_I: Configured from console by console
IP-EIGRP Topology Table for AS(100)/ID(192.168.1.1)
Codes: P - Passive, A - Active, U - Update, Q - Query, R - Reply,
       r - reply Status, s - sia Status
P 10.20.20.0/24, 1 successors, FD is 3196416
         via 192.168.1.2 (3196416/2684416), Serial0/1
         via 172.16.0.2 (297246976/28160), Tunnel0
P 10.10.10.0/24, 1 successors, FD is 281600
         via Connected, Ethernet0/1
P 192.168.1.0/24, 1 successors, FD is 2169856
         via Connected, Serial0/1
```

Obie ścieżki zdają się figurować w tabeli, jednak dystans na ścieżce tunelu jest ogromny w porównaniu z dystansem na interfejsie szeregowym. Aby dowiedzieć się, czemu tak jest, przyjrzyjmy się wirtualnemu interfejsowi tunelu:

```
Router-A#sho int tu 0
Tunnel0 is up, line protocol is up
  Hardware is Tunnel
  Internet address is 172.16.0.1/24
  MTU 1514 bytes, BW 9 Kbit, DLY 500000 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation TUNNEL, loopback not set
  Keepalive not set
  Tunnel source 10.100.100.100 (Loopback0), destination 10.200.200.200
  Tunnel protocol/transport GRE/IP, key disabled, sequencing disabled
  Checksumming of packets disabled,  fast tunneling enabled
  Last input 00:00:00, output 00:00:00, output hang never
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Last clearing of "show interface" counters never
Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
Queueing strategy: fifo
Output queue :0/0 (size/max)
5 minute input rate 0 bits/sec, 0 packets/sec
5 minute output rate 0 bits/sec, 0 packets/sec
   88293 packets input, 7429380 bytes, 0 no buffer
   Received 0 broadcasts, 0 runts, 0 giants, 0 throttles
   0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored, 0 abort
   80860 packets output, 6801170 bytes, 0 underruns
   0 output errors, 0 collisions, 0 interface resets
   0 output buffer failures, 0 output buffers swapped out
```

Zwróć uwagę na przepustowość i opóźnienie dla interfejsu tunelu. Mamy tu domyślnie ekstremalnie niską przepustowość oraz skrajnie wysokie opóźnienie. Ponieważ EIGRP używa tych metryk do określenia realnego dystansu, tunel wydaje się znacznie mniej optymalną ścieżką niż istniejące łącza szeregowe. W sumie jest to korzystne, ponieważ tunel może być zbudowany na dużej liczbie routerów i łączy. Tunel jest rozwiązaniem programowym, co oznacza, że opóźnienie przetwarzania dla interfejsu jest zmienne (inaczej niż dla interfejsu fizycznego). Tunele nie powinny być domyślnie najbardziej optymalnymi ścieżkami.

Aby udowodnić, że tunel działa, oraz pokazać, jak zachowuje się wirtualne łącze, prześledźmy trasę z Routera do najbliższego interfejsu szeregowego na Routerze D, czyli 192.168.3.2 na S0/1:

```
Router-A#trace 192.168.3.2
Type escape sequence to abort. //Naciśnij Escape, aby przerwać
Tracing the route to 192.168.3.2 //Śledzenie trasy do 192.168.3.2
  1 192.168.1.2 4 msec 4 msec 0 msec
  2 192.168.2.2 4 msec 4 msec 4 msec
  3 192.168.3.2 4 msec 4 msec 4 msec
Router-A#
```

A teraz prześledźmy trasę do zdalnego końca tunelu na Routerze D (172.16.0.2), czyli tym samym co powyżej. Jest on oddalony o trzy fizyczne przeskoki:

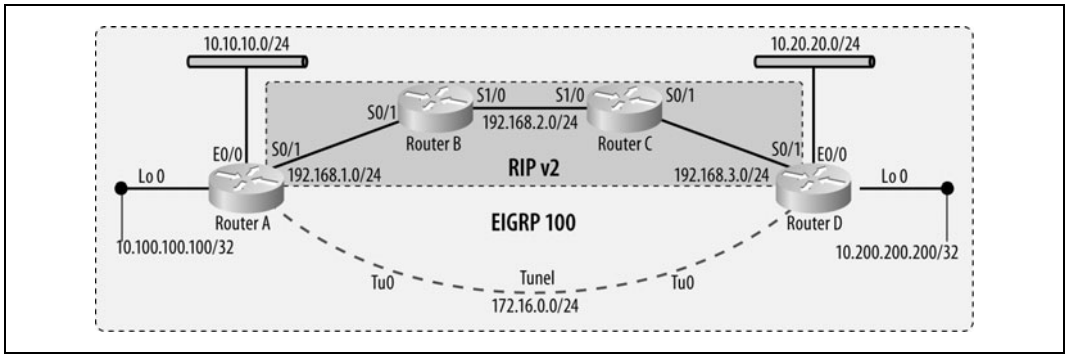
```
Router-A#trace 172.16.0.2
Type escape sequence to abort.
Tracing the route to 172.16.0.2
  1 172.16.0.2 4 msec 4 msec 4 msec
Router-A#
```

Router A postrzega drugi koniec tunelu, jakby był drugim końcem łącza. Jednak w rzeczywistości tunel jest konstrukcją logiczną składającą się z wielu urządzeń pośredniczących, które są niewidoczne. Tunel ukrywa w szczególności to, że na ścieżce znajdują się Routery B i C.

Tunele GRE i protokoły routingu

Zastosowanie protokołu routingu w tunelu GRE może być przyczyną powstawania interesujących problemów. Przykładem może być nasza sieć, której obecny kształt przedstawia rysunek 12.3. Łącza pomiędzy routerami aktualizują trasy, wykorzystując protokół RIPv2. Pozostałe interfejsy na Routerach A i D redystrybuują w RIP poprzez komendę `redistribute connected`. EIGRP działa na wszystkich interfejsach Routerów A i D z wyjątkiem łączy szeregowych, które pracują w RIPv2.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 12.3. Przykład routingu rekurencyjnego

Choć może wydać się to nieco dziwne, musisz brać pod uwagę możliwość, że Routery B i C są własnością usługodawcy i są przez niego zarządzane. Nie mamy nad nimi kontroli, a jedynym protokołem routingu, jaki obsługują, jest RIPv2. Na naszych routerach (A i D) uruchomiony jest protokół EIGRP, więc chcemy rutować między nimi, wykorzystując ten właśnie protokół.

Oto odpowiednie konfiguracje dla Routera A i Routera D (pamiętaj, że w tym scenariuszu Routery B i C pozostają poza naszą kontrolą):

Router A:

```
interface Loopback0
 ip address 10.100.100.100 255.255.255.255
!
interface Tunnel0
 ip address 172.16.0.1 255.255.255.0
 tunnel source Loopback0
 tunnel destination 10.200.200.200
!
interface Ethernet0/0
 ip address 10.10.10.1 255.255.255.0
!
interface Serial0/1
 ip address 192.168.1.1 255.255.255.0
router eigrp 100
 network 10.10.10.0 0.0.0.255
 network 10.100.100.0 0.0.0.255
 network 172.16.0.0 0.0.0.255
 no auto-summary
!
router rip
 version 2
 redistribute connected
 passive-interface Ethernet0/0
 passive-interface Loopback0
 passive-interface Tunnel0
 network 192.168.1.0
 no auto-summary
```

Router D:

```
interface Loopback0
 ip address 10.200.200.200 255.255.255.255
!
interface Tunnel0
 ip address 172.16.0.2 255.255.255.0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
 tunnel source Loopback0
 tunnel destination 10.100.100.100
!
interface Ethernet0/0
 ip address 10.20.20.1 255.255.255.0
!
interface Serial0/1
 ip address 192.168.3.2 255.255.255.0
!
router eigrp 100
 network 10.20.20.0 0.0.0.255
 network 10.200.200.0 0.0.0.255
 network 172.16.0.0 0.0.0.255
 no auto-summary
!
router rip
 version 2
 redistribute connected
 passive-interface FastEthernet0/0
 passive-interface Loopback0
 passive-interface Tunnel0
 network 192.168.3.0
 no auto-summary
```

Wszystko wygląda w porządku, a tunel zaczyna działać praktycznie natychmiast. Jednak już chwilę później pojawiają się na konsoli komunikaty błędów:

```
1d01h: %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel0, changed state to up
                                                //Protokół linii na interfejsie Tunel0, status zmieniony na aktywny
1d01h: %TUN-5-RECURDOWN: Tunnel0 temporarily disabled due to recursive routing
                                                //Tunel0 tymczasowo wyłączony z powodu routingu rekurencyjnego
1d01h: %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel0, changed state to down /
                                                /Protokół linii na interfejsie Tunel0, status zmieniony na nieaktywny
```

Błąd `Tunnel0 temporarily disabled due to recursive routing` wynika z tego, że miejsce docelowe dla tunelu zostało znalezione właśnie poprzez tunel. Jeśli tunel zostanie ręcznie wyłączony na Routerze A, interfejs pętli zwrotnej na Routerze B będzie odnajdowany poprzez protokół RIP, tak jak można się tego było spodziewać:

```
Router-A#sho ip route
Gateway of last resort is not set
     10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
R       10.20.20.0/24 [120/3] via 192.168.1.2, 00:00:07, Serial0/1
C       10.10.10.0/24 is directly connected, Ethernet0/1
C       10.100.100.100/32 is directly connected, Loopback0
R       10.200.200.200/32 [120/3] via 192.168.1.2, 00:00:07, Serial0/1
C    192.168.1.0/24 is directly connected, Serial0/1
R    192.168.2.0/24 [120/1] via 192.168.1.2, 00:00:07, Serial0/1
R    192.168.3.0/24 [120/2] via 192.168.1.2, 00:00:07, Serial0/1
```

Kiedy tunel zostanie ponownie uruchomiony i EIGRP zacznie działać, pętla zwrotna interfejsu zdalnego będzie widziana poprzez tunel:

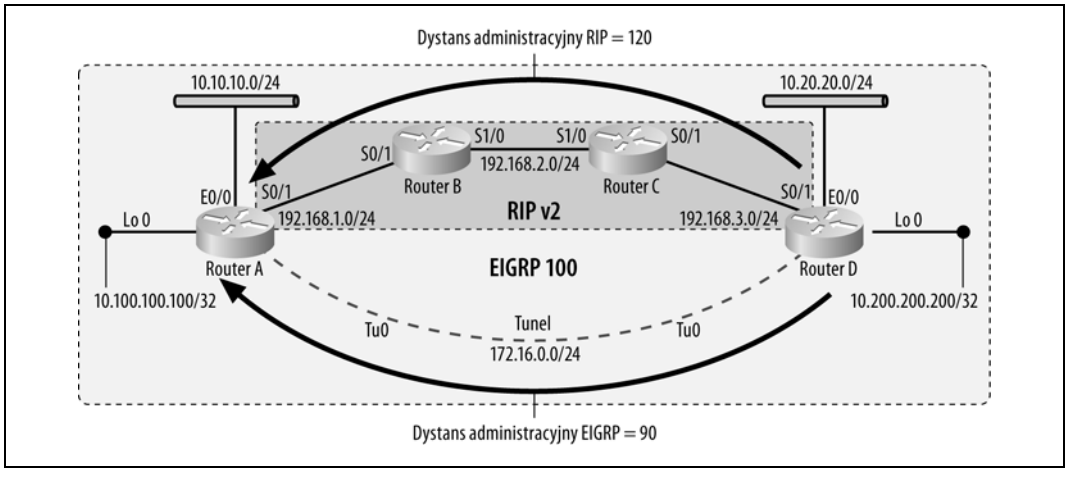
```
Router-A#sho ip route
Gateway of last resort is not set
     172.16.0.0/24 is subnetted, 1 subnets
C       172.16.0.0 is directly connected, Tunnel0
     10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
D       10.20.20.0/24 [90/297246976] via 172.16.0.2, 00:00:04, Tunnel0
C       10.10.10.0/24 is directly connected, Ethernet0/1
C       10.100.100.100/32 is directly connected, Loopback0
D       10.200.200.200/32 [90/297372416] via 172.16.0.2, 00:00:04, Tunnel0
C    192.168.1.0/24 is directly connected, Serial0/1
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
R    192.168.2.0/24 [120/1] via 192.168.1.2, 00:00:00, Serial0/1
R    192.168.3.0/24 [120/2] via 192.168.1.2, 00:00:00, Serial0/1
```

W tym momencie routery na obu końcach natychmiast wykryją problem i zamkną tunel. Trasa EIGRP zostanie utracona, a przywrócona zostanie trasa RIP. Wtedy router ponownie uruchomi tunel i cykl będzie się powtarzał w nieskończoność, dopóki coś się nie zmieni.

Powodem występowania problemu trasy rekurencyjnej jest dystans administracyjny wykorzystywanych protokołów routingu. RIP ma dystans administracyjny 120, podczas gdy EIGRP ma 90. Kiedy protokół z korzystniejszym dystansem administracyjnym uzyskuje informacje o trasie, jest ona umieszczana w tabeli routingu. Rysunek 12.4 przedstawia Router A, który uzyskuje trasy z różnych protokołów routingu.



Rysunek 12.4. Trasa EIGRP uzyskana poprzez tunel

W przypadku kiedy w tunelu działa EIGRP, problem stanowi to, że trasa do zdalnego końca tunelu uzyskiwana jest poprzez sam tunel. Tunel opiera się na trasie do interfejsu zdalnego, którą sam sobie dostarcza. Ponieważ jest to niedopuszczalne, router zamyka tunel. Niestety chwilę później tunel zostaje znowu uruchomiony, co powoduje nieustanną zmianę tras oraz niestabilne działanie tunelu.

Rozwiązanie tego problemu może być dwojakie. Można zaprzestać używania tuneli (rekomendowałbym to w tym przypadku) lub filtrować zdalną stronę tunelu tak, aby nie była dodawana do protokołu routingu, który biegnie przez tunel (EIGRP). Sprawdziłoby się również zainstalowanie VPN, ponieważ ukryłby on sieci „publiczne" przed tym, co znajduje się „wewnątrz" obu stron, łagodząc w ten sposób problem. Przyglądając się naszym konfiguracjom, zauważysz, że kłopot tkwi w tym, iż dodaliśmy sieci pętli zwrotnej do procesów EIGRP. Ich usunięcie rozwiąże nasz problem trasy rekurencyjnej:

```
Router-A (config)#router eigrp 100
Router-A(config-router)# no network 10.100.100.0 0.0.0.255
```

Oto nowa konfiguracja dla Routera A:

```
router eigrp 100
 network 10.10.10.0 0.0.0.255
 network 172.16.0.0 0.0.0.255
 no auto-summary
```

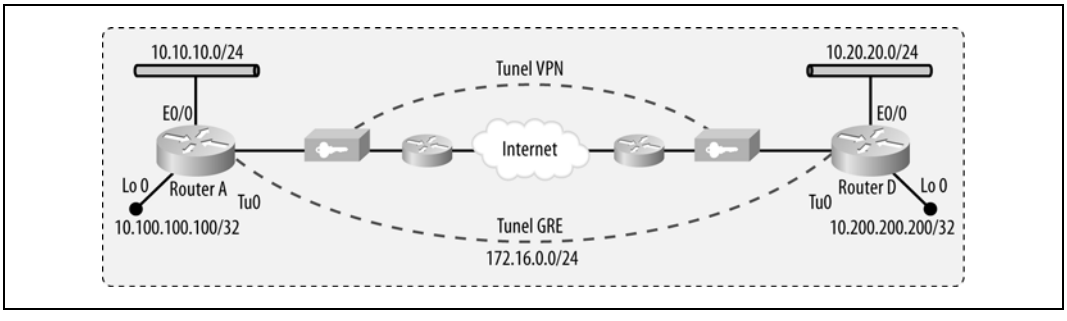

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zrobimy to samo na Routerze D dla jego sieci pętli zwrotnej. Będziemy mogli wtedy zobaczyć pożądany rezultat na Routerze A. Teraz trasa do zdalnego adresu pętli zwrotnej została pozyskana z protokołu RIP, a trasa do zdalnego Ethernetu z protokołu EIGRP:

```
Router-A#sho ip route
Gateway of last resort is not set
     172.16.0.0/24 is subnetted, 1 subnets
C       172.16.0.0 is directly connected, Tunnel0
     10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
D       10.20.20.0/24 [90/297246976] via 172.16.0.2, 00:03:23, Tunnel0
C       10.10.10.0/24 is directly connected, Ethernet0/1
C       10.100.100.100/32 is directly connected, Loopback0
R       10.200.200.200/32 [120/3] via 192.168.1.2, 00:00:11, Serial0/1
C    192.168.1.0/24 is directly connected, Serial0/1
R    192.168.2.0/24 [120/1] via 192.168.1.2, 00:00:12, Serial0/1
R    192.168.3.0/24 [120/2] via 192.168.1.2, 00:00:12, Serial0/1
```

Wykorzystanie tuneli GRE nie jest z reguły dobrym pomysłem, ponieważ komplikują one działanie sieci. Jeśli ktoś próbowałby zdiagnozować problem w sieci, którą właśnie stworzyliśmy, tunel pogłębiłby jedynie złożoność tego problemu. Zazwyczaj dodawanie do sieci tuneli GRE bez oczywistej potrzeby jest wynikiem złego planowania lub niewłaściwego projektu. Jednym z uzasadnionych powodów stosowania GRE jest routing w tunelach VPN (IPsec, powszechnie stosowany w przypadku VPN-ów, nie przekazuje pakietów multicast oraz broadcast, wymagany jest więc GRE).

Stworzenie tunelu GRE wewnątrz tunelu VPN zapewni Ci takie łącze, jakiego potrzebujesz. Dlaczego więc nie zastosować samego tunelu GRE? Pamiętaj, że GRE nie szyfruje danych. Rysunek 12.5 ilustruje typowy schemat tunelu GRE wewnątrz tunelu VPN.



Rysunek 12.5. Tunel GRE w tunelu VPN

Konfiguracja dla tego przykładu jest identyczna pod względem EIGRP oraz tunelu. Różnicą jest, że pośrodku nie ma protokołu RIP. Trasy do zdalnego końca tunelu VPN są domyślnymi trasami dla koncentratorów VPN, ponieważ będą one miały publiczny adres IP. Koncentratory VPN umożliwią odnajdywanie zdalnych adresów pętli zwrotnej routera poprzez trasy statyczne.

Jest to dość powszechne zastosowanie GRE w przypadku, gdy potrzebujesz, aby w tunelu VPN działały protokoły routingu. Nie likwiduje to jednak zagrożenia, jakim jest routing rekurencyjny. Musisz więc zadbać o to, żeby adresy interfejsów pętli zwrotnych, na których oparty jest interesujący nas tunel, nie były rozgłaszane przez EIGRP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

GRE i listy dostępu

GRE jest protokołem tej samej warstwy co TCP oraz UDP. Konfigurując firewall, aby nie blokował GRE, nie ustawiasz numerów portów, tak jak w przypadku telnetu lub SSH. Zamiast tego konfigurujesz go tak, żeby dopuszczał protokół 47. W routerach Cisco stosuje się argument `gre` w konfiguracji list dostępu:

```
R1(config)#access-list 101 permit ?
  <0-255>  An IP protocol number
  ahp      Authentication Header Protocol
  eigrp    Cisco's EIGRP routing protocol
  esp      Encapsulation Security Payload
  gre      Cisco's GRE tunneling
  icmp     Internet Control Message Protocol
  igmp     Internet Gateway Message Protocol
  igrp     Cisco's IGRP routing protocol
  ip       Any Internet Protocol
  ipinip   IP in IP tunneling
  nos      KA9Q NOS compatible IP over IP tunneling
  ospf     OSPF routing protocol
  pcp      Payload Compression Protocol
  pim      Protocol Independent Multicast
  tcp      Transmission Control Protocol
  udp      User Datagram Protocol
```

Firewalle PIX również obsługują argument `gre`:

```
PIX(config)#access-list In permit gre host 10.10.10.10 host 20.20.20.20
```

Protokół PPTP (ang. *Point-to-Point Tunneling Protocol*) wykorzystuje GRE, więc jeśli używasz PPTP do dostępu VPN, musisz również ustanowić w firewallu regułę dla GRE.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 13.

Protokoły redundancji

Wyeliminowanie wszystkich punktów awarii powinno być priorytetowym zadaniem każdego inżyniera sieciowego lub architekta przy projektowaniu sieci. Można by przyjąć proste założenie, że zdublowanie każdego urządzenia w sieci zapewni właściwą redundancję. Co jednak tak naprawdę czyni urządzenia redundantnymi?

Urządzenia takie jak firewalle lub te, które służą do równoważenia obciążenia (ang. *load balancers*), posiadają funkcje redundancji oraz odporności na problemy (ang. *fault-tolerance*) zaimplementowane w systemie operacyjnym. Posunięte jest to nawet do tego stopnia, że zmiany w konfiguracji urządzenia podstawowego (ang. *primary*) przesyłane są do urządzenia zapasowego (ang. *secondary*). Routery Cisco nie posiadają wprawdzie takiej funkcjonalności, ale nie dzieje się tak bez powodu. Mógłbyś zechcieć skonfigurować dla bramy domyślnej dwa routery w trybie *failover* (w przypadku awarii jednego drugi przejmuje jego zadania), jednak mogą mieć one podłączone różne łącza szeregowe. Może również zaistnieć taka sytuacja, że do jednego routera będzie podłączony inny usługodawca internetowy niż do drugiego. Konfiguracja routerów nie będzie więc taka sama, a co za tym idzie, synchronizacja konfiguracji nie będzie praktycznym rozwiązaniem.

W routerach potrzebujemy zazwyczaj funkcji polegającej na przejmowaniu przez jedno urządzenie zadań drugiego w danej sieci. Routery zasadniczo obsługują wiele protokołów i integrują wiele rodzajów technologii, a każda technologia może być skonfigurowana w preferowanym dla niej trybie *failover*. W przypadku Ethernetu najczęściej stosowanymi metodami są protokoły HSRP (ang. *Hot Standby Router Protocol*) oraz VRRP (ang. *Virtual Router Redundancy Protocol*). HSRP jest własnością Cisco, podczas gdy VRRP jest protokołem niezastrzeżonym, dzięki czemu jest dostępny również na urządzeniach innych producentów. Nowym rozwiązaniem w dziedzinie protokołów redundancji jest GLBP (ang. *Gateway Load Balancing Protocol*), który również stanowi własność Cisco. W tym rozdziale omówię HSRP i GLBP.

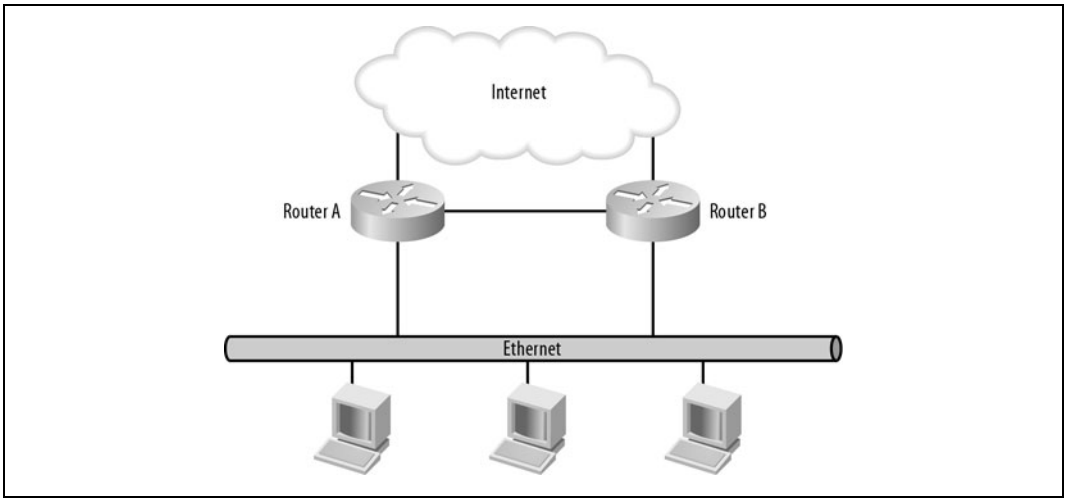
HSRP

Aby uruchomić HSRP, należy skonfigurować jeden lub kilka routerów, stosując polecenie `ip standby` na interfejsach, które utworzą grupę HSRP (ang. *HSRP group*). HSRP w najprostszej formie tworzą dwa routery, z których każdy posiada jeden interfejs w sieci. Jeden z routerów jest podstawowy, a drugi zapasowy. Jeśli router podstawowy zawiedzie, zapasowy przejmuje jego funkcje.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Szczegóły dotyczące protokołu Cisco HSRP znajdziesz w publikacji RFC 2281, zatytułowanej „Cisco Hot Standby Router Protocol (HSRP)".

Rysunek 13.1 przedstawia prosty schemat z redundantną parą routerów pełniących funkcję bramy domyślnej. W normalnym przypadku jeden z routerów byłby podstawowy, a drugi zapasowy — w terminologii HSRP powiedzielibyśmy, że jeden router jest **aktywny** (ang. *active*), a drugi **rezerwowy** (ang. *standby*). Schemat ten nie zapewnia jednak wystarczających informacji, pozwalających określić, jak zachowuje się sieć. Który router aktualnie przekazuje pakiety? A może obydwa? Który router jest podstawowy, a który zapasowy?



Rysunek 13.1. Prosty schemat HSRP

Aby skonfigurować HSRP dla tej sieci, musimy najpierw zdefiniować trzy rzeczy: adres IP interfejsu Ethernet Routera A, adres IP interfejsu Ethernet Routera B oraz wirtualny adres IP (VIP), który będzie pełnił funkcję bramy dla urządzeń w tej sieci.

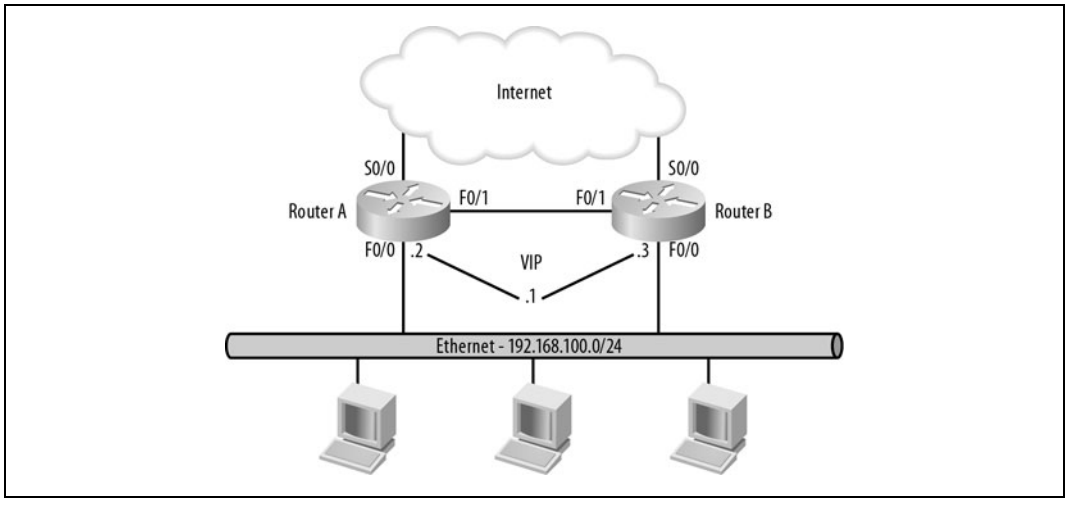
Adresy IP interfejsów routerów nie będą się zmieniać, podobnie jak VIP. Jedynym elementem, który będzie się zmieniać w przypadku awarii, jest właściciel adresu VIP. Adres wirtualny jest aktywny na tym routerze, który ma najwyższy priorytet. Domyślna wartość priorytetu wynosi 100, a jego możliwe do skonfigurowania wartości zawierają się w zakresie 0 – 255.

Wszystkie routery z tej samej grupy HSRP (domyślną grupą jest grupa 0) wysyłają pakiety HSRP na multicastowy adres 224.0.0.2, wykorzystując port UDP o numerze 1985. Wszystkie pakiety HSRP mają czas życia pakietu, czyli TTL (ang. *time-to-live*), wynoszący 1. Nie wyjdą więc poza lokalny segment Ethernetu.

Kiedy router z interfejsem skonfigurowanym w HSRP uruchamia ten interfejs (oraz za każdym razem, gdy interfejs jest aktywowany), HSRP wysyła pakiety hello i czeka, aż inny router HSRP zostanie znaleziony. Jeśli znajdzie się więcej niż jeden router HSRP, routery będą negocjować, aby ustalić, który z nich powinien zostać routerem aktywnym. Aktywnym routerem zostaje ten z najwyższym priorytetem. W przypadku remisu rozstrzyga najwyższy skonfigurowany adres IP routera.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Skonfigurujmy więc dla naszego przykładu trzy wymagane adresy IP, tak jak ilustruje to rysunek 13.2.



Rysunek 13.2. Przydział adresów IP

Możemy teraz zaimplementować najprostszą konfigurację HSRP:

Router A:

```
interface f0/0
 ip address 192.168.100.2 255.255.255.0
 standby ip 192.168.100.1
 standby preempt
```

Router B:

```
interface f0/0
 ip address 192.168.100.3 255.255.255.0
 standby ip 192.168.100.1
 standby priority 95
 standby preempt
```

Tak jak zawsze przypisaliśmy na każdym routerze adresy IP do interfejsów. Przypisaliśmy również obydwu routerom ten sam rezerwowy adres IP (ang. *standby IP address*) — jest to adres VIP.

Zauważ, że instrukcja `standby priority` dotyczy tylko Routera B. Skoro domyślny priorytet wynosi 100, to ustawiając dla Routera B priorytet 95, ustanowisz go routerem rezerwowym (95 to niższy priorytet niż 100).

Konfiguracje obu routerów zawierają polecenie `standby preempt`. Domyślnie HSRP nie ustawi podstawowego routera jako routera aktywnego, gdy powróci on do trybu online.

Aby uruchomić takie zachowanie, musisz skonfigurować routery do wywłaszczania (ang. *preempt*). Oznacza to, że kiedy router z wyższym priorytetem niż router aktywny wraca do trybu online, aktualny router aktywny ustępuje pola routerowi z wyższym priorytetem.

Komenda `show standby` służy do sprawdzenia statusu HSRP na routerach:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Router A:

```
Router-A>sho standby
FastEthernet0/0 - Group 0
  Local state is Active, priority 100, may preempt
  Hellotime 3 sec, holdtime 10 sec
  Next hello sent in 0.412
  Virtual IP address is 192.168.100.1 configured
  Active router is local
  Standby router is 192.168.100.3 expires in 7.484
  Virtual mac address is 0000.0c07.ac00
  2 state changes, last state change 23w3d
```

Router B:

```
Router-B>sho standby
FastEthernet0/0 - Group 0
  Local state is Standby, priority 95, may preempt
  Hellotime 3 sec, holdtime 10 sec
  Next hello sent in 1.398
  Virtual IP address is 192.168.100.1 configured
  Active router is 192.168.100.2 priority 100 expires in 9.540
  Standby router is local
  2 state changes, last state change 23w3d
```

Listing z Routera A wskazuje, że jest on routerem aktywnym o priorytecie 100 i ma zdolność odzyskiwania swojej pozycji. Widzimy również, że adres wirtualny to 192.168.100.1, a routerem rezerwowym jest 192.168.100.3. Przy większej liczbie routerów niż dwa bez tej informacji nie byłoby oczywiste, który router jest rezerwowy.

Wielu z nas umyka pewien istotny aspekt HSRP. Jeśli w protokole skonfigurowane są więcej niż dwa routery i routery aktywny oraz rezerwowy zostały już wybrane, pozostałe routery nie mogą przyjąć statusu aktywny lub rezerwowy, dopóki router rezerwowy nie zostanie routerem aktywnym. Publikacja RFC 2281 stwierdza:

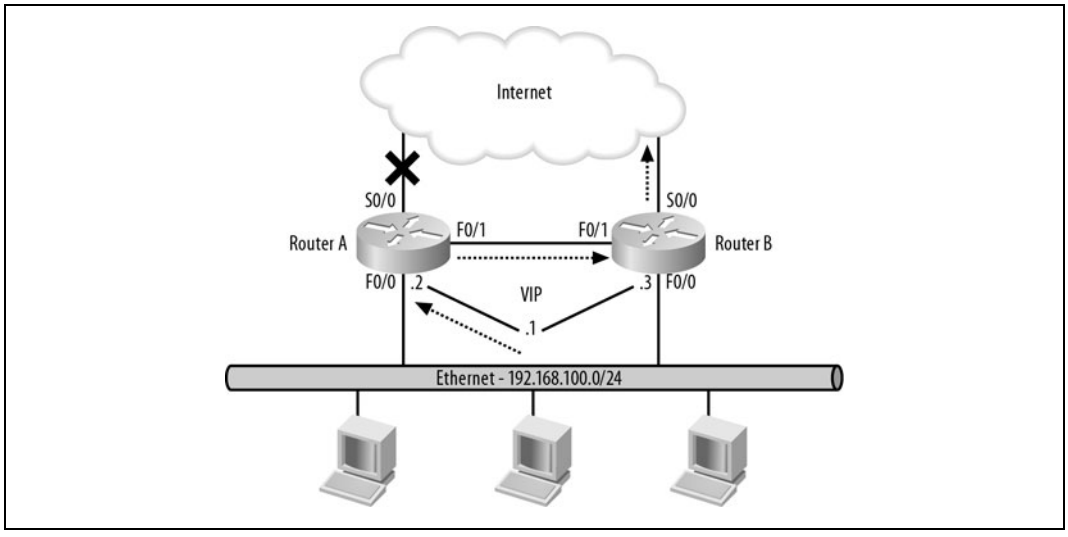
> Aby zminimalizować ruch sieciowy, od momentu, gdy protokół zakończył fazę wybierania urządzeń, tylko routery aktywny i rezerwowy wysyłają periodyczne wiadomości HSRP. Jeśli router aktywny ma awarię, jego funkcje przejmuje router rezerwowy, który staje się routerem aktywnym. Kiedy router rezerwowy ma awarię lub staje się routerem aktywnym, kolejny router jest wybierany routerem rezerwowym.

Śledzenie interfejsu HSRP

O ile HSRP jest doskonałym rozwiązaniem w przypadku awarii routera lub interfejsu łączącego z siecią lokalną, o tyle jedna z jego podstawowych funkcji rozwiązuje problemy występujące w innych obszarach. Rysunek 13.3 przedstawia problem bardziej złożony niż ten opisany poprzednio. W tym scenariuszu awarii ulega link szeregowy łączący Router A z internetem. Ponieważ zarówno router, jak i interfejsy Ethernet nadal działają, HSRP może wysyłać i odbierać pakiety hello. W takiej sytuacji Router A wciąż pozostaje routerem aktywnym.

Ponieważ sieć jest elastyczna, pakiety nadal będą miały dostęp do internetu poprzez interfejsy F0/1. Po co jednak dodawać kolejny przeskok, skoro go nie potrzebujemy? Jeśli moglibyśmy w jakikolwiek sposób wpłynąć na priorytet HSRP, opierając się na statusie innego interfejsu, to przenieślibyśmy adres wirtualny z Routera A na Router B z uwagi na status łącza S0/0. To właśnie umożliwia nam funkcja **śledzenia interfejsu HSRP** (ang. *HSRP interface tracking*).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 13.3. Awaria podstawowego łącza internetowego w przypadku braku śledzenia interfejsu

Dodając kilka prostych komend do naszej konfiguracji HSRP, możemy stworzyć schemat, w którym interfejsy Ethernet będą się zamieniać w przypadku awarii łącza szeregowego:

Router A:

```
interface f0/0
 ip address 192.168.100.2 255.255.255.0
 standby ip 192.168.100.1
 standby preempt
 standby track Serial0/0 10
```

Router B:

```
interface f0/0
 ip address 192.168.100.3 255.255.255.0
 standby ip 192.168.100.1
 standby priority 95
 standby preempt
 standby track Serial0/0 10
```

Na każdym routerze dodaliśmy komendę `standby track Serial0/0 10` dla interfejsu Ethernet. Komenda ta powoduje, że HSRP obniża o 10 wartość priorytetu interfejsu Ethernet w przypadku wyłączenia interfejsu Serial0/0.

Widziałem wiele sieci, w których jeden router ma priorytet 100, a drugi priorytet 90. Kiedy zawiedzie śledzony interfejs na routerze podstawowym, spowoduje to powstanie sytuacji remisowej, w której IOS musi wybrać router aktywny na podstawie najwyższego skonfigurowanego adresu IP. Choć może się wydawać, że nie jest to problem w przypadku jedynie dwóch routerów, to jednak w tych okolicznościach ruch w sieci może być kierowany nie tam, gdzie byś tego oczekiwał.

Konfigurowanie wartości obniżania priorytetu jest bardzo przydatną funkcją. Jeśli każdy router posiadałby trzy łącza do internetu, mógłbyś np. zmniejszyć priorytet o 3 dla każdego interfejsu. W naszym przykładzie, jeśli zawiódłby jeden link, Router A pozostałby routerem aktywnym. Jednak jeśli zawiodłyby dwa linki szeregowe, priorytet routera zostałby obniżony łącznie o 6,

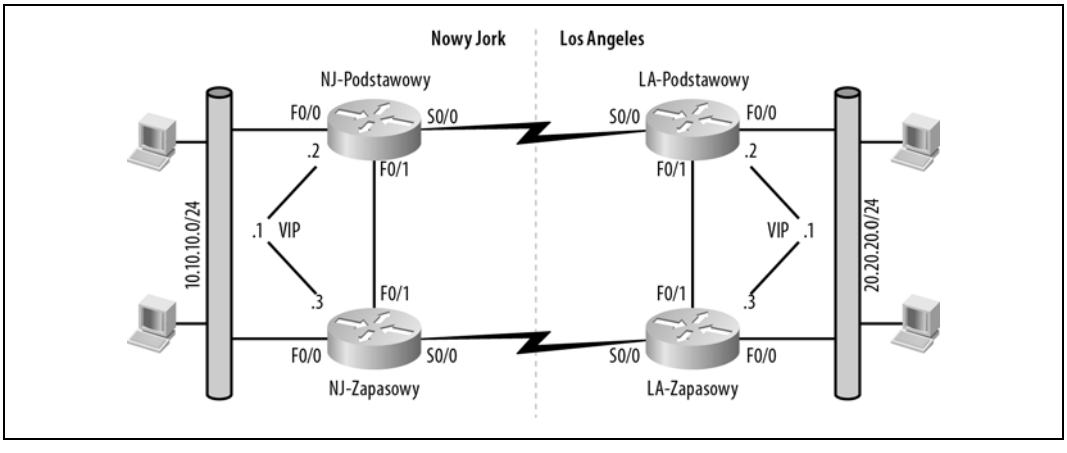
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

więc wyniósłby 94. Router A miałby wtedy niższy priorytet od Routera B, którego priorytet to 95. W tej sytuacji routerem aktywnym zostałby Router B. Jeśli jednak w tym przykładzie chcielibyśmy, aby routerem aktywnym zawsze był router z największą liczbą działających linków szeregowych, zwiększmy z 3 na 6 wartość, o jaką będzie spadał priorytet (oczywiście awaria routera lub interfejsu Ethernet wpływa na pozostałe routery w taki sam sposób, jaki opisałem w podstawowym przykładzie).

Kiedy HSRP nie wystarcza

HSRP jest niesamowitym narzędziem, a połączony z funkcją śledzenia interfejsu może być środkiem do uzyskania niemal całkowitej redundancji. Są jednak sytuacje, w których protokół HSRP nie jest wystarczający. Przykład, który tutaj przedstawię, jest jednym z moich ulubionych pytań podczas rozmów kwalifikacyjnych, ponieważ tylko osoba z praktycznym doświadczeniem w złożonych sieciach mogła się z tym spotkać.

Rysunek 13.4 przedstawia pozornie prostą konfigurację HSRP. Dwie lokalizacje, Nowy Jork i Los Angeles, są połączone poprzez dwie sieci T1. Routery po każdej ze stron połączone są interfejsami F0/1, a HSRP jest zaimplementowane ze śledzeniem interfejsu na interfejsach F0/0. Koncepcja zakłada, że w przypadku awarii któregoś z routerów podstawowych router zapasowy przejmuje jego funkcje. Ponadto jeśli zawiedzie podstawowy link T1, wykorzystany zostanie link zapasowy, ponieważ włączone jest śledzenie interfejsu.



Rysunek 13.4. Scenariusz failover dla dwóch linków z wykorzystaniem HSRP

Poniżej podstawowe konfiguracje Ethernet dla każdego routera:

NJ-Podstawowy:

```
interface f0/0
 ip address 10.10.10.2 255.255.255.0
 standby ip 10.10.10.1
 standby preempt
 standby track Serial0/0 10
```

NJ-Zapasowy:

```
interface f0/0
 ip address 10.10.10.3 255.255.255.0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
 standby ip 10.10.10.1
 standby priority 95
 standby preempt
 standby track Serial0/0 10
```

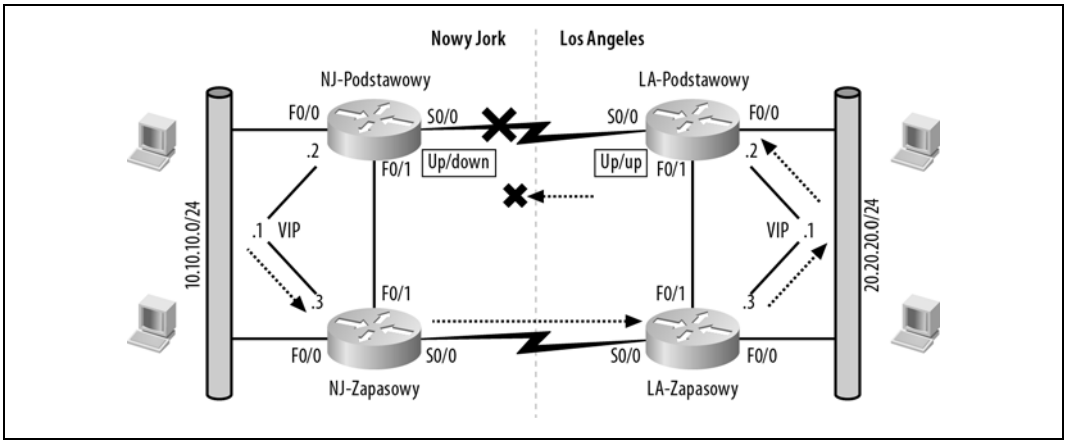
LA-Podstawowy:

```
interface f0/0
 ip address 20.20.20.2 255.255.255.0
 standby ip 20.20.20.1
 standby preempt
 standby track Serial0/0 10
```

LA-Zapasowy

```
interface f0/0
 ip address 20.20.20.3 255.255.255.0
 standby ip 20.20.20.1
 standby priority 95
 standby preempt
 standby track Serial0/0 10
```

W przypadku całkowitej awarii linku T1 łączącego NJ-Podstawowy z LA-Podstawowym routery NJ i LA rozpoznają problem i uruchomione zostaną routery zapasowe. W rzeczywistości jednak problemy bywają bardziej złożone niż te z rozważań teoretycznych i ten schemat nie funkcjonuje tak dobrze, jak byśmy chcieli. Rysunek 13.5 pokazuje, co może pójść źle.



Rysunek 13.5. Ograniczenia HSRP

Załóżmy, że łącze pomiędzy Nowym Jorkiem i Los Angeles cierpi z powodu częściowej awarii. Coś spowodowało, że interfejs szeregowy podstawowego routera NJ przeszedł w stan up/down (włączony/wyłączony), natomiast interfejs szeregowy routera podstawowego LA pozostał w stanie up/up (włączony/włączony). Spotkałem się z tym nie jeden raz, w różnych rodzajach układów.

HSRP reaguje na wyłączenie interfejsu po stronie sieci w Nowym Jorku ustanowieniem zapasowego routera NJ routerem aktywnym. Dzieje się tak, ponieważ interfejs szeregowy na routerze podstawowym NJ jest śledzony. Pakiety są przekazywane do routera zapasowego NJ, a potem dalej przez T1 do routera zapasowego LA, który przekazuje je do ich miejsc docelowych. Z pakietami zwrotnymi jest jednak problem. Ponieważ router podstawowy LA nie rozpoznał awarii podstawowego linku T1, pozostał routerem aktywnym. Pakiety zwrotne są wysyłane

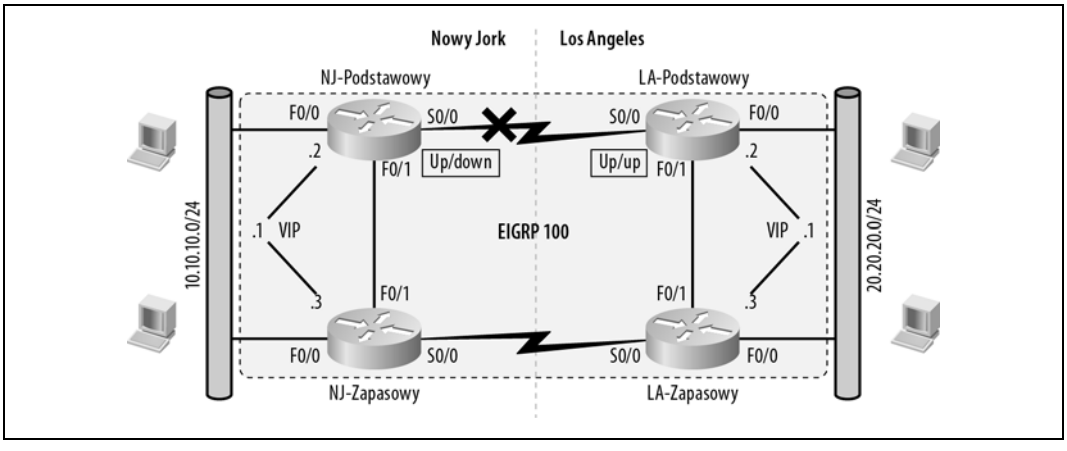
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Obszar MAE (ang. *Metropolitan Area Ethernet*), czyli inaczej Metro-E, jest podatny na występowanie tego typu sytuacji. Ponieważ link Metro-E opiera się z reguły na technologii transportu danych SONET[1] konwertowanej na Ethernet, integralność łącza jest lokalna po każdej ze stron. Jeśli odepniesz jedną stronę obwodu Metro-E, druga oddalona strona w przypadku większości instalacji nie zostanie wyłączona.

do routera podstawowego LA, a ponieważ router ten przyjmuje, że link jest nadal aktywny, przekazuje pakiety dalej przez interfejs S0/0. Pakiety te giną, gdyż druga strona linku jest wyłączona.

Bardziej niezawodnym rozwiązaniem niż metoda linku awaryjnego (ang. *link-failover*) jest zastosowanie protokołu bramy wewnętrznej na wszystkich routerach. Protokół taki jak OSPF czy EIGRP ustanawia relacje przylegania dla sąsiadów po obu stronach linku. Kiedy link zawodzi, protokół routingu „wie", że oddalony sąsiad jest nieosiągalny, więc usuwa link z tabeli routingu.

Takie rozwiązanie wykorzystuje więc oprócz HSRP również protokół routingu. Rysunek 13.6 przedstawia tę samą sieć, ale wzbogaconą o protokół EIGRP. Teraz, kiedy zawiedzie router podstawowy po stronie NJ na linku podstawowym, EIGRP traci relację przylegania pomiędzy routerem podstawowym NJ i routerem podstawowym LA, a link usuwany jest z tabel routingu na wszystkich routerach. Ponieważ EIGRP modyfikuje tabele routingu, może znaleźć trasę omijającą wyłączony link, pomimo że jedna strona raportuje status up/up. Oczywiście HSRP jest w tym przykładzie nadal wymagane, ponieważ EIGRP nie posiada funkcji umożliwiającej to, aby dwa routery zachowywały się jak pojedyncza brama w lokalnych segmentach Ethernet.



Rysunek 13.6. Poprawiony schemat failover z wykorzystaniem EIGRP

Nexus i HSRP

Tak jak w przypadku dużej części protokołów dla urządzeń Nexus musisz na początku uruchomić funkcję HSRP poleceniem `feature hsrp`.

[1] Jest to standard sieci światłowodowych stosowany w Ameryce Północnej. Jego europejskim odpowiednikiem jest Synchroniczna Hierarchia Cyfrowa (ang. *Synchronous Digital Hierarchy — SDH*) — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

NX-OS dodaje pewną funkcjonalność do HSRP, jeśli korzystasz z kanałów vPC. W skrócie, kiedy pakiet jest odbierany na porcie rutującym (a nie przełączającym) i jest skierowany do adresu wirtualnego HSRP skonfigurowanego na routerze, pakiet zakończy swoją drogę na tym routerze, nawet jeśli nie jest to aktywny router HSRP. Pakiety otrzymywane na interfejsach warstwy drugiej będą przekazywane do aktywnego routera.

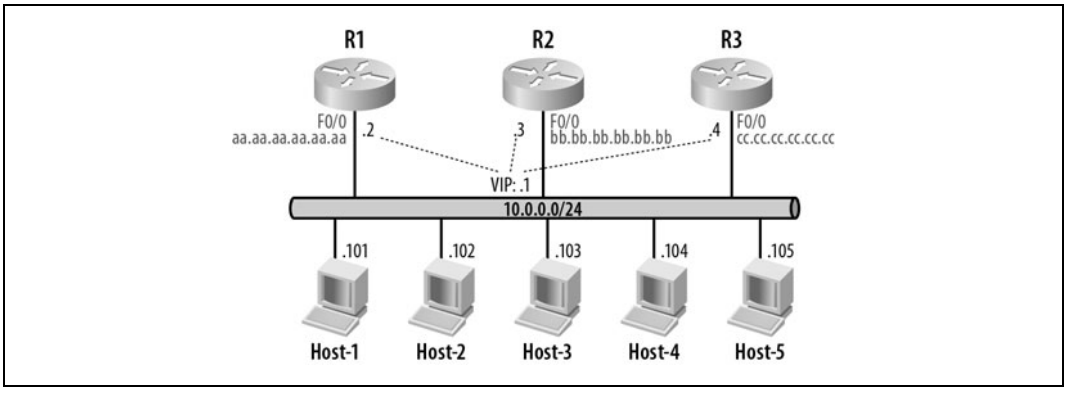
Czemu jest to ważne? Przede wszystkim pomoże to wyeliminować ruch na łączach pomiędzy routerami. Po drugie, jeśli nie jesteś świadomy takiego zachowania, możesz dostać szału, próbując zrozumieć, gdzie płyną pakiety, kiedy myślałeś, że wszystko działa w tradycyjny sposób. Jest to ostrzeżenie.

Kanał vPC pomiędzy dwoma przełącznikami Nexus będzie przepuszczał ruch zarówno przez routery aktywne, jak i routery zapasowe. Dokumentacja Cisco zaleca, aby aktywny router HSRP był tym samym urządzeniem co podstawowe urządzenie w parze vPC.

GLBP

GLBP (ang. *Gateway Load Balancing Protocol*) jest protokołem zbliżonym do HSRP i VRRP, posiada jednak dodatkowo funkcję równoważenia obciążenia pomiędzy członkami grupy. W przypadku HSRP oraz VRRP mamy do czynienia tylko z jednym aktywnym członkiem grupy, podczas gdy reszta urządzeń nie transmituje danych. W GLBP każdy członek grupy może przekazywać ruch sieciowy i ruch ten może być odpowiednio rozkładany między urządzeniami. Rozkładanie ruchu między członków grupy umożliwi Ci np. posiadanie trzech aktywnych bram, z których jedna przekazuje 50% ruchu, a pozostałe dwie po 25%. Na interfejsie routera mogą być maksymalnie 1024 grupy GLBP, z których każda może posiadać do czterech wirtualnych spedytorów (ang. *forwarders*). Przyjrzyjmy się przykładowi.

Rysunek 13.7 przedstawia prostą sieć, w której znajduje się pięć hostów i trzy routery. Każdy z trzech routerów należy do GLBP, więc ruch będzie aktywnie przekazywany przez wszystkie trzy. Skonfigurujemy obciążenia 50% dla R1 oraz po 25% dla R2 i R3.



Rysunek 13.7. Przykładowa sieć GLBP

Bramy skonfigurowane w GLBP zwane są **spedytorami** (ang. *forwarders*). Każdy router, który przekazuje pakiety, określany jest mianem aktywnego wirtualnego spedytora (ang. *Active Virtual Forwarder* — AVF). Jeden z routerów wybierany jest aktywną bramą wirtualną (ang. *Active Virtual Gateway* — AVG). AVG jest odpowiedzialna za kontrolę grupy i zasadniczo

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

stanowi mózg operacji. AVG kontroluje proces rozkładania obciążenia, odpowiada na zapytania ARP oraz przydziela wirtualne adresy MAC routerom AVF. Uważaj tutaj na terminologię, gdyż określenie **aktywna** (ang. *active*) oznacza co innego niż w HSRP.

AVG w grupie GLBP jest wybierana na podstawie priorytetów. Priorytet AVG może być wywłaszczony, co oznacza, że członek o wyższym priorytecie będzie przejmował kontrolę, gdy zostanie uruchomiony i przejdzie do trybu online.

Równoważenie obciążenia może się odbywać z wykorzystaniem algorytmów: *round-robin* (karuzela), *weighted load balancing* (ważone rozłożenie obciążenia) oraz *host-dependent* (zależny od hosta). *Weighted load balancing* zastosujemy, jeśli będziemy chcieli rozłożyć obciążenie ruchem procentowo. Metoda *host-dependent* zakłada, że ten sam wirtualny adres MAC jest zawsze dostarczany do tego samego adresu MAC hosta. Dzięki temu host zawsze korzysta z tej samej fizycznie bramy, chyba że brama ta jest offline. Jeśli nie zdefiniowano żadnego sposobu rozłożenia obciążenia, GLBP zachowuje się jak HSRP, a router AVG funkcjonuje jako jedyny spedytor.

Używanie sformułowań takich jak „procentowy udział w ruchu sieciowym" oraz „rozłożenie obciążenia" może być użyteczne, zwłaszcza jeśli rozmawiasz z przełożonymi. Są to jednak zwroty technicznie niedokładne. GLBP nie rozkłada obciążenia ruchem, ale raczej **udziały** (ang. *load-shares*), decydując o tym, która brama jest ogłaszana, gdy pojawia się zapytanie ARP. Dlaczego to rozróżnienie jest takie ważne? Załóżmy, że Host-5 z rysunku 13.7 jest korporacyjnym serwerem pocztowym. Jeśli wszystkie pozostałe hosty są po prostu stacjami roboczymi, jest bardzo prawdopodobne, że bez względu na sposób dystrybucji wirtualnych adresów MAC przez GLBP serwer pocztowy będzie pochłaniał większość zasobów. Można na to spojrzeć jeszcze inaczej: załóżmy, że R1 będzie domyślną bramą dla 50% hostów. Jeśli jednak bramą przypisaną do serwera pocztowego jest R3, to prawdopodobnie właśnie R3 będzie obsługiwał 99% ruchu sieciowego.

Przyjrzyjmy się proponowanej konfiguracji dla naszego przykładu. Poniżej podstawowe konfiguracje GLBP dla trzech routerów:

Konfiguracja R1:

```
interface FastEthernet0/0
 ip address 10.0.0.2 255.255.255.0
 glbp 10 ip 10.0.0.1
 glbp 10 preempt
 glbp 10 weighting 50
 glbp 10 load-balancing weighted
```

Bez skonfigurowania priorytetu R1 przyjmuje domyślną wartość 100.

Konfiguracja R2:

```
interface FastEthernet0/0
 ip address 10.0.0.3 255.255.255.0
 glbp 10 ip 10.0.0.1
 glbp 10 priority 50
 glbp 10 preempt
 glbp 10 weighting 25
 glbp 10 load-balancing weighted
```

Konfiguracja R3:

```
interface Ethernet0/0
 ip address 10.0.0.4 255.255.255.0
 glbp 10 ip 10.0.0.1
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
glbp 10 priority 25
glbp 10 preempt
glbp 10 weighting 25
glbp 10 load-balancing weighted
```

W przypadku urządzeń Nexus musisz najpierw uruchomić funkcję GLBP poleceniem `feature glbp`.

Metoda ważenia (ang. *weighting*) jest wykorzystywana zarówno do dystrybucji obciążenia, jak i do śledzenia obiektu (ang. *object tracking*). To ostatnie opiszę w następnym podrozdziale. Wartości *weighting* mogą zawierać się w przedziale od 1 do 254. W jaki sposób router rozkłada obciążenie, jeśli suma wartości *weighting* w routerach jednej grupy jest inna niż 100?

Oto, co można znaleźć na ten temat w dokumentacji Cisco:

> Wielkość obciążenia dla określonego AVF zależy od wartości *weighting* rozgłaszanej przez bramę, do której należy AVF.
>
> Każdy wirtualny spedytor w bramie stosuje bieżącą wartość *weighting* bramy, bez względu na to, jak wiele wirtualnych spedytorów jest aktywnych w danej bramie.

oraz:

> Jest to zdolność GLBP do określania wagi dla każdego urządzenia przy obliczaniu podziału obciążenia, które nastąpi wraz z przydziałem adresu MAC. Każdy router GLBP w grupie będzie ogłaszał swoją wartość *weighting* wraz z przydzielonym adresem MAC. Na tej podstawie działać będzie AVG.
>
> Załóżmy, że mamy w grupie dwa routery: A i B. Router A posiada zdolność przekazywania dwukrotnie większej liczby pakietów niż router B. Wartość *weighting* routera A powinna być więc dwukrotnie większa niż wartość *weighting* routera B.

Obciążenie wyznacza się, dzieląc indywidualne wagi routerów przez sumę wag wszystkich routerów w grupie. Przykładowo, jeśli skonfigurujesz trzy routery, przypisując im wartości wagi odpowiednio 60/30/30, to otrzymasz następującą dystrybucję obciążenia:

R1: 60 / (60+30+30) = 0,50 (50%)

R2: 30 / (60+30+30) = 0,25 (25%)

R3: 30 / (60+30+30) = 0,25 (25%)

Jeśli nie zamierzasz korzystać z algorytmu *weighted load balancing*, zalecam stosowanie możliwie najprostszych wartości. Jak zobaczysz za chwilę, wartości te mogą być zmieniane przez funkcję śledzenia obiektów, więc w Twoim interesie jest, aby były one jak najmniej skomplikowane.

Jak widzisz, konfiguracja jest bardzo zbliżona do konfiguracji HSRP. Pewne różnice dotyczą aktywnych interfejsów. Oto listing z polecenia `show glbp` na R1:

```
R1#sho glbp
FastEthernet0/0 - Group 10
  State is Active
    8 state changes, last state change 00:09:19
  Virtual IP address is 10.0.0.1
  Hello time 3 sec, hold time 10 sec
    Next hello sent in 1.496 secs
  Redirect time 600 sec, forwarder time-out 14400 sec
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Preemption enabled, min delay 0 sec
Active is local
Standby is 10.0.0.3, priority 50 (expires in 9.496 sec)
Priority 100 (default)
Weighting 50 (configured 50), thresholds: lower 1, upper 50
Load balancing: weighted
Group members:
  0012.43b9.2020 (10.0.0.4)
  0012.43b9.2c70 (10.0.0.2) local
  001c.588b.f6a8 (10.0.0.3)
There are 3 forwarders (1 active)
Forwarder 1
  State is Listen
    8 state changes, last state change 00:00:45
  MAC address is 0007.b400.0a01 (learnt)
  Owner ID is 001c.588b.f6a8
  Redirection enabled, 597.840 sec remaining (maximum 600 sec)
  Time to live: 14397.840 sec (maximum 14400 sec)
  Preemption enabled, min delay 30 sec
  Active is 10.0.0.3 (primary), weighting 25 (expires in 7.840 sec)
Forwarder 2
  State is Active
    1 state change, last state change 00:00:44
  MAC address is 0007.b400.0a02 (default)
  Owner ID is 0012.43b9.2c70
  Redirection enabled
  Preemption enabled, min delay 30 sec
  Active is local, weighting 50
Forwarder 3
  State is Listen
  MAC address is 0007.b400.0a03 (learnt)
  Owner ID is 0012.43b9.2020
  Redirection enabled, 599.836 sec remaining (maximum 600 sec)
  Time to live: 14399.836 sec (maximum 14400 sec)
  Preemption enabled, min delay 30 sec
  Active is 10.0.0.4 (primary), weighting 25 (expires in 9.836 sec)
```

GLBP jest użytecznym protokołem, jednak uważam, że jest mocno frustrujący w przypadku diagnozowania problemów. Popatrz na paragraf dotyczący aktywnego spedytora:

```
Forwarder 2
  State is Active
    1 state change, last state change 00:00:44
  MAC address is 0007.b400.0a02 (default)
  Owner ID is 0012.43b9.2c70
  Redirection enabled
  Preemption enabled, min delay 30 sec
  Active is local, weighting 50
```

Możemy stwierdzić, że spedytor 2 jest aktywny, ale nie wiemy, który router jest tym spedytorem. Nie przypisaliśmy im numerów porządkowych 1, 2, 3, jednak zrobił to za nas router. GLBP podaje nam adres MAC właściciela spedytora, możemy więc porównać go z listą członków grupy umieszczoną na wcześniejszym listingu:

```
Group members:
  0012.43b9.2020 (10.0.0.4)
  0012.43b9.2c70 (10.0.0.2) local
  001c.588b.f6a8 (10.0.0.3)
```

Możesz pomyśleć, że się czepiam, ale i tak uważam, iż ta informacja powinna być zawarta w paragrafie dotyczącym spedytora. Jednak odkładając nerwy na bok, GLBP jest nieocenione,

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

jeśli chcesz korzystać z więcej niż jednej bramy naraz, posiadając przy tym pojedynczy adres bramy wirtualnej.

Kolejną przyjemną komendą GLBP jest `show glbp brief`, która umożliwia szybki przegląd konfiguracji i sposobu działania GLBP:

```
R1#sho glbp brief
Interface   Grp  Fwd Pri State    Address         Active router   Standby router
Fa0/0       10   -   100 Active   10.0.0.1        local           10.0.0.3
Fa0/0       10   1   -   Listen   0007.b400.0a01  10.0.0.3        -
Fa0/0       10   2   -   Active   0007.b400.0a02  local           -
Fa0/0       10   3   -   Listen   0007.b400.0a03  10.0.0.4        -
```

Śledzenie obiektu w GLBP

Podobnie jak w przypadku śledzenia interfejsu HSRP można skonfigurować GLBP do śledzenia statusu drugiego interfejsu. Jest nawet lepiej, gdyż GLBP wykorzystuje funkcję zwaną **śledzeniem obiektu** (ang. *object tracking*). Ma ona większe możliwości niż tylko śledzenie prostego statusu up/down.

Na początek musisz skonfigurować obiekt, który ma być śledzony. Poniżej skonfiguruję nowy obiekt o numerze 2, którym jest status protokołu linii na interfejsie F0/1:

```
R1(config)#track 2 interface FastEthernet0/1 line-protocol
```

Śledzenie obiektu jest potężnym narzędziem, które umożliwia więcej niż tylko śledzenie statusu interfejsu. Przykładowo możesz wykorzystać je do śledzenia czasu odpowiedzi dla łącza, stosując RTR (ang. *Response Time Reporter*), a następnie wyłączać bramę w GLBP, kiedy czas odpowiedzi jest zbyt długi. Możesz też użyć śledzenia obiektu dla wielu innych protokołów niż GLBP. Bądź jednak ostrożny. Możesz niezamierzenie skomplikować konfigurację routera do tego stopnia, że ktoś inny będzie miał spore problemy, by zorientować się, co się dzieje, gdy wystąpi jakiś problem.

Kiedy już zdefiniowaliśmy obiekt, możemy oznaczyć go parametrem `glbp weighting`:

```
R1(config-if)# glbp 10 weighting track 2 decrement 10
```

Teraz, gdy obiekt 2 (protokół linii) zostanie wyłączony, jego waga pomniejszy się o 10. Ale to nie wszystko. Możemy śledzić więcej niż jeden obiekt. Dodajmy kolejny, wskazując tym razem interfejs szeregowy:

```
R1(config)#track 3 interface s0/0/0 line-protocol
```

Dodajmy go teraz do grupy interfejsów GLBP:

```
R1(config-if)# glbp 10 weighting track 3 decrement 10
```

Oto dotychczasowa konfiguracja dla f0/0:

```
interface FastEthernet0/0
 ip address 10.0.0.2 255.255.255.0
 glbp 10 ip 10.0.0.1
 glbp 10 preempt
 glbp 10 weighting 50
 glbp 10 load-balancing weighted
 glbp 10 weighting track 2 decrement 10
 glbp 10 weighting track 3 decrement 10
end
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

A to nasze śledzone obiekty:

```
track 2 interface FastEthernet0/1 line-protocol
track 3 interface Serial0/0/0 line-protocol
```

Teraz w przypadku awarii protokołu linii dla interfejsu f0/1 lub s0/0/0 brama GLBP obniży ich wagę o 10. Jeśli wyłączone zostaną obydwa jednocześnie, waga spadnie o 20. Ale przecież to nie HSRP, więc obniżenie wagi zmniejszy jedynie natężenie ruchu na tym routerze. Żeby wszystko zadziałało tak, jak chcemy, musimy poinstruować GLBP, że w przypadku spadku wartości wagi poniżej pewnego poziomu ten router nie będzie dłużej spedytorem. Zrobimy to poprzez dodanie argumentów `lower` (mniejszy) oraz `upper` (większy) do polecenia `glbp weighting`:

```
R1(config-if)# glbp 10 weighting 50 lower 35 upper 45
```

Teraz kiedy waga spadnie poniżej 35, GLBP przestanie korzystać z tego spedytora. Bądź jednak ostrożny, gdyż skonfigurowałem również drugi warunek. Jeśli tylko jeden z interfejsów powróci do trybu online, GLBP nie uruchomi tego spedytora, ponieważ waga nie wzrośnie powyżej wartości 45.

Spójrz na listing polecenia `show glbp`, po tym jak przywrócony do trybu online został jeden z dwóch śledzonych interfejsów, które miały awarię:

```
R1#sho glbp
FastEthernet0/0 - Group 10
  State is Active
    8 state changes, last state change 21:47:52
  Virtual IP address is 10.0.0.1
  Hello time 3 sec, hold time 10 sec
    Next hello sent in 1.756 secs
  Redirect time 600 sec, forwarder time-out 14400 sec
  Preemption enabled, min delay 0 sec
  Active is local
  Standby is 10.0.0.3, priority 50 (expires in 8.932 sec)
  Priority 100 (default)
  Weighting 40, low (configured 50), thresholds: lower 35, upper 45
    Track object 2 state Up decrement 10
    Track object 3 state Down decrement 10
  Load balancing: weighted
[fragment usunięty]
  Forwarder 2
    State is Listen
      6 state changes, last state change 00:05:40
    MAC address is 0007.b400.0a02 (default)
    Owner ID is 0012.43b9.2c70
    Redirection enabled
    Preemption enabled, min delay 30 sec
    Active is 10.0.0.4 (secondary), weighting 25 (expires in 8.816 sec)
```

Jeśli chcielibyśmy, żeby GLBP uruchomił spedytora nawet wtedy, kiedy tylko jeden ze śledzonych interfejsów powróci do stanu online, należałoby ustawić wartość 35 dla obydwu argumentów:

```
R1(config-if)# glbp 10 weighting 50 lower 35 upper 35
```

Teraz po przywróceniu tylko jednego z dwóch linków, które uległy awarii, nasza brama jest aktywnym spedytorem:

```
R1#sho glbp
FastEthernet0/0 - Group 10
  State is Active
    11 state changes, last state change 00:08:52
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  Virtual IP address is 10.0.0.1
  Hello time 3 sec, hold time 10 sec
    Next hello sent in 1.200 secs
  Redirect time 600 sec, forwarder time-out 14400 sec
  Preemption enabled, min delay 0 sec
  Active is local
  Standby is 10.0.0.3, priority 50 (expires in 9.212 sec)
  Priority 100 (default)
  Weighting 40 (configured 50), thresholds: lower 35, upper 35
    Track object 2 state Up decrement 10
    Track object 3 state Down decrement 10
  Load balancing: weighted
  Group members:
    0012.43b9.2020 (10.0.0.4)
    0012.43b9.2c70 (10.0.0.2) local
    001c.588b.f6a8 (10.0.0.3)
  There are 3 forwarders (1 active)
[fragment usunięty]
Forwarder 2
   State is Active
     7 state changes, last state change 00:00:06
   MAC address is 0007.b400.0a02 (default)
   Owner ID is 0012.43b9.2c70
   Redirection enabled
   Preemption enabled, min delay 30 sec
   Active is local, weighting 40
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 14.

Mapy routingu

Mapy routingu (ang. *route maps*) są zmorą wielu osób przygotowujących się do egzaminów certyfikacyjnych. Myślę, że powodem tego jest sposób, w jaki mapy routingu są skonstruowane. Są one rozwiązaniem nieco przestarzałym w porównaniu do bardziej popularnych funkcji, takich jak listy dostępu (ang. *access lists*). Dlaczego uważam mapy routingu za przestarzałe? Przyjrzyjmy się temu.

Lista dostępu podaje funkcję dla każdego wpisu bezpośrednio w tym wpisie. Dla przykładu poniższa linia poleceń dopuszcza pakiety IP z dowolnego źródła do dowolnego miejsca przeznaczenia:

```
access-list 101 permit ip any any
```

Składnia jest nieskomplikowana i dodatkowo samodokumentująca się. Lista dostępu 101 dopuszcza pakiety IP z dowolnego źródła do dowolnego miejsca. Proste!

Z kolei mapa routingu pozwalająca osiągnąć ten sam cel mogłaby wyglądać następująco:

```
route-map GAD permit 10
 match ip address 101
```

Aby określić, do czego służy mapa routingu, musisz najpierw sprawdzić, jaka funkcja zawarta jest w liście dostępu 101, a następnie zorientować się, w jaki sposób mapa routingu ją stosuje. Przedstawiona mapa routingu również dopuszcza pakiety IP z dowolnego źródła do dowolnego miejsca przeznaczenia, ale w przeciwieństwie do powyższej listy dostępu jej intencja nie jest tak oczywista.

Po co dodawać mapę dostępu do wystarczająco prostej listy dostępu? Po pierwsze, są instancje, w których listy dostępu nie są bezpośrednio dostępne. Przykładowo BGP wykorzystuje mapy routingu i w wielu przypadkach nie dopuszcza bezpośredniego zastosowania list dostępu. Po drugie, mapy routingu są bardziej elastyczne niż listy dostępu. Pozwalają Ci tworzyć reguły dla sporej liczby rzeczy, których nie obsługują listy dostępu:

```
R1(config)#route-map GAD permit 10
R1(config-route-map)# match ?
  as-path       Match BGP AS path list //Dopasowanie na podstawie listy AS path
  clns          CLNS information //Dopasowanie na podstawie informacji CLNS
  community     Match BGP community list //Dopasowanie na podstawie list community BGP
  extcommunity  Match BGP/VPN extended community list
                                          //Dopasowanie na podstawie rozszerzonych list community BGP/VPN
  interface     Match first hop interface of route //Dopasowanie na podstawie interfejsu pierwszego skoku
  ip            IP specific information //Dopasowanie na podstawie informacji dotyczących IP
  length        Packet length //Dopasowanie na podstawie długości pakietu
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  metric        Match metric of route //Dopasowanie na podstawie metryki ścieżki
  route-type    Match route-type of route //Dopasowanie na podstawie typu ścieżki
  tag           Match tag of route //Dopasowanie na podstawie znacznika ścieżki
```

Mapy routingu są szczególnie przydatne w protokołach routingu. Wykorzystując je, możesz dokonywać filtrowania na podstawie typów tras, znaczników tras, prefiksów, rozmiarów pakietu, a nawet źródła lub następnego przeskoku pakietu.

Mapy routingu mogą również zmieniać pakiety, podczas gdy listy dostępu nie mają takiej możliwości. Polecenie `set` mapy routingu pozwala Ci zmieniać różnego rodzaju elementy pakietu, który jest wysyłany. Możesz zmieniać interfejs, do którego pakiet jest wysyłany, następny przeskok pakietu oraz wartości QoS (ang. *quality of service*), takie jak priorytet *IP precedence*:

```
R1(config-route-map)#set ?
  as-path           Prepend string for a BGP AS-path attribute //Dopisuje do AS-path zadaną wartość
  automatic-tag     Automatically compute TAG value //Automatyczne przekazanie informacji o znaczniku scieżki
  clns              OSI summary address //Adresy sumaryczne OSI
  comm-list         Set BGP community list (for deletion) //Ustawia w BGP listę community (do usunięcia)
  community         BGP community attribute //Ustawia atrybut community w BGP
  dampening         Set BGP route flap dampening parameters
                                  //Ustawia parametry mechanizmu chroniącego przed niestabilnymi ścieżkami w BGP
  default           Set default information //Ustawia informację domyślną
  extcommunity      BGP extended community attribute //Ustawia rozszerzony atrybut community w BGP
  interface         Output interface //Interfejs wyjścia
  ip                IP specific information //Informacje odnoszące się do IP
  level             Where to import route //Gdzie zaimportować trasę
  local-preference  BGP local preference path attribute //Atrybut local-preference ścieżki BGP
  metric            Metric value for destination routing protocol
                                  //Wartość metryki dla docelowego protokołu routingu
  metric-type       Type of metric for destination routing protocol
                                  //Typ metryki dla docelowego protokołu routingu
  origin            BGP origin code //Kod źródła BGP
  tag               Tag value for destination routing protocol //Wartość tag dla docelowego protokołu routingu
  weight            BGP weight for routing table //Atrybut weight w tabeli routingu
```

Elementy charakterystyczne dla IP mogą być modyfikowane, jeśli są dostępne w kategorii `ip`:

```
R1(config-route-map)#set ip ?
  default     Set default information //Ustawia informację domyślną
  df          Set DF bit //Ustawia bit DF
  next-hop    Next hop address //Adres następnego przeskoku
  precedence  Set precedence field //Ustawia pole priorytetu
  qos-group   Set QOS Group ID //Ustawia identyfikator grupy QOS
  tos         Set type of service field //Ustawia pole rodzaju usługi
```

Polityka routingu (ang. *policy routing*) to termin określający wykorzystanie map routingu do zmiany informacji o tym, dokąd kierowany jest pakiet. Stosuj tę funkcję ostrożnie, gdyż wiąże się ona z przełączaniem procesów, co może znacząco obciążyć procesor routera. Przełączanie procesów opisane jest w rozdziale 15.

Budowa mapy routingu

Mapy routingu zbudowane są z **klauzul** (ang. *clauses*), od nich też biorą swoją nazwę. Nazwa znajduje się w każdej klauzuli. Klauzule są ponumerowane, aby określić kolejność, w jakiej powinny być sprawdzane, oraz żeby umożliwić Ci wykonanie lub ominięcie tylko niektórych działań bez konieczności ponownego wprowadzania całej mapy routingu. Domyślnym numerem klauzuli jest 10, jest więc dobrym pomysłem, aby numerować je w interwałach co 10. Pozwoli

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

to na wstawianie wielu klauzul bez konieczności przeprojektowywania całej mapy routingu. Możesz dodawać poszczególne klauzule w dowolnym momencie. Analizator składniowy (ang. *parser*) umieści je w konfiguracji w odpowiedniej kolejności.

Każda klauzula może być klauzulą `permit` (zezwalającą) lub `deny` (zabraniającą). Domyślnie klauzule mają ustawiony parametr `permit`. Sposób, w jaki wartości `permit` i `deny` wpływają na przetwarzanie mapy routingu, zależy od jej zastosowania. Następny podrozdział przedstawia przykład polityki routingu z wykorzystaniem map routingu.

W każdej klauzuli można stosować dwa typy komend:

`match`
: Wybiera trasę lub pakiety na podstawie zdefiniowanych kryteriów.

`set`
: Modyfikuje informacje w pakietach lub tabelach protokołu routingu na podstawie zdefiniowanych kryteriów.

Komendy `match` są odczytywane w kolejności wprowadzenia ich do klauzuli. Wpisy `match` mogą być odczytywane w dwojaki sposób: kilka parametrów dla komendy `match` w pojedynczej linii będzie traktowane jak logiczny warunek „LUB" (ang. *or*), natomiast kilka komend `match` w osobnych liniach poleceń będzie traktowane jak logiczny warunek „I" (ang. *and*).

Poniższy kod konfiguruje mapę routingu GAD z domyślnymi wartościami klauzuli `permit` oraz `10`. Pierwszy warunek `match` sprawdza, czy adres IP pasuje do listy dostępu 101 „LUB" 102 „LUB" 103. Drugi warunek `match` sprawdza, czy długość pakietu podana w bajtach mieści się w przedziale 200 – 230:

```
route-map GAD permit 10
 match ip address 101 102 103
 match length 200 230
 set ip next-hop 10.10.10.10
```

Ze względu na naturę funkcjonowania map routingu wymagane jest, by jedna z trzech list dostępu (ACL) została dopasowana „I" aby długość pakietu odpowiadała drugiemu warunkowi, żeby polecenie `set` zostało wykonane. Jeśli nie zostaną spełnione równocześnie obydwa te warunki, żadna akcja nie zostanie podjęta dla tej klauzuli. Zostanie wtedy odczytana kolejna klauzula z wyższym numerem.

Kiedy w danej klauzuli nie ma żadnego warunku `match`, oznacza to, że wszystkie pakiety lub trasy „pasują". W takim przypadku komenda `set` jest wykonywana dla każdego pakietu lub trasy.

Jeśli nie ma ani jednej komendy `set`, nie jest podejmowane żadne działanie oprócz tego, które określa sama klauzula. Jest to użyteczne, kiedy chcesz ograniczyć redystrybucję w protokołach routingu — ponieważ nie chcesz wtedy niczego zmieniać, nie ma potrzeby stosowania komendy `set`. Mapa routingu zezwoli na redystrybucję trasy, dopóki nie zostanie napotkana klauzula `deny`.

Poniższa mapa routingu może być zastosowana dla instrukcji `redistribute`. Dopuści ona wszystkie trasy, które są zgodne z listą dostępu 101 „LUB" listą dostępu 102. Wszystkie inne trasy zostaną odrzucone:

```
route-map GAD permit 10
 match ip address 101 102
route-map GAD deny 20
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kiedyś można było w mapach routingu definiować warunek „I", umieszczając odniesienie do każdej listy dostępu w osobnej linii polecenia `match`. Taki przykład zawiera pierwsze wydanie książki **Sieciowy wojownik**. Jeśli postąpiłbyś podobnie w najnowszych wersjach IOS, router zmieniłby Twoje wpisy tak, aby mapa routingu wykonywała warunek „LUB" dla tych list dostępu. Oto oryginalny kod z pierwszego wydania:

```
R1-PBX(config)#route-map GAD permit 10
R1-PBX(config-route-map)#match ip address 101
R1-PBX(config-route-map)#match ip address 102
R1-PBX(config-route-map)#route-map GAD deny 20
```

Obecnie router skonwertowałby tę konfigurację następująco:

```
route-map GAD permit 10
 match ip address 101 102
!
route-map GAD deny 20
```

Analogicznie moglibyśmy wykorzystać mapę routingu do zablokowania określonych tras i dopuszczenia wszystkich pozostałych przez proste odwrócenie kolejności klauzul `permit` oraz `deny`:

```
route-map GAD deny 10
 match ip address 101 102
route-map GAD permit 20
```

Przykład polityki routingu

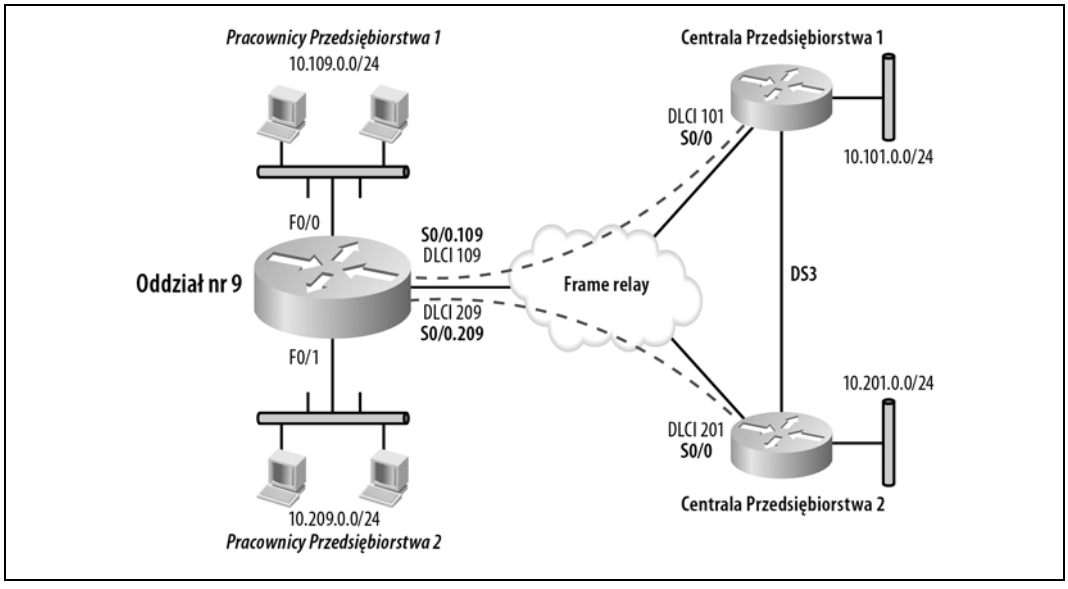
Polityka routingu (ang. *policy routing*) polega na rutowaniu pakietów na podstawie zasad innych niż te, które narzuca standardowo protokół routingu. Przykładowo przy tej technice możesz wysyłać pakiety do miejsc przeznaczenia różniących się od tych, które wyznacza protokół routingu uruchomiony w routerze. Istnieją oczywiście pewne ograniczenia, ale funkcja ta może pomóc Ci znaleźć wyjście z niektórych patowych sytuacji. Rysunek 14.1 ilustruje przykład inspirowany prawdziwym problemem, z którym się spotkałem.

Dwa przedsiębiorstwa, Przedsiębiorstwo 1 i Przedsiębiorstwo 2, są partnerami. Żeby zaoszczędzić pieniądze, partnerzy zdecydowali się stworzyć sieć, w której każdy oddział będzie osobnym biurem podłączonym bezpośrednio do centrali obu przedsiębiorstw jednocześnie. Postanowili także podzielić się kosztami wszystkich routerów, z których każdy będzie obsługiwał sieć w jednym oddziale. Jeden interfejs Ethernet każdego routera posłużył do podłączenia pracowników Przedsiębiorstwa 1, a drugi interfejs Ethernet do podłączenia pracowników Przedsiębiorstwa 2. Pomimo że pracownicy każdego z przedsiębiorstw siedzieli fizycznie w tych samych pomieszczeniach biurowych, zastrzeżono, by nie mogli kontaktować się między sobą poprzez tę sieć. W związku z tym musieliśmy stworzyć listy dostępu dla interfejsów Ethernet, aby zapobiec interakcji pomiędzy sieciami obu biur.

Projekt ten jest doskonałym przykładem sytuacji, w której polityka i pieniądze stają ponad zasadami dobrej inżynierii. Jednak naszym zadaniem nie było oceniać, ale sprawić, by sieć zadziałała w taki sposób, w jaki życzył sobie tego klient.

Aby bardziej skomplikować zadanie, pracownicy mieli korzystać jedynie z łącza Frame Relay, osobnego dla każdego przedsiębiorstwa. Każde z tych łączy zostało wykupione dla pracowników odpowiednio przez te przedsiębiorstwa. Problem w tym, że pracownicy obu przedsiębiorstw w każdym oddziale korzystali jednocześnie z serwerów w obu centralach. Innymi

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 14.1. Przykład polityki routingu

słowy, jeśli pracownik Przedsiębiorstwa 1 w Oddziale 9 musiał skorzystać z serwera w centrali Przedsiębiorstwa 2, nie mógł skorzystać z linku łączącego Oddział 9 z Przedsiębiorstwem 2. Zamiast tego musiał użyć linku dostarczonego przez Przedsiębiorstwo 1. Należało więc zapewnić (i zabezpieczyć) trasę dla tego żądania poprzez link DS3[1] łączący centrale obu przedsiębiorstw. Takie rozwiązanie trzeba było zastosować dla ponad 300 oddziałów, z których każdy miał być skonfigurowany w ten sam sposób. Nie mogliśmy też instalować dodatkowych urządzeń.

Oto tabela routingu z routera w Oddziale 9:

```
Branch-9#sho ip route
[fragment usunięty]
     172.16.0.0/24 is subnetted, 2 subnets
C       172.16.201.0 is directly connected, Serial0/0.209
C       172.16.101.0 is directly connected, Serial0/0.109
     10.0.0.0/24 is subnetted, 4 subnets
C       10.109.0.0 is directly connected, FastEthernet0/0
D       10.101.0.0 [90/2172416] via 172.16.101.1, 00:16:02, Serial0/0.109
D       10.201.0.0 [90/2172416] via 172.16.201.1, 00:16:06, Serial0/0.209
C       10.209.0.0 is directly connected, FastEthernet1/0
```

Jak widzisz, protokół routingu (w tym przypadku EIGRP) robi to, do czego został zaprojektowany. Sieć 10.101.0.0/24 jest dostępna poprzez najkrótszą ścieżkę — bezpośrednie łącze na S0/0.109. Analogicznie sieć 10.201.0.0/24 jest dostępna poprzez najkrótszą ścieżkę na S0/0.209.

Problemem jest to, że routing oparty jest na adresach docelowych. Jeśli użytkownik jednej z lokalnie podłączonych sieci Ethernet chce dostać się do sieci centrali któregoś z przedsiębiorstw, router nie uwzględnia pochodzenia pakietów. Protokół routingu zapewnia po prostu najbardziej optymalną ścieżkę do sieci docelowej. Dzięki mapom routingu oraz polityce routingu jesteśmy w stanie to zmienić.

[1] Standard stosowany w Ameryce Północnej; europejskim odpowiednikiem jest E3 — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Konfiguracja dla linków Ethernet w Oddziale 9 jest prosta (usunąłem listy dostępu, które uniemożliwiają komunikację między sieciami Ethernet, aby uniknąć zamieszania; listy te nie są związane z tematem dyskusji):

```
interface FastEthernet0/0
 ip address 10.109.0.1 255.255.255.0
interface FastEthernet0/1
 ip address 10.209.0.1 255.255.255.0
```

Musimy dodać do każdego interfejsu Ethernet mapę polityki routingu zawierającą instrukcję dla routera, żeby zmieniał następny przeskok dla każdego pakietu wysyłanego do biur centralnych, opierając się na adresie źródłowym zawartym w pakiecie.

Na początek musimy zdefiniować nasze mapy routingu. Zastosowany w nich schemat logiczny przedstawiony jest w poniższym fragmencie pseudokodu:

```
JEŚLI {sieć źródłowa to 10.109.0.0/24     ORAZ sieć docelowa to 10.201.0.0/24}
WTEDY wyślij pakiet przez interfejs S0/0.109
JEŚLI {sieć źródłowa to 10.209.0.0/24
    ORAZ sieć docelowa to 10.101.0.0/24}
WTEDY wyślij pakiet przez interfejs S0/0.209
```

W terminologii map routingu musieliśmy dopasować poleceniem `match` adres docelowy pakietu, a następnie ustawić poleceniem `set` następny przeskok. Aby osiągnąć nasz cel, zastosujemy te mapy do interfejsów wejścia.

Żeby dopasować adresy IP w mapach routingu, musimy stworzyć listy dostępu, do których będziemy się odwoływać. Przygotowałem dwie listy dostępu dla sieci docelowych. Lista dostępu 101 dopasowuje sieć 10.101.0.0/24 (Przedsiębiorstwo 1), a lista dostępu 102 dopasowuje sieć 10.201.0.0/24 (Przedsiębiorstwo 2):

```
access-list 101 permit ip any 10.101.0.0 0.0.0.255
access-list 101 remark <[ Sieć Przedsiębiorstwa-1 ]>
!
access-list 102 permit ip any 10.201.0.0 0.0.0.255
access-list 102 remark <[ Sieć Przedsiębiorstwa-2 ]>
```

Powróćmy teraz do naszego pseudokodu i zaktualizujmy go o informacje z konfiguracji routera:

```
JEŚLI {sieć źródłowa to 10.109.0.0/24 (int F0/0)
    ORAZ sieć docelowa to 10.201.0.0/24 (ACL 102)}
WTEDY wyślij pakiet przez interfejs S0/0.109
JEŚLI {sieć źródłowa to 10.209.0.0/24 (int F0/1)
    ORAZ sieć docelowa to 10.101.0.0/24 (ACL 101)}
WTEDY wyślij pakiet przez interfejs S0/0.209
```

Kiedy sieci docelowe są już zdefiniowane, możemy stworzyć klauzule mapy routingu, które będą się do nich odnosić. Po dopasowaniu sieci docelowej trzeba zmienić interfejs, na który przełączane będą pakiety w obrębie routera. Pierwsza mapa routingu wymusza, aby pakiety skierowane do centrali Przedsiębiorstwa 1 przechodziły przez łącze Przedsiębiorstwa 1:

```
route-map Przedsiębiorstwo-1 permit 10
 match ip address 102
 set interface Serial0/0.109
```

Druga mapa wymusza analogiczne zachowanie dla pakietów skierowanych do centrali Przedsiębiorstwa 2, przesyłając je przez łącze Przedsiębiorstwa 2:

```
route-map Przedsiębiortwo-2 permit 10
 match ip address 101
 set interface Serial0/0.209
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W tym przykładzie fragment pseudokodu dotyczący sieci źródłowej jest dopasowany poprzez zastosowanie mapy routingu do interfejsu, do którego podłączona jest ta sieć. Konfigurujemy to komendą interfejsu `ip policy`.

Zastosowaliśmy tutaj nieco niewłaściwą logikę. Wiedzieliśmy, że pakiety pracowników Przedsiębiorstwa 1 skierowane do centrali tego przedsiębiorstwa będą rutowane prawidłowo bez konieczności ich modyfikowania podobnie jak w przypadku Przedsiębiorstwa 2. Jedynym momentem, w którym musieliśmy zaingerować, był moment, kiedy pracownicy jednego z przedsiębiorstw łączyli się z serwerem w centrali drugiego przedsiębiorstwa. Wtedy i tylko wtedy musieliśmy zmuszać pakiety, żeby wybierały inne łącze. Aby to osiągnąć, zastosowaliśmy mapę routingu Przedsiębiorstwa 2 do interfejsu Ethernet Przedsiębiorstwa 1 i vice versa. Zrobiliśmy to na routerze Oddziału 9:

```
interface FastEthernet0/0
 description <[ Użytkownicy Przedsiębiorstwa-1 ]>
 ip address 10.109.0.1 255.255.255.0
 ip policy route-map Przedsiębiorstwo-1 half-duplex

interface FastEthernet0/1
 description <[ Użytkownicy Przedsiębiorstwa-2 ]>
 ip address 10.209.0.1 255.255.255.0
 ip policy route-map Przedsiębiorstwo-2 half-duplex
```

Polityka routingu jest stosowana, kiedy interfejs otrzymuje pakiet. Z tego powodu reguły w tym przykładzie musiały być zaimplementowane na interfejsach Ethernet.

Monitorowanie polityki routingu

Kiedy już skonfigurujesz politykę routingu, skąd będziesz wiedział, czy działa? W poprzednim przykładzie, dotyczącym routingu opartego na źródle pakietów, zmienialiśmy sposób ich rutowania. Nie zmieniała się jednak tabela routingu IP. Normalnie spojrzelibyśmy na tabelę routingu, aby określić, gdzie router wysyła pakiety. Jednak przy stosowaniu polityki routingu tabela routingu przestaje być najbardziej wiarygodnym źródłem informacji.

Polityka routingu pomija tabelę routingu, co może być dezorientujące, kiedy diagnozujesz problemy związane z routingiem. Jeśli pakiety nie są rutowane w sposób, jakiego byś oczekiwał, sprawdź reguły polityki routingu na interfejsach.

Żeby sprawdzić, jakie reguły polityki routingu są przypisane do określonych interfejsów, użyj komendy `show ip policy`:

```
Branch-9#sho ip policy
Interface          Route map
FastEthernet0/0    Przedsiębiorstwo-1 FastEthernet0/1    Przedsiębiorstwo-2
```

Inną metodą sprawdzenia, czy włączona jest polityka routingu, jest skorzystanie z komendy `show ip interface`. Komenda ta dostarcza dużej ilości informacji, przydatne więc będzie ich przefiltrowanie argumentem `include Policy`:

```
Branch-9#sho ip int f0/0 | include Policy
  Policy routing is enabled, using route map Przedsiębiorstwo-1
  BGP Policy Mapping is disabled
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Problem z tymi komendami polega na tym, że informują one jedynie, iż reguły polityki routingu zostały zastosowane, nie pokazują jednak, w jaki sposób działają. Polecenie `show route-map` wyświetli wszystkie mapy routingu skonfigurowane w routerze, podając przy tym pewne przydatne statystyki dotyczące tego, ile razy mapa routingu została wykorzystana dla celów polityki routingu. Jest to informacja skumulowana, możesz więc jedynie założyć, że skoro wartości liczników rosną, to mapa routingu działa w sposób, w jaki powinna:

```
Branch-9#sho route-map
route-map Przedsiębiorstwo-1, permit, sequence 10
  Match clauses:
    ip address (access-lists): 102
  Set clauses:
    interface Serial0/0.109
    ip next-hop 172.16.101.1
  Policy routing matches: 626 packets, 65304 bytes
route-map Przedsiębiorstwo-2, permit, sequence 10
  Match clauses:
    ip address (access-lists): 101
  Set clauses:
    interface Serial0/0.209
  Policy routing matches: 656 packets, 68624 bytes
```

Jeszcze inną możliwością zweryfikowania, czy router stosuje zdefiniowane reguły polityki routingu, jest użycie komendy `debug ip policy`.

Uważaj, kiedy stosujesz komendę `debug`, gdyż może ona wpłynąć na działanie routera. Pamiętaj, że polityka routingu stosowana jest dla każdego pakietu, który jest odbierany na interfejsie. Zrób najpierw testy w środowisku laboratoryjnym, zanim zaczniesz stosować tę komendę dla sieci produkcyjnej.

Tutaj mamy sytuację, w której stacja robocza w sieci pracowniczej Przedsiębiorstwa 2 w Oddziale 9 pinguje sieć centrali Przedsiębiorstwa 1 (10.101.0.1):

```
D       10.101.0.0 [90/2172416] via 172.16.101.1, 03:21:29, Serial0/0.109
```

Według tablicy routingu pakiety te powinny być rutowane przez interfejs S0/0.109. Jednak kiedy użytkownik pinguje 10.101.0.1, informacja z debugowania pokazuje zupełnie co innego:

```
04:49:24: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1, len 100, FIB policy match
04:49:24: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1, g=172.16.101.1, len 100,
FIB policy routed
04:49:25: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1, len 100, FIB policy match
04:49:25: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1, g=172.16.101.1, len 100,
FIB policy routed
04:49:25: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1, len 100, FIB policy match
04:49:25: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1, g=172.16.101.1, len 100,
FIB policy routed
04:49:25: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1, len 100, FIB policy match
04:49:25: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1
Branch-9#, g=172.16.101.1, len 100, FIB policy routed
04:49:25: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1, len 100, FIB policy match
04:49:25: IP: s=10.209.0.2 (FastEthernet0/1), d=10.101.0.1, g=172.16.101.1, len 100,
FIB policy routed
```

Zwróć uwagę na wersy zaznaczone wytłuszczonym drukiem. Każdy z nich dotyczy pojedynczego pakietu ping. Najpierw pojawia się wpis `FIB policy match`. Wskazuje to, że instrukcja `match` w mapie routingu powiodła się. Następna linia zawiera frazę `FIB policy routed`. To z kolei informuje, że do pakietu zastosowano reguły polityki routingu zamiast zwykłego routingu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Poniżej przykład pakietów, które nie zostały dopasowane do tablicy routingu i z tego powodu były rutowane normalnie.

```
04:52:35: IP: s=10.209.0.2 (FastEthernet0/1), d=10.201.0.1, len 100, FIB policy
rejected(no match) - normal forwarding
04:52:35: IP: s=10.209.0.2 (FastEthernet0/1), d=10.201.0.1, len 100, FIB policy
rejected(no match) - normal forwarding
04:52:35: IP: s=10.209.0.2 (FastEthernet0/1), d=10.201.0.1, len 100, FIB policy
rejected(no match) - normal forwarding
04:52:35: IP: s=10.209.0.2 (FastEthernet0/1), d=10.201.0.1, len 100, FIB policy
rejected(no match) - normal forwarding
04:52:35: IP: s=10.209.0.2 (FastEthernet0/1), d=10.201.0.1, len 100, FIB policy
rejected(no match) - normal forwarding
```

Ponownie zaznaczyłem wytłuszczonym drukiem wpis dotyczący pojedynczego pakietu. Tym razem widzimy frazę `FIB policy rejected(no match) - normal forwarding`, która wskazuje, że pakiet nie pasował do żadnej klauzuli w mapie routingu, został więc przekazany dalej w normalny sposób.

Inny przykład działania map routingu znajdziesz w rozdziale 11.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 15.

Algorytmy przełączania w routerach Cisco

Termin **przełączanie** (ang. *switching*), użyty w kontekście routerów, opisuje proces przenoszenia pakietów z jednego interfejsu na drugi w obrębie routera. Transportowane pakiety docierają do jednego z interfesjów i muszą być przeniesione na inny interfejs na podstawie informacji routingu przechowywanych w routerze.

Routing jest procesem wyboru ścieżek oraz przekazywania pakietów do ich miejsc przeznaczenia, znajdujących się fizycznie poza routerem. **Przełączanie** to wewnętrzne przekazywanie pakietów pomiędzy interfejsami.

Podobnie jak istnieją różne protokoły routingu służące do wyznaczania zewnętrznych ścieżek dla pakietów, tak istnieje też kilka metod wewnętrznego przełączania. Te algorytmy przełączania, inaczej **ścieżki** (ang. *paths*), są istotnym sposobem zwiększania (lub zmniejszania) wydajności routera.

Jednym z najważniejszych czynników wpływających na to, jak szybko pakiet dostanie się od źródła do miejsca przeznaczenia, jest opóźnienie przetwarzania danych (ang. *processing delay*) obecne w każdym routerze wzdłuż całej trasy. Każda z dostępnych metod przełączania ma zdecydowanie inny wpływ na wydajność routera. Wybranie właściwej metody — i umiejętność diagnozowania problemów z nią związanych — pomoże doświadczonemu administratorowi utrzymać najwyższą efektywność sieci.

Router musi przenosić pakiety z jednego interfejsu do drugiego, podobnie jak robi to switch. Sposób przenoszenia pakietów między interfejsami określany jest z wykorzystaniem **bazy informacji routingu** (ang. *routing information base* — RIB), która jest stworzona manualnie lub przez protokoły routingu warstwy trzeciej. RIB jest zasadniczo tabelą routingu (szczegóły znajdziesz w rozdziale 9.). Przełączniki Nexus mogą mieć wiele baz RIB, zwanych instancjami VRF (ang. *Virtual Route Forwarding*). Dokładniejsze informacje dotyczące VRF znajdują się w rozdziale 18.

Istnieje wiele typów przełączania w routerach Cisco. Dzielą się one na dwie kategorie: **przełączanie procesowe** (ang. *process switching*) oraz **przełączanie kontekstu z obsługą przerwań** (ang. *interrupt context switching*). Przełączanie zadań wiąże się z wywoływaniem przez procesor tego zadania, które uzyskuje aktualnie dostęp do RIB, oraz z oczekiwaniem na następne zaplanowane wykonanie tego zadania. Przełączanie kontekstu z obsługą przerwań polega na tym, że procesor przerywa bieżące zadanie, aby przełączyć pakiet. Przełączanie kontekstu dzieli się z kolei na 3 rodzaje:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

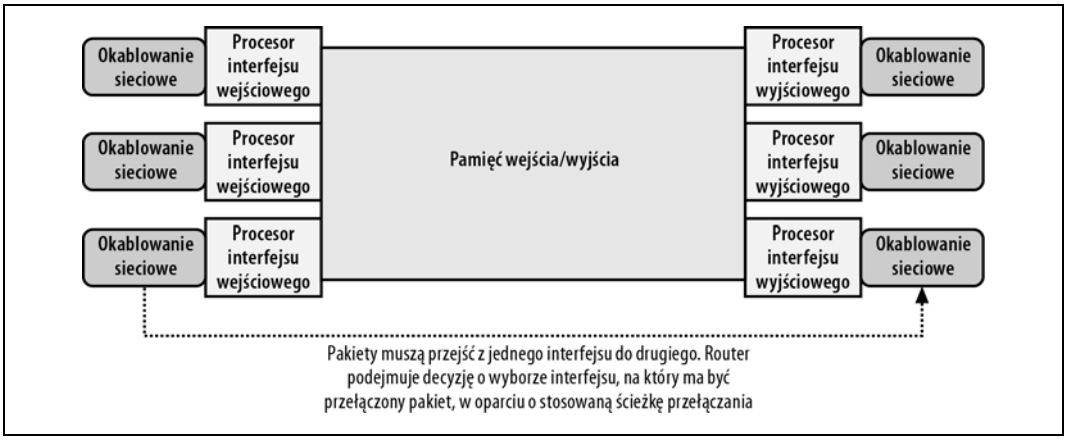
- szybkie przełączanie (ang. *fast switching*),
- optymalne przełączanie (ang. *optimum switching*),
- ekspresowe przekazywanie Cisco (ang. *Cisco Express Forwarding* — CEF).

Każda z metod wykorzystuje różne techniki do wyznaczania interfejsu docelowego. Generalnie rzecz biorąc, proces przełączania pakietu składa się z poniższych etapów:

1. Określenie, czy miejsce docelowe pakietu jest osiągalne.
2. Określenie następnego przeskoku dla miejsca docelowego oraz tego, na który interfejs pakiet powinien być przełączony, aby tam dotrzeć.
3. Zmodyfikowanie nagłówka MAC pakietu tak, aby mógł dotrzeć do miejsca przeznaczenia.

Informacje zawarte w tym rozdziale niekoniecznie mogą być dla Ciebie przydatne w codziennej pracy przy zarządzaniu siecią. Jednak poznanie, jak działają różne metody, pozwoli Ci zrozumieć, dlaczego powinieneś wybrać taką ścieżkę, a nie inną. Wiedza na temat przełączania wewnętrznego routera pozwoli Ci również zrozumieć różnice w zachowaniu routera wynikające z wyboru różnych ścieżek przełączania.

Rysunek 15.1 przedstawia uproszczony schemat procesu logicznego zachodzącego wewnątrz routera. Znajdują się tu trzy interfejsy, z których każdy ma dostęp do pamięci wejścia/wyjścia (ang. *input/output memory*). Kiedy pakiet zostaje dostarczony do jakiegoś interfejsu, router musi zdecydować, do którego interfejsu powinien być dalej przesłany. Kiedy ta decyzja zostanie już podjęta, nagłówki MAC pakietów są modyfikowane, a pakiety są przesyłane dalej.



Rysunek 15.1. Zasada przełączania pakietów w routerze

Tabela routingu zawiera wszystkie informacje niezbędne do wyznaczenia odpowiedniego interfejsu, ale aby pozyskać te dane, trzeba posłużyć się przełączaniem procesowym. Jest to jednak mało wydajne. Zazwyczaj preferowana jest więc metoda przełączania kontekstu z obsługą przerwań.

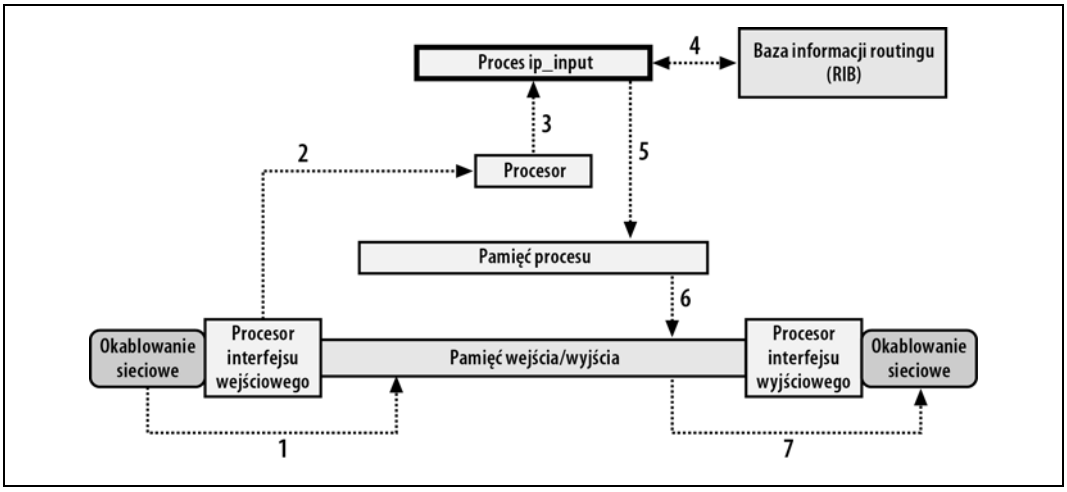
Liczba etapów składających się na proces przekazywania pakietu zależy od tego, jaką ścieżkę przełączania zastosujemy. Metoda przechowywania i pobierania następnego przeskoku oraz informacji o interfejsie również zależy od ścieżek przełączania. Dodatkowo różne modele routerów charakteryzują się innym sposobem działania pamięci oraz miejscem podejmowania decyzji.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przełączanie procesowe

Pierwotna metoda określania, do którego interfejsu powinien zostać przesłany pakiet, nazywa się **przełączaniem procesowym** (ang. *process switching*). Jest to metoda najłatwiejsza do zrozumienia, ponieważ jej zachowanie jest takie, jakiego byś prawdopodobnie oczekiwał.

Przy przełączaniu procesowym, kiedy dostarczany jest pakiet, harmonogram zadań (ang. *scheduler*) wywołuje proces, który analizuje tabelę routingu. Określa w ten sposób, do którego interfejsu pakiet powinien zostać przełączony, a następnie przełącza pakiet. Dzieje się tak dla każdego pakietu, na każdym interfejsie. Rysunek 15.2 ilustruje etapy przełączania procesowego:



Rysunek 15.2. Przełączanie procesowe

1. Procesor interfejsu wykrywa pakiet i przenosi go do pamięci wejścia/wyjścia.
2. Procesor interfejsu generuje **przerwanie odbioru** (ang. *receive interrupt*). W tym czasie procesor centralny (CPU) określa rodzaj pakietu (IP) i jeśli to konieczne, kopiuje go do pamięci procesora (zależy to od platformy). Następnie procesor umieszcza pakiet w odpowiedniej kolejce wejścia procesów i uwalnia przerwanie. Ten proces dla pakietów IP nazywa się **ip_input**.
3. Następnie harmonogram zadań „zauważa" obecność pakietu w kolejce wejścia dla procesu **ip_input** i wpisuje go na listę procesów do wykonania.
4. Podczas wykonywania procesu **ip_input** sprawdzany jest w bazie RIB następny przeskok oraz interfejs wyjścia. Teraz proces **ip_input** odwołuje się do pamięci podręcznej ARP, aby pobrać adres warstwy drugiej dla następnego przeskoku.
5. Proces modyfikuje nagłówek MAC pakietu, dodając odpowiednie adresy, i umieszcza pakiet w kolejce wyjścia właściwego interfejsu.
6. Pakiet jest przenoszony z kolejki wyjścia interfejsu wyjściowego do kolejki transmisji tego interfejsu. W tym kroku realizowany jest również wychodzący QoS.
7. Procesor interfejsu wyjściowego „zauważa" pakiet w swojej kolejce i wysyła go dalej przez okablowanie sieciowe.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W tym procesie jest kilka kluczowych elementów, które czynią go dość powolnym. Po pierwsze, procesor oczekuje na następne zaplanowane wykonanie procesu **ip_input**. Po drugie, gdy proces **ip_input** zostanie już uruchomiony, odnosi się on do bazy RIB. Operacja ta jest bardzo powolna. Proces **ip_input** posiada ten sam priorytet co inne procesy w routerze, takie jak protokoły routingu oraz interfejs serwera sieciowego HTTP.

Zaletą przełączania procesowego jest to, że jest on dostępny na każdej platformie routerów Cisco, bez względu na typ i datę produkcji sprzętu. Pakiety, których źródłem lub miejscem przeznaczenia jest sam router, są zawsze przetwarzane metodą przełączania procesowego. Są to takie pakiety jak pułapki SNMP oraz pakiety telnet przeznaczone dla routera.

Jak możesz sobie wyobrazić, w przypadku dużych routerów lub routerów, które przekazują dużą liczbę pakietów, przełączanie procesowe może być bardzo obciążające. Nawet przy małych routerach przełączanie procesowe może powodować problemy z wydajnością. Spotkałem się z routerami z serii 2600 obsługującymi tylko pojedynczą sieć T1[1], w których wykorzystanie CPU podczas przełączania procesowego wynosiło średnio 60 – 80 procent.

Przełączanie procesowe nigdy nie powinno być stosowane, jeśli możemy wybrać inną metodę. Każda inna metoda zapewni dużo większą wydajność.

Przełączanie kontekstu z obsługą przerwań

Przełączanie kontekstu z obsługą przerwań jest o wiele szybsze niż przełączanie procesowe. Zwiększenie szybkości działania uzyskane jest w dużej mierze dzięki temu, że proces **ip_input** jest rzadko przywoływany. Przełączanie kontekstowe działa natomiast w ten sposób, że przerywa bieżący proces na routerze, aby przełączyć pakiet. Przełączanie kontekstowe zazwyczaj pomija bazę RIB i działa na podstawie tabel równoległych, które są skonstruowane efektywniej (szczegóły dotyczące tych tabel różnią się w zależności od wykorzystywanej ścieżki przełączania). Możesz również zaoszczędzić sporo czasu na tym, że procesor nie musi czekać na zakończenie bieżącego procesu.

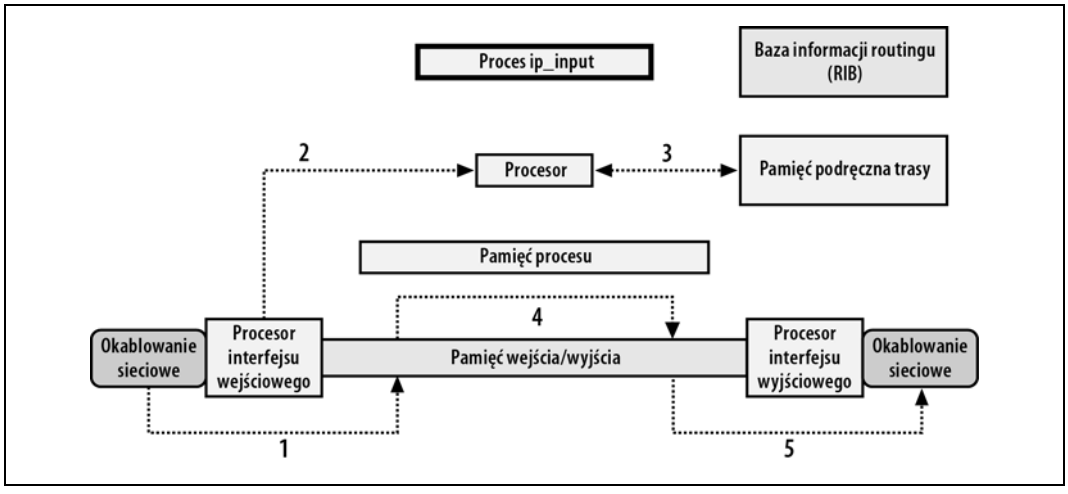
Podstawowe etapy przełączania kontekstu z obsługą przerwań przedstawia rysunek 15.3.

Przełączanie kontekstu z obsługą przerwań opiszę dużo szerzej, omawiając przy tym również różne ścieżki przełączania: szybkie przełączanie, optymalne przełączanie oraz CEF. Przełączanie kontekstu składa się z poniższych etapów:

1. Procesor interfejsu wykrywa pakiet i przenosi go do pamięci wejścia/wyjścia.
2. Procesor interfejsu generuje przerwanie odbioru. W tym czasie procesor centralny (CPU) określa rodzaj pakietu (IP) i zaczyna przełączać pakiet.
3. Procesor przeszukuje pamięć podręczną trasy w poszukiwaniu następujących informacji, które zostaną użyte do zmodyfikowania nagłówka MAC pakietu:
   a) Czy miejsce docelowe jest osiągalne?
   b) Jaki powinien być interfejs wyjściowy?

[1] Jest to standard wykorzystywany w Ameryce Północnej; jego europejskim odpowiednikiem jest sieć E1 — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 15.3. Przełączanie kontekstu z obsługą przerwań

c) Jaki jest następny przeskok?

d) Na co powinien być skonwertowany adres MAC?

4. Pakiet jest kopiowany do kolejki transmisji lub do kolejki wyjścia interfejsu wyjściowego. Przerwanie odbioru zostaje zakończone, a pierwotny proces może być kontynuowany.
5. Procesor interfejsu wyjściowego wykrywa pakiet w swojej kolejce i wysyła go dalej przez okablowanie sieciowe.

Oczywistą różnicą jest, że w tym przypadku mamy tylko pięć etapów w porównaniu do siedmiu etapów przełączania procesowego. Duże znaczenie ma tu to, że bieżący proces na routerze jest przerywany, podczas gdy poprzednia metoda wymagała oczekiwania na następne zaplanowane wykonanie procesu **ip_input**.

Baza RIB jest również całkowicie pomijana w tym modelu, a niezbędne informacje są pozyskiwane z innych źródeł. W przykładzie przedstawionym na rysunku 15.3 źródło to nazwane jest **pamięcią podręczną trasy** (ang. *route cache*). Jak przekonasz się później, każda ścieżka przełączania posiada swój sposób uzyskiwania, przechowywania oraz pobierania tej informacji. Stosowanie odmiennej metody jest tym, co odróżnia indywidualne ścieżki przełączania wewnątrz grupy przełączania kontekstu z obsługą przerwań.

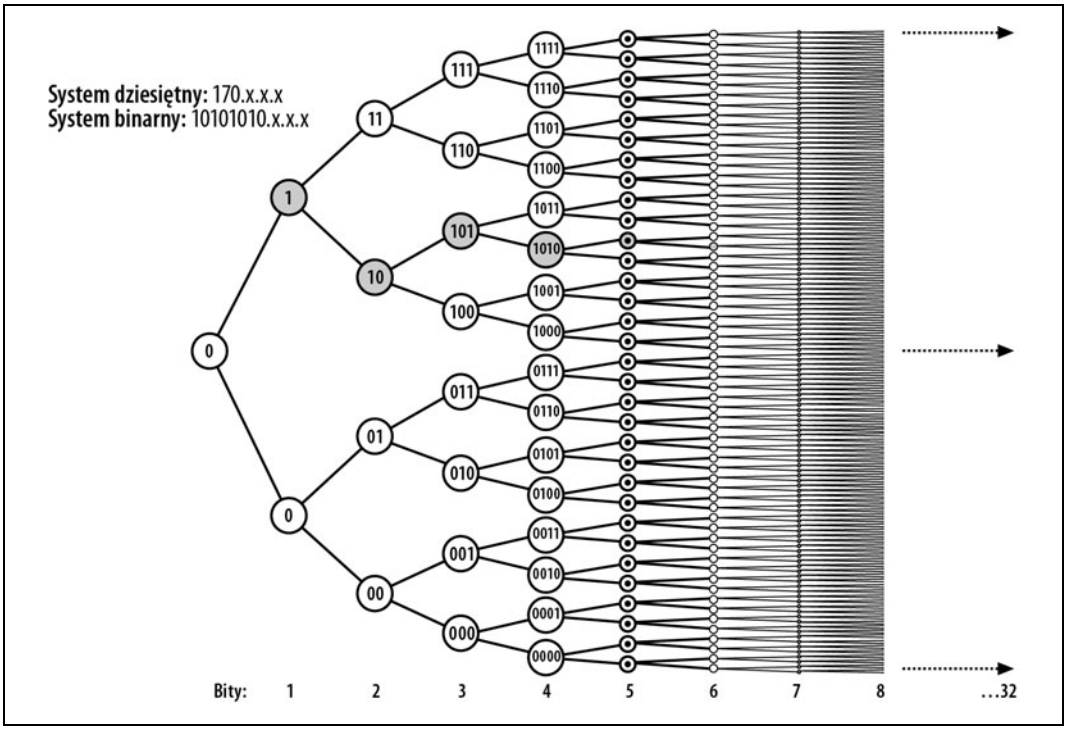
Szybkie przełączanie

Szybkie przełączanie (ang. *fast switching*) polega na tym, że do pierwszego pakietu w konwersacji stosowana jest metoda przełączania procesowego, a następnie uzyskana informacja o adresie docelowym dla pakietu jest przechowywana w tabeli zwanej pamięcią podręczną trasy. Szybkie przełączanie wykorzystuje format **drzewa binarnego** (ang. *binary tree*) do zapisywania i pobierania informacji z pamięci podręcznej trasy.

Rysunek 15.4 pokazuje przykład drzewa binarnego w takiej postaci, w jakiej mogłoby być zastosowane w szybkim przełączaniu. Każda kolejna gałąź dodaje wartość 0 lub 1 do wartości poprzedniego poziomu, zaczynając od wartości 0 — *the root*. Następnie mamy dwa rozgałęzienia: dolne z wartością 0 i górne z wartością 1. Każdy z tych węzłów dzieli się na dalsze

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

dwie gałęzie, z których dolna zawsze zyskuje kolejne zero, a górna zawsze kolejną jedynkę. Rysunek 15.4 przedstawia osiem poziomów drzewa, jednak właściwe drzewo dla adresów IP powinno mieć 32 poziomy, czyli tyle, ile jest bitów w adresie IP.



Rysunek 15.4. Drzewo binarne dla szybkiego przełączania

Węzły zaznaczone na rysunku 15.4 szarym kolorem odpowiadają adresowi IP 170.x.x.x (wartość binarna dla 170 to 10101010 — to powinno ułatwić Ci wizualizację tego przykładu). Przedstawiona została jedynie część adresu IP, czyli właśnie 170.x.x.x, ponieważ więcej czytelnych węzłów poza ośmioma bitami nie zmieściło się na rysunku.

Zaletą tej konstrukcji, w porównaniu z przeszukiwaniem bazy RIB, jest szybkość dostępu. Informacje dotyczące następnego przeskoku oraz zmian adresów MAC przechowywane są w każdym węźle drzewa. Ponieważ takie drzewo jest modelem deterministycznym, odnalezienie określonych wpisów nie zajmuje wiele czasu.

Wadą takiego wdrożenia jest sam rozmiar tabeli oraz to, że dane dla każdego adresu są składowane w węzłach, przez co rozmiar tych danych nie jest stały. Skoro każdy węzeł może mieć inny rozmiar, to tabela może być nieefektywna.

Pamięć podręczna trasy nie jest związana bezpośrednio z tabelą routingu i jest aktualizowana tylko wtedy, gdy pakiety są przełączane metodą przełączania procesowego. Innymi słowy, pamięć podręczna trasy jest aktualizowana, kiedy pierwszy pakiet dla danego miejsca przeznaczenia jest przełączany. Od tego momentu pamięć podręczna trasy jest wykorzystywana, a pozostałe pakiety są przetwarzane w trybie szybkim. Aby dane w pamięci podręcznej trasy były cały czas aktualne, $^{1}/_{20}$ całej pamięci podręcznej trasy musi być **zrzucana** (ang. *aged out*) co minutę. Do odbudowania informacji stosuje się przełączanie procesowe.

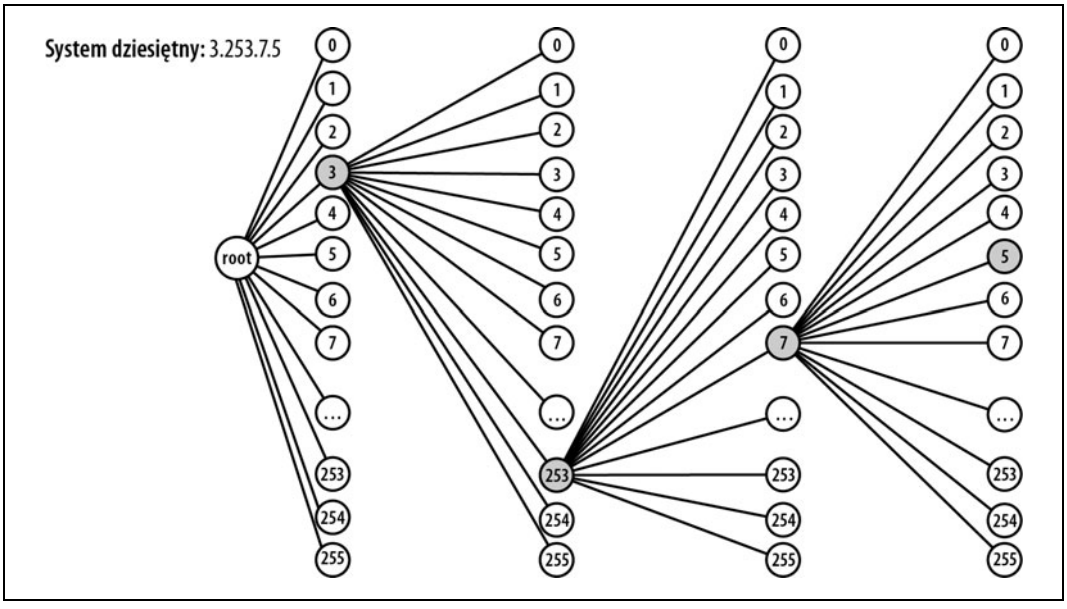
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ponieważ tabela ARP nie jest bezpośrednio związana z zawartością pamięci podręcznej trasy, zmiany w tabeli ARP mogą skutkować unieważnieniem jedynie części pamięci podręcznej trasy. Aby usunąć te rozbieżności, również stosuje się przełączanie procesowe.

Przełączanie optymalne

W przełączaniu optymalnym (ang. *optimum switching*) zamiast drzewa binarnego stosuje się **drzewo wielodrożne** (ang. *multiway tree*) do zapisywania i pobierania danych z pamięci podręcznej trasy. W odniesieniu do adresów IP drzewo wielodrożne działa znacznie szybciej, ponieważ każdy oktet przyjmuje tylko jedną z 256 możliwych wartości. Drzewo binarne jest przeznaczone do obsługi wartości dowolnego typu, więc jego potencjalny rozmiar nie jest ograniczony.

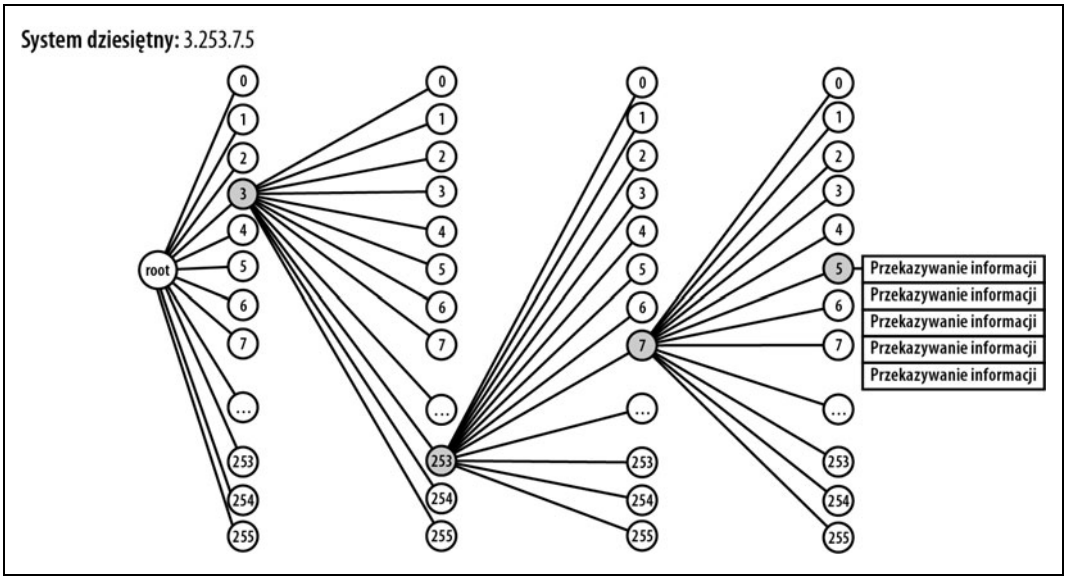
Rysunek 15.5 przedstawia przykład drzewa wielodrożnego, które mogłoby być zastosowane w przełączaniu optymalnym. Pień drzewa (ang. *root*) rozgałęzia się na 256 węzłów ponumerowanych od 0 do 255. Każdy z tych węzłów rozgałęzia się na kolejne 256 węzłów. Schemat ten powtarza się dla 4 poziomów — jeden poziom na każdy oktet. Węzły zaznaczone szarym kolorem pokazują, w jaki sposób może zostać dopasowany adres IP. Adres IP z tego przykładu to 3.253.7.5.



Rysunek 15.5. Drzewo wielodrożne przełączania optymalnego

Ten rodzaj drzewa umożliwia znacznie szybsze wyszukiwanie adresów IP, ponieważ tabela jest zoptymalizowana pod kątem formatu adresu IP. Informacja dla każdej trasy lub adresu IP przechowywana jest w węźle końcowym, tak jak pokazano to na rysunku 15.6. Skoro rozmiar tej informacji jest zmienny oraz z uwagi na to, że nie każdy węzeł musi zawierać jakąkolwiek informację, całkowity rozmiar tabeli jest również zmienny. Jest to pewna wada, gdyż przeszukiwanie drzewa nie jest tak efektywne, jak mogłoby być, gdyby każdy węzeł posiadał stały, znany rozmiar.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 15.6. Dane przechowywane w drzewie wielodrożnym przełączania optymalnego

Ponieważ odpowiednie dane są przechowywane w węzłach i nie mają bezpośredniego związku z bazą RIB lub pamięcią podręczną ARP, wpisy ulegają z czasem przedawnieniu i są odbudowywane w procesie przetwarzania procesowego.

Przełączanie optymalne jest dostępne tylko w routerach z wyższej półki, takich jak modele z serii 7500 czy 7600.

CEF

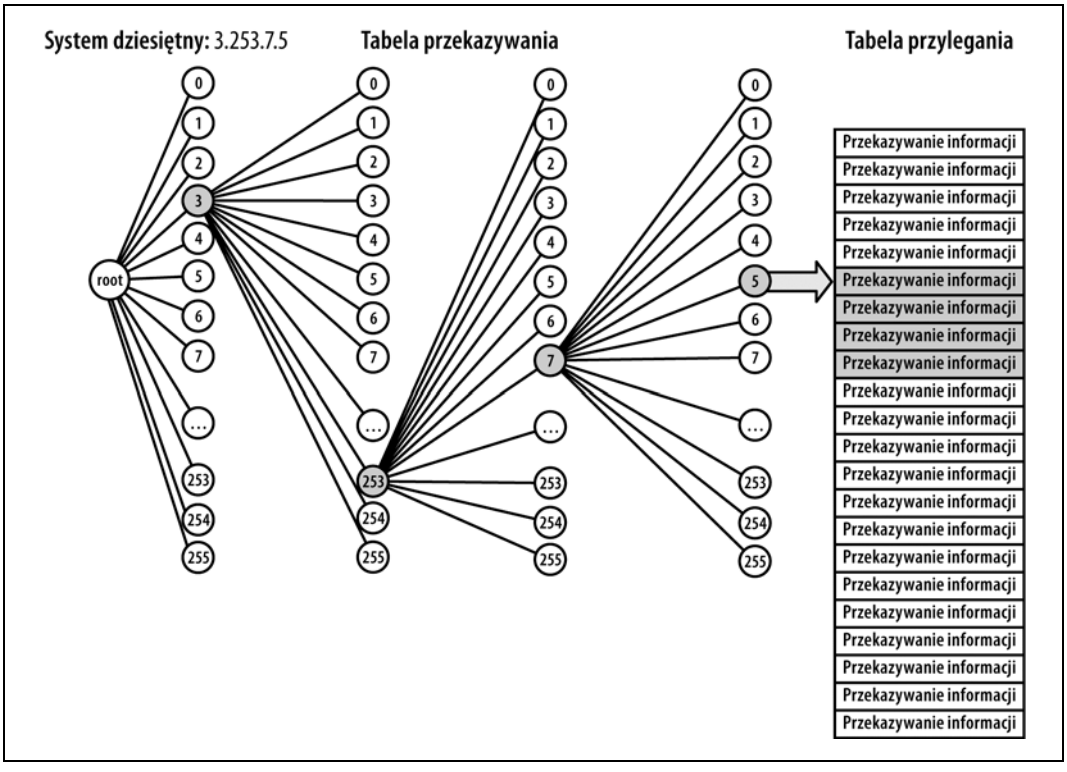
Jeśli chodzi o ścieżki przełączania, CEF jest najlepszym wyborem dla każdego routera obsługującego tę metodę. Prawie wszystkie nowoczesne routery i przełączniki domyślnie stosują przełączanie CEF, szczególnie jeśli pracują w centrach danych.

CEF idzie o krok dalej w stosunku do drzewa wielodrożnego przełączania optymalnego, wprowadzając koncepcję **drzewa trie**. Podstawowe założenia są takie same jak w przypadku drzewa wielodrożnego lub binarnego, ale dane nie są przechowywane w węzłach. Tutaj raczej każdy węzeł staje się odnośnikiem do innej tabeli zawierającej właściwe dane.

Termin **drzewo trie** pochodzi od angielskiego słowa **retrieve** (pobierać) i jest wymawiany jak angielskie słowo **tree** (drzewo). Niektórzy preferują wymowę taką jak wymowa angielskiego słowa **try** (próbować), aby odróżnić ten termin od słowa „tree".

W przełączaniu CEF drzewo trie zwane jest **tabelą przekazywania** (ang. *forwarding table*). Każdy węzeł ma ten sam stały rozmiar i nie przechowuje danych. Pozycja węzła w samym drzewie jest odniesieniem do kolejnej tabeli, zwanej **tabelą przylegania** (ang. *adjacency table*). Tabela ta przechowuje takie istotne dane jak substytut nagłówka MAC oraz informacje o następnym przeskoku dla węzłów. Rysunek 15.7 przedstawia tabele CEF.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 15.7. Tabele przekazywania i przylegania CEF

Jedną z największych zalet przełączania CEF jest to, że tabele konstruowane są bez przełączania procesowego. Obie tabele mogą zostać utworzone bez konieczności oczekiwania, aż pakiet zostanie przesłany do miejsca przeznaczenia. Ponadto skoro tabela przekazywania jest tworzona oddzielnie w stosunku do tabeli przylegania, błąd w jednej z tabel nie prowadzi do dezaktualizacji drugiej. W przypadku modyfikacji w pamięci podręcznej ARP zmienia się tylko tabela przylegania, więc tabela przekazywania nie musi stawać się nieaktualna lub przestarzała.

Przełączanie CEF obsługuje rozkładanie obciążenia, opierając się na ścieżkach o równych kosztach, jak sam zobaczysz w następnym podrozdziale. Rozkładanie obciążenia na poziomie przełączania jest o wiele lepsze od rozkładania obciążenia przez protokoły routingu, ponieważ funkcjonują one na znacznie wyższym poziomie. Najnowsze wersje IOS dodają przełączanie CEF do rozkładania obciążenia protokołów routingu, tak więc przestaje to być problemem.

Ścieżki przełączania — konfiguracja i zarządzanie

Konfiguracja ścieżek przełączania odbywa się zarówno globalnie, jak i na poziomie interfejsu, co zapewnia elastyczność konfiguracji różnych ścieżek przełączania na każdym interfejsie. Przykładowo możesz potrzebować wyłączyć CEF na określonym interfejsie, aby zdiagnozować jakiś problem.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przełączanie procesowe

Aby zmusić router do korzystania z metody przełączania procesowego, wyłącz wszystkie inne metody przełączania.

Przedstawię teraz, jak radzi sobie router Cisco 2621XM z przepływem 600k danych poprzez interfejs szeregowy s0/1:

```
R1#sho int s0/1 | include minute
  5 minute input rate 630000 bits/sec, 391 packets/sec
  5 minute output rate 627000 bits/sec, 391 packets/sec
```

Normalna metoda przełączania dla interfejsu na tym routerze to CEF. Aby sprawdzić, jaka ścieżka przełączania jest zastosowana na interfejsie s0/1, użyj polecenia `show ip interface s0/1 | include switching`:

```
R1#sho ip interface s0/1 | include switching
  IP fast switching is enabled //Szybkie przełączanie jest włączone
  IP fast switching on the same interface is enabled
  IP Flow switching is disabled
  IP CEF switching is enabled //Przełączanie CEF jest włączone
  IP CEF Fast switching turbo vector
  IP multicast fast switching is enabled
  IP multicast distributed fast switching is disabled
```

Zauważ, że włączone jest zarówno szybkie przełączanie, jak i CEF. CEF w pierwszej kolejności będzie próbowało przełączyć pakiet. Jeśli nie będzie w stanie, **przesunie** go (ang. *punt*) do następnej dostępnej ścieżki przełączania — w tym przypadku do szybkiego przełączania. Jeżeli szybkie przełączanie nie poradzi sobie z pakietem, router zastosuje przełączanie procesowe. Kiedy wszystkie pozostałe ścieżki przełączania są wyłączone, router musi stosować przełączanie procesowe dla wszystkich pakietów.

Aby wyłączyć wszystkie ścieżki przełączania kontekstu z obsługą przerwań, zastosuj polecenie `no ip route-cache`:

```
R1(config-if)#no ip route-cache
R1(config-if)# ^Z
R1# sho ip interface s0/1 | include switching
  IP fast switching is disabled //Szybkie przełączanie jest wyłączone
  IP fast switching on the same interface is disabled
  IP Flow switching is disabled
  IP CEF switching is disabled //Przełączanie CEF jest wyłączone
  IP Fast switching turbo vector
  IP multicast fast switching is disabled
  IP multicast distributed fast switching is disabled
```

Wykonałem to dla wszystkich interfejsów routera. Przyjrzyjmy się teraz wykorzystaniu procesora w routerze, wpisując komendę `show processes cpu history`:

```
R1#sho proc cpu hist
     444444444444444444444444411111
     444422222333332222200000555556666644444444444444444444444
100
 90
 80
 70
 60
 50
 40  ************************
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
 30  ************************
 20  ****************************
 10  **********************************
    0....5....1....1....2....2....3....3....4....4....5....5....
              0    5    0    5    0    5    0    5    0    5
               CPU% per second (last 60 seconds) //CPU% na sekundę (ostatnie 60 sekund)
```

Przy przełączaniu CEF router wykorzystywał średnio 4 procent CPU. Kiedy wyłączyłem buforowanie wszystkich tras, router powrócił do metody przełączania procesowego dla wszystkich pakietów. Przy przepływie zaledwie 600k danych router wykorzystuje teraz 40 procent CPU do przekazywania pakietów.

Kiedy router wykorzystuje przełączanie procesowe dla większości swoich pakietów IP, najwyżej na liście będzie zawsze proces **ip_input**. Możesz to sprawdzić, wpisując polecenie `show processes cpu sorted`:

```
R1#sho proc cpu sort
CPU utilization for five seconds: 48%/20%; one minute: 44%; five minutes: 40%
 PID Runtime(ms)   Invoked      uSecs   5Sec   1Min   5Min TTY Process
  31   177614152 132817658       1337 26.60% 24.06% 21.94%   0 IP Input
   3       60540     20255       2988  1.47%  0.22%  0.05%   0 Exec
   4     1496508    210836       7097  0.16%  0.06%  0.05%   0 Check heaps
   1           0         1          0  0.00%  0.00%  0.00%   0 Chunk Manager
   2       12488    328220         38  0.00%  0.00%  0.00%   0 Load Meter
   5         220       332        662  0.00%  0.00%  0.00%   0 Pool Manager
   6           0         2          0  0.00%  0.00%  0.00%   0 Timers
   7           0         8          0  0.00%  0.00%  0.00%   0 Serial Backgroun
   8       12108    327897         36  0.00%  0.00%  0.00%   0 ALARM_TRIGGER_SC
```

Jest to przydatne w diagnozowaniu problemów, kiedy router lub switch wielowarstwowy działają dość wolno. Jeśli zauważysz, że **ip_input** konsumuje większą część zasobów CPU podaną w minutach (jak na powyższym listingu), musisz zadać sobie pytanie, dlaczego router stosuje przełączanie procesowe. Sprawdź swoją konfigurację i poszukaj problemów.

Pakiety, które są zniekształcone, mogą spowodować, że router zmieni metodę na przełączanie procesowe. Spotkałem się z sytuacją, w której sprzęt dostawcy dla obszarów MAE zniekształcał pakiety tak bardzo, że router musiał zastosować przełączanie procesowe. Z podobnym problemem miałem do czynienia w przypadku awarii układów ASIC w przełącznikach. Chcę przez to powiedzieć, że nie tylko to, co dzieje się wewnątrz routera, ale także zewnętrzne urządzenia i zdarzenia mogą powodować, że router powróci do metody przełączania procesowego.

Szybkie przełączanie

W przypadku interfejsu, na którym ustawione jest przełączanie procesowe, możemy zastosować polecenie interfejsu `ip route-cash`, aby włączyć szybkie przełączanie:

```
R1(config)#int s0/1
R1(config-if)# ip route-cache
```

Uruchomi to szybkie przełączanie, ale bez włączania CEF:

```
R1#sho ip int s0/1 | include swi
  IP fast switching is enabled //Szybkie przełączanie jest włączone
  IP fast switching on the same interface is enabled
  IP Flow switching is disabled
  IP CEF switching is disabled //Przełączanie CEF jest wyłączone
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  IP Fast switching turbo vector
  IP multicast fast switching is enabled
  IP multicast distributed fast switching is disabled
```

Dla interfejsu stosującego CEF wyłączenie CEF spowoduje powrót do metody szybkiego przełączania:

```
R1(config)#int s0/1
R1(config-if)# no ip route-cache cef
```

Nie możesz wyłączyć szybkiego przełączania bez jednoczesnego wyłączenia CEF, możesz jednak uruchomić szybkie przełączanie bez jednoczesnego uruchamiania CEF.

Powracając do przykładu z poprzedniego podrozdziału, w przypadku ponownego uruchomienia szybkiego przełączania na routerze stosującym przełączanie procesowe użycie CPU spadnie z powrotem do normalnego poziomu:

```
R1#sho proc cpu hist
                 222224444444444444444444444444444444444444444444
     444444444444000007777711111888889999988888999993333344446
100
 90
 80
 70
 60
 50                 *****     ********************         *
 40                 ****************************************
30                 *****************************************
 20             ***********************************************
 10             ***********************************************
    0....5....1....1....2....2....3....3....4....4....5....5....
              0    5    0    5    0    5    0    5    0    5
               CPU% per second (last 60 seconds) //CPU% na sekundę (ostatnie 60 sekund)
```

CEF

Przełączanie CEF jest domyślnie ustawione na wszystkich nowoczesnych modelach routerów Cisco. Gdybyś jednak potrzebował uruchomić CEF na starszym modelu routera lub gdyby po prostu zostało ono wyłączone, istnieją dwa sposoby, które pozwolą Ci to zrobić. Pierwszym jest użycie globalnej komendy `ip cef`, która włączy CEF na każdym interfejsie obsługującym tę metodę:

```
R1(config)#ip cef
```

Zanegowanie komendy wyłącza CEF globalnie:

```
R1(config)#no ip cef
```

Aby włączyć lub wyłączyć CEF na pojedynczym interfejsie, skorzystaj z komendy `ip route-cash cef`:

```
R1(config)#int s0/1
R1(config-if)# ip route-cache cef
```

Zanegowanie tej komendy wyłącza CEF na danym interfejsie.

Dla pakietów CEF będzie stosowało rozkładanie obciążenia na łącza o równym koszcie. Domyślnie rozkładanie obciążenia oparte jest na miejscach docelowych pakietów (ang. *per-destination basis*). Oznacza to, że każdy pakiet skierowany do określonego miejsca przeznaczenia będzie

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

korzystał z tego samego łącza. CEF umożliwia również skonfigurowanie rozłożenia obciążenia na podstawie pakietu (ang. *per-packet basis*). Może być to korzystne np. wtedy, kiedy istnieje tylko jeden host na oddalonym końcu łącza lub duży serwer, który konsumuje znaczną część przepustowości dla każdego pojedynczego linku.

Niektóre protokoły, jak np. VoIP, nie tolerują rozłożenia obciążenia *per-packet*, ponieważ pakiety mogą przybywać w niewłaściwej kolejności. Kiedy stosujesz taki protokół, zawsze upewnij się, że rozkład obciążenia odbywa się *per-destination*, albo zastosuj protokół wyższego poziomu, taki jak Multilink-PPP.

Aby ustawić dla interfejsu metodę rozkładania obciążenia *per-packet*, wprowadź polecenie interfejsu `ip load-sharing per-packet`:

```
R1(config-if)#ip load-sharing per-packet
```

Z kolei aby skonfigurować dla interfejsu rozkład obciążenia *per-destination*, użyj polecenia `ip load-sharing per-destination`:

```
R1(config-if)#ip load-sharing per-destination
```

Aby wyświetlić w czytelnym formacie tabele CEF, skorzystaj z komendy `show ip cef`:

```
R1#sho ip cef
Prefix              Next Hop            Interface
0.0.0.0/32          receive
10.1.1.0/24         attached            Loopback1
10.1.1.0/32         receive
10.1.1.1/32         receive
10.1.1.255/32       receive
10.2.2.0/24         attached            Loopback2
10.2.2.0/32         receive
10.2.2.2/32         receive
10.2.2.255/32       receive
10.3.3.0/24         192.168.1.2         Serial0/1
10.4.4.0/24         192.168.1.2         Serial0/1
10.5.5.0/24         attached            FastEthernet0/1
10.5.5.0/32         receive
10.5.5.1/32         10.5.5.1            FastEthernet0/1
10.5.5.5/32         receive
10.5.5.255/32       receive
10.10.10.0/24       192.168.1.2         Serial0/1
192.168.1.0/24      attached            Serial0/1
192.168.1.0/32      receive
192.168.1.1/32      receive
192.168.1.255/32    receive
224.0.0.0/4         drop
224.0.0.0/24        receive
255.255.255.255/32  receive
```

Zazwyczaj nie potrzebujesz sprawdzać, jak funkcjonuje CEF, chyba że zostaniesz o to poproszony przez pracowników działu wsparcia technicznego Cisco (ang. *Cisco Technical Assistance Center*). Jedynym przypadkiem, w którym musiałem wykorzystać tę komendę, był problem z CEF, który powodował wysyłanie pakietów przez interfejsy inne niż te określone w tabeli routingu. Powinieneś przyjąć jako generalną zasadę, że nigdy nie należy wyłączać CEF.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

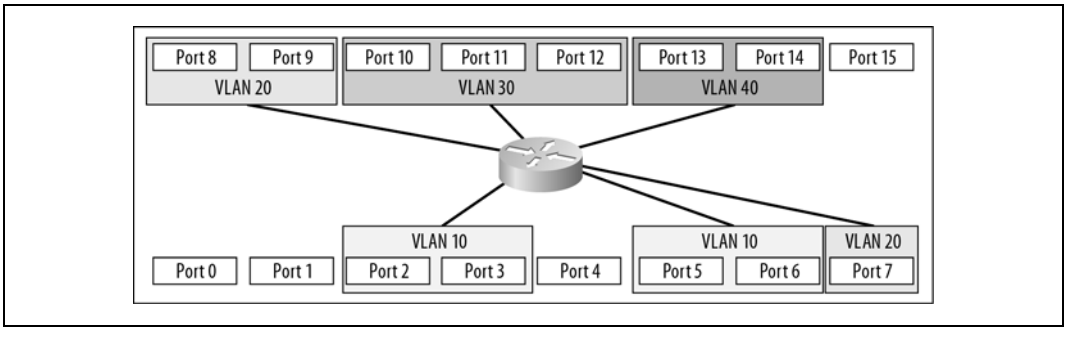
ROZDZIAŁ 16.

Przełączniki wielowarstwowe

W tradycyjnym ujęciu przełączniki operują w warstwie drugiej modelu ISO/OSI. Pierwsze przełączniki wielowarstwowe były zwane **przełącznikami warstwy trzeciej** (ang. *Layer-3 switches*), ponieważ oferowały możliwość routingu między VLAN-ami. Obecnie przełączniki mają praktycznie takie możliwości jak router, łącznie z testowaniem protokołów oraz operowaniem w wyższych warstwach aż do warstwy siódmej. Tak więc przełączniki działające powyżej warstwy drugiej są dzisiaj określane **przełącznikami wielowarstwowymi** (ang. *multilayer switches*).

Główną zaletą przełącznika wielowarstwowego jest routing pomiędzy VLAN-ami, który jest możliwy dzięki wprowadzeniu wirtualnych interfejsów w obrębie przełącznika. Te wirtualne interfejsy są przypisane do VLAN-ów i nazywane **wirtualnymi interfejsami przełączania** (ang. *switched virtual interfaces* — SVI).

Rysunek 16.1 przedstawia zasady routingu wewnątrz przełącznika. Na początek musisz przypisać porty do określonych VLAN-ów. Następnie tworzysz interfejsy SVI, które umożliwią przypisanie VLAN-om adresów IP. Wirtualny interfejs staje się interfejsem wirtualnego routera, zapewniając tym samym routing VLAN-ów.



Rysunek 16.1. Routing między VLAN-ami w obrębie przełącznika

Obecnie większość przełączników wielowarstwowych nie posiada wyodrębnionych pod względem sprzętowym routerów. Funkcje routera zaimplementowane są w obwodach samego przełącznika lub w jednostce zarządzającej (tj. CPU) w przypadku przełącznika modularnego. Starsze wersje przełączników, takie jak switch Cisco 4000 typu *chassis*, posiadają moduł warstwy

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

trzeciej, który został dodany, aby zapewnić funkcjonalność wielowarstwową. Takie moduły nie są już teraz potrzebne, ponieważ funkcjonalność warstwy trzeciej oferowana jest przez większość jednostek zarządzających.

W przełącznikach *chassis-based* ze starszymi wersjami modułów zarządzających router jest wyodrębnionym urządzeniem z własnym systemem operacyjnym. Funkcję routera w tych przełącznikach pełnią „karty-córki" modułów zarządzających, zwane **kartami MSFC** (ang. *multilayer switch function card*). W tych urządzeniach operacje warstwy drugiej są kontrolowane przez system operacyjny CatOS, natomiast routing w warstwie trzeciej przez system operacyjny IOS. Taka konfiguracja, zwana **trybem hybrydowym** (ang. *hybrid mode*), może być początkowo nieco dezorientująca. Dla niektórych posiadanie osobnych systemów operacyjnych dla każdej funkcji jest bardziej sensowne niż ich połączenie w jednym systemie. Jednak dla większości model pojedynczego OS — zwanego **trybem natywnym** (ang. *native mode*) w przełącznikach *chassis* — jest prawdopodobnie dużo łatwiejszy. Switche Nexus posiadają system operacyjny zwany NX-OS, który jest zbliżony do IOS.

W przypadku przełączników *chassis-based* w trybie hybrydowym fizyczne interfejsy oraz VLAN-y muszą zostać skonfigurowane w CatOS. Aby ustanowić routing między nimi, musisz przejść do IOS i skonfigurować interfejsy SVI dla VLAN-ów, które stworzyłeś w CatOS.

Inną opcją dla niektórych przełączników jest zmiana portu przełączającego w port rutujący — umożliwi to bezpośrednie adresowanie tego portu protokołem warstwy trzeciej, takim jak IP. Można tego dokonać na przełącznikach z natywnym systemem IOS lub IOS działającym w trybie hybrydowym. Przełączniki Nexus 7000 mają domyślnie ustawione porty w trybie routera.

Zdarza się, że jestem zmuszony zestawić łącze warstwy trzeciej pomiędzy dwoma przełącznikami wielowarstwowymi. Konfiguracja portu, VLAN-u oraz interfejsu SVI jest wieloetapowa, zwłaszcza jeśli wziąć pod uwagę, że VLAN nie będzie nigdy zawierał żadnych portów poza linkiem między przełącznikami. W takich przypadkach konwersja portu przełączania na port rutujący jest łatwiejsza.

Nie możesz przypisać adresu IP do fizycznego interfejsu, jeśli jest on skonfigurowany na porcie przełączającym (domyślne ustawienie większości switchy Catalyst):

```
3750-IOS(config)#int g1/0/17
3750-IOS(config-if)#ip address 10.10.10.1 255.255.255.0
                          ^
% Invalid input detected at '^' marker. //Wykryto niewłaściwy wpis od miejsca oznaczonego '^'
```

Aby zmienić port przełączający przełącznika wielowarstwowego w port rutujący, zastosuj dla tego portu polecenie `no switchport`:

```
3750-IOS(config)#int g1/0/17
3750-IOS(config-if)#no switchport
```

Kiedy już to zrobisz, będziesz mógł przypisać adres IP do fizycznego interfejsu:

```
3750-IOS(config)#int g1/0/17
3750-IOS(config-if)#ip address 10.10.10.1 255.255.255.0
```

W przełącznikach Nexus 7000 porty są domyślnie ustawione w trybie routera. Jeśli chcesz zmienić je w porty przełączające, użyj komendy interfejsu `switchport`:

```
NX-7K-1(config)# int e8/1-2
NX-7K-1(config-if-range)# switchport
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Porty Ethernet w routerach bywają kosztowne i nie oferują odpowiedniej gęstości portów. Dodanie modułów przełącznika, zapewniających kilka dodatkowych interfejsów, poprawiło ich zagęszczenie. Nic jednak nie może się równać z elastycznością przełączników wielowarstwowych, jeśli chodzi o Ethernet.

Bądź ostrożny, kiedy będziesz wykorzystywał przełącznik w funkcji routera. Projektowałem niegdyś sieć MPLS o zasięgu krajowym. Ponieważ wszystkimi przekaźnikami dla sieci był Metro Ethernet, zdecydowałem, że można zaoszczędzić nieco pieniędzy przez wykorzystanie istniejących przełączników 3750 na końcach każdej gałęzi. Wszystko szło doskonale, dopóki nie spróbowałem skonfigurować kolejkowania z niskim opóźnieniem (ang. *low-latency queuing* — LLQ) dla QoS. Skoro przełączniki nie są routerami, nie obsługują zazwyczaj QoS dla WAN, a raczej CoS dla LAN. Komendy konfiguracji LLQ nie były obsługiwane przez przełączniki 3750. Miałem więc pecha i musiałem prosić o zakup kolejnego sprzętu w dalszej fazie projektu.

Konfiguracja interfejsów SVI

Wirtualne interfejsy przełączania są konfigurowane odmiennie w zależności od platformy przełącznika oraz zainstalowanego systemu operacyjnego. Przyjrzyjmy się temu.

IOS (4500, 6500, 3550, 3750 itp.)

Poniżej listing dla komendy `show ip interface brief` zastosowanej w przełączniku 3750:

```
3750-IOS#sho ip int brief
Interface              IP-Address      OK? Method Status                Protocol
Vlan1                  unassigned      YES NVRAM  administratively down down
GigabitEthernet1/0/1   unassigned      YES unset  down                  down
GigabitEthernet1/0/2   unassigned      YES unset  down                  down
GigabitEthernet1/0/3   unassigned      YES unset  down                  down
[fragment usunięty]
GigabitEthernet1/0/47  unassigned      YES unset  up                    up
GigabitEthernet1/0/48  unassigned      YES unset  up                    up
GigabitEthernet1/0/49  unassigned      YES unset  down                  down
GigabitEthernet1/0/50  unassigned      YES unset  down                  down
GigabitEthernet1/0/51  unassigned      YES unset  down                  down
GigabitEthernet1/0/52  unassigned      YES unset  down                  down
Port-channel1          unassigned      YES unset  down                  down
```

Pierwszy interfejs jest wirtualnym interfejsem przełączania dla VLAN-u 1. To SVI nie może być usunięte dla 3750, gdyż jest domyślnie wykorzystywane do zarządzania. Patrząc na tabele VLAN-ów dzięki poleceniu `show vlan`, możesz zobaczyć, że oprócz domyślnego VLAN-u 1 są skonfigurowane jeszcze trzy inne VLAN-y:

```
3750-IOS#sho vlan
VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
1    default                          active    Gi1/0/1, Gi1/0/2, Gi1/0/3
                                                Gi1/0/4, Gi1/0/5, Gi1/0/6
                                                Gi1/0/7, Gi1/0/8, Gi1/0/9
                                                Gi1/0/10, Gi1/0/11, Gi1/0/13
                                                Gi1/0/14, Gi1/0/15, Gi1/0/16
[fragment usunięty]
                                                Gi1/0/44, Gi1/0/45, Gi1/0/46
                                                Gi1/0/49, Gi1/0/50, Gi1/0/51
                                                Gi1/0/52
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
2    VLAN0002                         active    Gi1/0/12
3    VLAN0003                         active
4    VLAN0004                         active
VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
100  VLAN0100                         active
1002 fddi-default                     act/unsup
1003 token-ring-default               act/unsup
1004 fddinet-default                  act/unsup
1005 trnet-default                    act/unsup
```

Są to VLAN-y wyłącznie warstwy drugiej, dlatego też przełącznik nie może ustanowić routingu między nimi. Aby przełącznik miał dostęp do tych VLAN-ów w wyższych warstwach, musisz stworzyć interfejs SVI dla każdego z nich.

Żeby skonfigurować interfejsy SVI na przełączniku Nexus 7000, musisz najpierw włączyć tę funkcję globalną komendą `interface-vlan`.

Jedynym etapem wymaganym przy tworzeniu interfejsu SVI jest jego zdefiniowanie. SVI stworzysz, wpisując po prostu globalną komendę konfiguracji `interface vlan` *`vlan_nr`*. *`Vlan_nr`* nie musi odpowiadać numerowi żadnego z istniejących VLAN-ów. Przykładowo zdefiniowanie SVI dla VLAN-u 200, który nie występuje na naszym przełączniku, nadal spowoduje utworzenie tego interfejsu SVI. Możesz nawet przypisać temu interfejsowi adres IP oraz włączyć go poleceniem `no shutdown`:

```
3750-IOS(config)#int vlan 200
3750-IOS(config-if)#
00:11:40: %LINEPROTO-5-UPDOWN: Line protocol on Interface Vlan200, changed state to
down
3750-IOS(config-if)#ip address 10.200.0.1 255.255.255.0
3750-IOS(config-if)#no shut
```

Interfejs przybierze początkowo status *down/down* (nieaktywny/nieaktywny), ponieważ w warstwie drugiej nie ma jeszcze VLAN-u, który by go obsługiwał. Typ sprzętowy to **EtherSVI**, co wskazuje, że jest to logiczny interfejs SVI:

```
3750-IOS#sho int vlan 200
Vlan200 is down, line protocol is down
  Hardware is EtherSVI, address is 001c.b0a9.6541 (bia 001c.b0a9.6541)
  Internet address is 10.200.0.1/24
  MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input never, output never, output hang never
  Last clearing of "show interface" counters never
  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  5 minute input rate 0 bits/sec, 0 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
     0 packets input, 0 bytes, 0 no buffer
     Received 0 broadcasts (0 IP multicasts)
     0 runts, 0 giants, 0 throttles
     0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
     0 packets output, 0 bytes, 0 underruns
     0 output errors, 0 interface resets
     0 output buffer failures, 0 output buffers swapped out
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kiedy dodamy na przełączniku VLAN 200, interfejs się „podniesie”:

```
3750-IOS(config-if)#vlan 200
3750-IOS#sho ip int brie | include Vlan200
Vlan200                10.200.0.1      YES manual up                    up
```

Nie przypisaliśmy do tego VLAN-u żadnych portów, jednak interfejs SVI dla VLAN-u przyjął status *up* (aktywny) i operuje w warstwie trzeciej. Możemy nawet wykonać test `ping` nowego interfejsu:

```
3750-IOS#ping 10.200.0.1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.200.0.1, timeout is 2 seconds:
!!!!!
```

Gdybyśmy zechcieli, moglibyśmy dodać również protokół routingu i zrobić prawie wszystko, co można zrobić standardowo z interfejsem operującym w warstwie trzeciej.

Jaki jest cel posiadania interfejsu SVI bez przypisania do niego żadnych fizycznych portów? Jednym z przykładów może być stworzenie zarządzanej sieci innej niż VLAN 1 dla wszystkich Twoich urządzeń. Nie potrzebowałbyś przypisywać do VLAN-u żadnych fizycznych portów poza portem trunkowym do pozostałych urządzeń. W ten sposób mógłbyś zarządzać ruchem na wyodrębnionym VLAN-ie oddzielnie w stosunku do zarządzania ruchem w sieci produkcyjnej. Mógłbyś również podłączyć do przełącznika VLAN-y poprzez trunki, co pozwoliłoby wykorzystać przełącznik jako bramę domyślną, mimo że nie ma tu żadnych portów dedykowanych dla VLAN-u.

Tryb hybrydowy (4500, 6500)

W przełączniku obsługującym hybrydowy tryb IOS system IOS nie jest integralną częścią funkcjonalności przełącznika. CatOS wykorzystywany jest do przełączania, które przynależy do warstwy drugiej. IOS jest stosowany jedynie do zarządzania kartami MSFC oraz warstwą trzecią.

VLAN tworzysz w CatOS, a SVI dla tego VLAN-u definiujesz w IOS. Nazywasz VLAN w CatOS, ale opis dla jego interfejsu SVI dodajesz w IOS. Wszystko, co jest funkcjonalnością warstwy drugiej, musisz konfigurować w CatOS. Funkcje warstwy trzeciej i wyższych warstw muszą być konfigurowane w IOS.

W przełącznikach obsługujących tylko IOS zawsze istnieje wirtualny interfejs VLAN-u 1. Nie ma to jednak miejsca w przypadku przełączników działających w trybie hybrydowym, ponieważ VLAN 1 nie wymaga domyślnie funkcjonalności warstwy trzeciej.

Stworzyłem poniżej dwa VLAN-y na switchu 6509 pracującym w trybie hybrydowym:

```
CatOS-6509: (enable)sho vlan
VLAN Name                             Status    IfIndex Mod/Ports, Vlans
---- -------------------------------- --------- ------- ------------------------
1    default                          active    9       1/1-2
                                                        2/1-2
                                                        3/5-48
                                                        4/1-48
                                                        5/1-48
10   Lab-VLAN                         active    161
20   VLAN0020                         active    162     3/1-4
1002 fddi-default                     active    10
1003 token-ring-default               active    13
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
1004 fddinet-default                 active    11
1005 trnet-default                   active    12
```

Pierwszy z nich, VLAN 10, nazwałem `Lab-VLAN`. Drugi, VLAN 20, ma domyślną nazwę `VLAN0020`. Aby skonfigurować te VLAN-y w IOS, musimy najpierw połączyć się z kartą MSFC, stosując komendę `session`. Po komendzie musisz podać numer modułu, z którym chcesz się połączyć. Jeśli potrzebujesz sprawdzić, jaki jest numer modułu, skorzystaj z polecenia `show module`:

```
CatOS-6509: (enable)sho mod
Mod Slot Ports Module-Type               Model               Sub Status
--- ---- ----- ------------------------- ------------------- --- --------
1   1    2     1000BaseX Supervisor      WS-X6K-SUP2-2GE     yes ok
15  1    1     Multilayer Switch Feature WS-F6K-MSFC2        no  ok
2   2    2     1000BaseX Supervisor      WS-X6K-SUP2-2GE     yes standby
16  2    1     Multilayer Switch Feature WS-F6K-MSFC2        no  ok
3   3    48    10/100BaseTX Ethernet     WS-X6348-RJ-45      no  ok
4   4    48    10/100BaseTX Ethernet     WS-X6348-RJ-45      no  ok
5   5    48    10/100BaseTX Ethernet     WS-X6348-RJ-45      no  ok
```

Pierwsza karta MSFC zgłasza się w slocie 15. Jest to normalne, kiedy moduł zarządzający znajduje się w slocie 1, ponieważ MSFC jest „kartą-córką" w stosunku do modułu zarządzającego. Przełącznik przypisuje numer wewnętrznego slotu do karty MSFC. Teraz możemy się z nią połączyć:

```
CatOS-6509: (enable)session 15
Trying Router-15...
Connected to Router-15.
Escape character is '^]'.

MSFC-6509> en
MSFC-6509#
```

Innym sposobem połączenia się z MSFC jest użycie polecenia `switch console`.

Teraz, kiedy jesteśmy już w IOS, zobaczmy, co MSFC „sądzi" na temat naszych dwóch VLAN-ów:

```
MSFC-6509#sho ip int brief
Interface                  IP-Address      OK? Method Status                Protocol
```

Nie ma aktywnych interfejsów SVI na MSFC — nie jest aktywny nawet VLAN 1. Dodajmy SVI dla VLAN-u 20 i zobaczmy, co się stanie:

```
MSFC-6509#conf t
Enter configuration commands, one per line. End with CNTL/Z.
MSFC-6509(config)# int vlan 20
MSFC-6509(config-if)# ip address 10.20.20.1 255.255.255.0
MSFC-6509(config-if)# no shut
MSFC-6509(config-if)# ^Z
MSFC-6509#
17w2d: %LINK-3-UPDOWN: Interface Vlan20, changed state to down
17w2d: %LINEPROTO-5-UPDOWN: Line protocol on Interface Vlan20, changed state to down
MSFC-6509#
MSFC-6509# sho ip int brief
Interface                  IP-Address      OK? Method Status            Protocol
Vlan20                     10.20.20.1      YES manual down              down
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Teraz SVI jest już widoczny, ma jednak status *down* (nieaktywny). SVI nie „podniesie" się, dopóki nie ma aktywnego portu w tym VLAN-ie w warstwie drugiej. Często o tym zapominam i po dodaniu interfejsów SVI przechodzę do tworzenia VLAN-ów tylko po to, by zorientować się, że żaden z nich nie będzie aktywny. Aby to zilustrować, przypiszę w CatOS adres IP do interfejsu zarządzania SC0 i umieszczę go w VLAN-ie 20. To umieści aktywne urządzenie w VLAN-ie:

```
CatOS-6509: (enable)set int sc0 20 10.20.20.20 255.255.255.0
Interface sc0 vlan set, IP address and netmask set.
```

Teraz, kiedy jest już jakieś aktywne urządzenie w VLAN-ie 20, jego interfejs SVI będzie aktywny w karcie MSFC:

```
MSFC-6509#sho ip int brief
Interface                  IP-Address      OK? Method Status            Protocol
Vlan20                     10.20.20.1      YES manual up                up
```

NX-OS (Nexus 7000, 5000)

Konfiguracja interfejsów SVI dla Nexusa jest bardzo zbliżona do konfiguracji w IOS. Główną różnicą, którą zawsze napotykam przy pierwszej konfiguracji dla danego przełącznika, jest to, że domyślnie interfejsy SVI nie są dozwolone:

```
NX-7K-1(config)# int vlan 10
                 ^
Invalid command (interface name) at '^' marker. //Niewłaściwa komenda (nazwa interfejsu) od miejsca
                                                // oznaczonego '^'
```

Podobnie jak w przypadku wielu technologii na platformie Nexus obsługa SVI musi zostać najpierw włączona. Konkretna funkcja, którą należy uruchomić, to `interface-vlan`:

```
NX-7K-1(config)# feature interface-vlan
```

Kiedy już funkcja zostanie włączona, możesz dowolnie konfigurować SVI, podobnie jak w IOS:

```
NX-7K-1(config)# int vlan 10
NX-7K-1(config-if)# ip address 10.10.10.10.24
```

Modele przełączników wielowarstwowych

Cisco oferuje szereg modeli przełączników wielowarstwowych. Różnice między nimi są jednak często bardzo zatarte, ponieważ routery takie jak te z serii 7600 mogą współpracować z niektórymi modułami przełączającymi switchy z serii 6500. Routery z serii 3800 oraz 2800 również zapewniają wsparcie dla niewielkich modułów przełączających zdolnych obsłużyć wiele interfejsów Ethernet. Moduły te zachowują się właściwie tak, jakby były wyodrębnionymi przełącznikami 3750.

Mimo to nie ma wciąż żadnego magicznego urządzenia „wszystko w jednym". Musisz dokonać wyboru pomiędzy przełącznikiem z ograniczonymi funkcjami routera a routerem z ograniczonymi możliwościami przełącznika. Podstawową różnicą jest wewnętrzna struktura systemu oraz lista obsługiwanych modułów. Router zaprojektowany jest odmiennie od przełącznika, chociaż to stwierdzenie nie jest do końca prawdziwe, jeśli brać pod uwagę urządzenia takie jak CSR-1 (na szczęście przeciętny inżynier sieciowy prawdopodobnie nigdy nie natknie się na te potwory). Router zasadniczo nastawiony jest na struktury WAN, podczas gdy przełącznik

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

raczej na sieci LAN. Nie ma modułów dla przełączników 6500 umożliwiających połączenie WAN dla standardu T1[1]. Choć mógłbyś zastosować moduły Ethernet przełączników 6500 do routera z serii 7600, to wydajność magistrali routera nie będzie tak wysoka jak wydajność magistrali przełącznika.

Przełączniki wielowarstwowe można podzielić na podstawie typu obudowy (ang. *chassis*). Na najniższej półce znajdują się przełączniki w obudowie rackowej o wysokości 1 U, przeznaczone do węzłów dystrybucji okablowania oraz niewielkich instalacji. Niektóre z tych przełączników mogą być łączone w stos na kilka sposobów w zależności od modelu. Znajdą się też modele 1 U, które wyróżniają się zwiększoną prędkością magistrali, obsługującą nawet łącza uplink 10 Gb/s.

Następnymi w hierarchii są nieduże przełączniki typu *chassis-based*. Ta grupa składa się z modeli z serii 4500. Switche te zaprojektowane są do obsługi większych węzłów dystrybucji okablowania, a nawet do spełniania podstawowych funkcji na poziomie rdzenia sieci. Mogą obsługiwać wiele źródeł zasilania oraz kilka modułów zarządzających. Są przeznaczone do instalacji, które ma cechować wysoka dostępność.

Przełączniki 6500 zajmowały dotąd najwyższą półkę w przedstawionej klasyfikacji. Dostępne są w różnych wielkościach obudowy i są bardzo popularne z uwagi na możliwości ich rozbudowy, elastyczność oraz wydajność. Obecnie czołówkę w klasie przełączników stosowanych w centrach danych stanowi linia switchy Cisco Nexus.

Więcej informacji na temat rodzajów przełączników znajdziesz w rozdziale 2. Przełączniki Cisco z serii 6500 są omówione szczegółowo w rozdziale 17., a seria 3750 w rozdziale 19. Rozdział 18., opisujący przełączniki Nexus, również powinien dostarczyć Ci przyjemnej lektury.

Napiszę teraz kilka słów na temat wyboru właściwego urządzenia do realizacji określonego zadania. Wielu z moich klientów kupuje przełączniki, które określam mianem przełączników do szafy rackowej, i stosuje je w centrach danych. Doskonałym przykładem jest Cisco 3750. To bardzo atrakcyjny przełącznik, gdyż może być łączony w stos, co umożliwia uzyskanie całkiem przyzwoitego zagęszczenia portów. Musisz być jednak ostrożny, ponieważ ten model został w rzeczywistości zaprojektowany do podłączania użytkowników końcowych. Oczywiście zadaniem przełącznika jest przekazywanie pakietów i nie dba on o to, co jest do niego podłączone, ale 3750 posiada tylko jedno źródło zasilania. Można nieco zniwelować tę niedogodność, stosując systemy zasilania redundantnego, takie jak RPS (ang. *Redundant Power System*), oraz przez wykorzystanie *multichassis EtherChannel*. RPS jest jednak dodatkową jednostką zajmującą miejsce w szafie rackowej. Rozważ również, że 3750 obsługuje tylko 127 instancji Spanning Tree. W przypadku gdy domyślnie zainstalowane jest RPVST, nie będziesz mógł dodać VLAN-u 128, stosując jednocześnie Spanning Tree do eliminowania pętli. Zaufaj mi. Jeśli o tym wspominam, to z pewnością nie będzie to zabawna sytuacja, kiedy już się w niej znajdziesz.

Cisco posiada w ofercie przełączniki, które obsługują wymagania wyższej dostępności spotykane w centrach danych. Switche takie jak seria 4900 posiadają podwójne źródło zasilania i są zaprojektowane jako urządzenia nadrzędne, jeśli chodzi o montaż w szafie rackowej. Linia Nexus przeznaczona jest specjalnie do stosowania w centrach danych. Posiada też odpowiedni

[1] T1 jest standardem stosowanym w Ameryce Północnej; jego europejskim odpowiednikiem jest E1 — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

system chłodzenia z przepływem powietrza z przodu na tył oraz podwójne źródło zasilania. Przełączniki Nexus 5000 oraz 5500 umożliwiają podłączenie dodatkowych urządzeń FEX, co pozwala uzyskać zagęszczenie portów przekraczające możliwości switchy 6509. Niektóre modele z tych serii Nexusa nie obsługują jednak warstwy trzeciej.

Jeśli to możliwe, staraj się unikać stosowania niewłaściwych przełączników w celu zaoszczędzenia paru złotych. Inaczej może się to skończyć tym, że będziesz musiał opłacić kogoś takiego jak ja, żeby wybawił Cię z kłopotów, w które sam się wpakowałeś. Uwierz mi, zainwestowanie w dobry przełącznik wychodzi taniej.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 17.

Przełączniki wielowarstwowe Cisco 6500

Przełączniki z serii Cisco 6500 to prawdopodobnie najbardziej powszechnie stosowane przełączniki klasy *enterprise* na świecie. Odpowiednie wyważenie takich cech jak możliwości rozbudowy, moc, funkcjonalność oraz zagęszczenie portów sprawia, że jest to urządzenie, z którym trudno konkurować. Seria 6500 dostępna jest w różnych rozmiarach, od 3 do 13 slotów. Są nawet wersje zgodne z systemem NEBS (ang. *Network Equipment Building System*), przeznaczone do zastosowań w strukturach telekomunikacyjnych, w których obowiązują bardzo rygorystyczne wymagania odnośnie do specyfikacji sprzętu.

Architektura urządzeń 6500 nie jest nowa, została jednak zaprojektowana z myślą o zapewnieniu szerokich możliwości rozbudowy. Dzięki temu przełącznik, który obsługiwał pierwotnie prędkość płyty *backplane* 32 Gb/s, teraz standardowo obsługuje 720 Gb/s. Istnieje nawet sposób łączenia w klaster dwóch przełączników 6500, co umożliwia osiągnięcie teoretycznej prędkości 1440 Gb/s. Uważam jednak, że, jak zobaczysz później, jest to raczej chwyt marketingowy niż rzeczywista funkcjonalność.

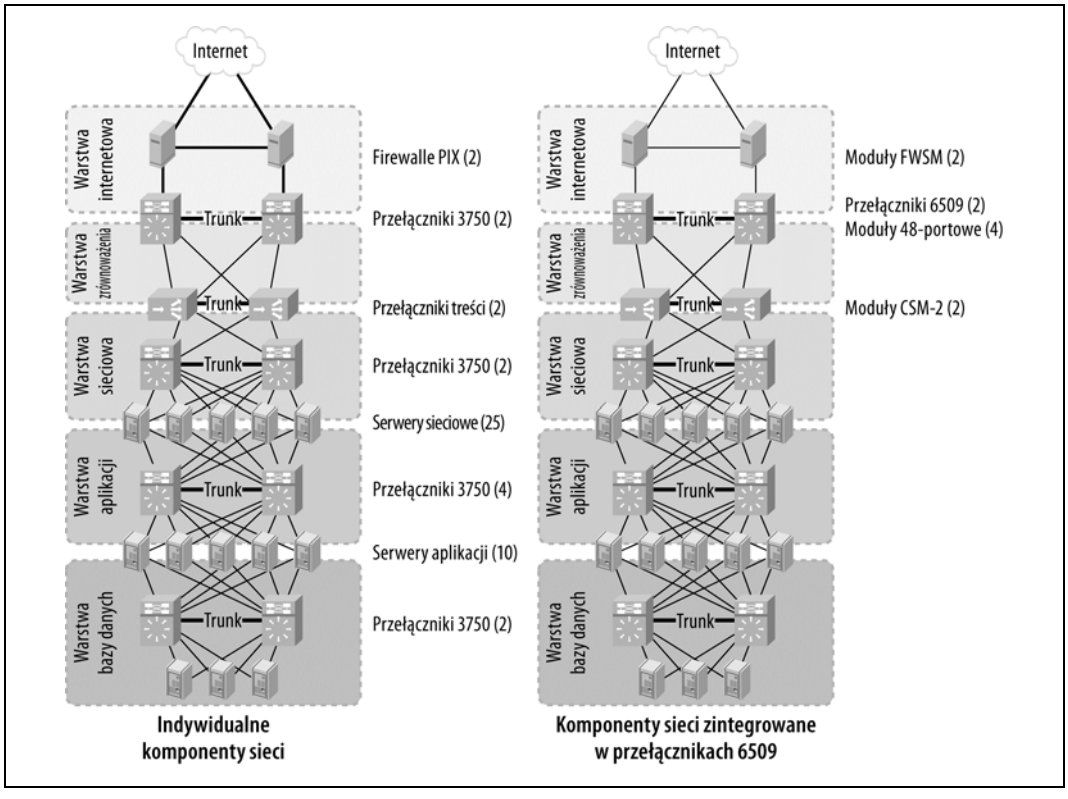
Wszechstronność platformy 6500 była zawsze motorem do lokowania urządzeń z tej serii w wielu różnych pozycjach w jeszcze szerszej gamie rozwiązań. Przełączniki 6509 są często stosowane w centrach sieci korporacyjnych, w warstwach dostępowych sieci dużych przedsiębiorstw, w funkcji rdzenia dla stron internetowych e-commerce, a nawet jako bramy telekomunikacyjne dla rozległych implementacji VoIP.

Podobnie rzecz ma się z elastycznością tej platformy, która spowodowała spopularyzowanie przełączników 6500 w mniejszych przedsiębiorstwach. Seria 6500 obejmuje moduły FWSM (ang. *Firewall Services Modules*), moduły CSM (ang. *Content Switching Modules*) oraz moduły NAM (ang. *Network Analysis Modules*). Cała infrastruktura sieci, z elementami takimi jak bezpieczeństwo, rozkładanie obciążenia oraz monitorowanie sprzętu, może być zawarta w pojedynczej obudowie.

Wraz z implementacją kart MSFC (dołączaną w najnowszych modułach zarządzających) urządzenia 6500 stały się przełącznikami wielowarstwowymi. Zastosowanie pracującego natywnie systemu IOS umożliwiło wykorzystanie 6500 jako routera z potencjałem ponad 300 interfejsów Ethernet przy zachowaniu funkcjonalności i prędkości przełącznika.

W trybie natywnym IOS przełącznik 6500 operuje w sposób bardzo zbliżony do mniejszych switchy 3550 oraz 3750, jednak z większą elastycznością i mocą. Rysunek 17.1 ilustruje, jakie korzyści odnieść może wielopoziomowa strona internetowa e-commerce dzięki zastosowaniu rozwiązania 6509 *chassis-based* zamiast szeregu poszczególnych elementów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 17.1. Indywidualne kontra zintegrowane komponenty sieci

Po pierwsze, ponieważ wszystkie warstwy mogą zostać skonsolidowane w pojedynczym urządzeniu wykorzystującym VLAN-y, większość przełączników nie jest już potrzebna i została zastąpiona modułami Ethernet w obudowie urządzenia 6509.

Po drugie, ponieważ niektóre warstwy nie potrzebują wielu portów, dostępne porty mogą być lepiej wykorzystane. Moduł nie jest przypisany do określonej warstwy, może więc zostać rozdysponowany w dowolny sposób w zależności od potrzeb. Jeśli przypisalibyśmy do każdej warstwy fizyczne przełączniki, mielibyśmy wiele niewykorzystanych portów na każdym przełączniku, zwłaszcza w wyższych warstwach.

Kolejną korzyścią jest ograniczenie liczby umów serwisowych ze względu na umieszczenie wszystkich komponentów w pojedynczej obudowie (choć niektóre moduły, jak FWSM lub ACE, wymagają oddzielnych umów serwisowych). Ponadto z uwagi na scentralizowanie wszystkich urządzeń wystarczą teraz jedynie dwa gniazda zasilania dla każdego przełącznika. Dostarczyć należy zasilanie co najmniej 220 V, 20 A.

Niektóre nowsze zasilacze do obudów 6500e wymagają więcej niż jednego kanału zasilania na jeden moduł. Zasilacz AC 6000W wymaga dwóch gniazd zasilania na moduł. Z kolei zasilacz AC 8700W będzie już wymagał trzech gniazd na pojedynczy moduł, co daje sześć gniazd zasilania na obudowę!

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Głównymi zaletami jest ułatwiona obsługa administracyjno-techniczna oraz prędkość. Każdy moduł jest dostępny bezpośrednio oraz z poziomu przełącznika, a w przypadku modułu ACE jego konfiguracja jest elementem Cisco IOS samego przełącznika (pod warunkiem, że jest to natywny IOS). Każdy moduł obsługuje funkcję *hot-swap*, czyli możliwość wymiany lub podłączenia modułu bez konieczności wyłączania całego przełącznika. Jedynym ograniczeniem jest tutaj konieczność wyłączenia niektórych modułów przed ich usunięciem. Ponadto moduły komunikują się między sobą poprzez płytę *backplane*, oferują więc znacznie większą prędkość niż ich samodzielnie działające odpowiedniki. Przykładowo moduł FWSM może osiągnąć przepustowość nawet ponad 5,5 Gb/s. Dopiero niedawno, wraz z wprowadzeniem ASA 5580-40, samodzielne urządzenia osiągnęły wyższe przepustowości, dochodzące aż do 10 Gb/s. Urządzenie ASA 5580-20 obsługiwało jedynie przepustowość 2,5 Gb/s. O ile ta różnica przetwarzania jest w dużej mierze związana z konstrukcją modułów FWSM oraz firewalli ASA, o tyle nie należy zapominać, że urządzenia samodzielne muszą komunikować się przez interfejsy Ethernet, podczas gdy moduły komunikują się bezpośrednio poprzez płytę *backplane* przełącznika. Stało się to mniejszym problemem, od kiedy urządzenia ASA obsługują Ethernet 10 Gb.

Przełączniki z serii 6500 zostały zaprojektowane tak, aby zapewniać wysoki stopień redundancji. Obsługują one podwójne adaptery zasilania oraz podwójne moduły zarządzające. Moduły zarządzające MSFC mogą działać jako indywidualne routery lub w trybie pojedynczego routera. Adaptery zasilania można skonfigurować w trybie *failover* albo jako działające równolegle. W tym drugim przypadku zapewniają większą moc na potrzeby modułów.

Jeśli połączysz funkcje tego przełącznika, skalowalność oraz elastyczność z dodatkową odpornością na problemy charakterystyczną dla sieci o wysokiej dostępności, otrzymasz na wyciągnięcie ręki światowej klasy architekturę. Mimo że przełączniki Nexus są najlepszym wyborem w centrach danych, to nadal nie są w stanie dorównać wszechstronności switchy 6500.

Architektura

Przełączniki 6500 wyewoluowały z serii 6000. Seria 6000 oferuje prędkość szyny płyty *backplane* na poziomie 32 Gb/s, podczas gdy seria 6500 posiada dodatkową szynę, zwaną **fabric bus** lub **crossbar switching bus**. Szyna ta umożliwia zwiększenie prędkości *backplane* do 720 Gb/s lub nawet powyżej tej wartości.

Zastosowanie struktury przełączania krzyżowego (ang. *crossbar switching*) w serii 6500 zapewnia niewiarygodny stopień elastyczności nowych obudów. Moduły z obudów 6000 mogą być wciąż wykorzystywane, ponieważ szyna 32 Gb/s w przełącznikach 6500 jest taka sama jak szyna w przełącznikach 6000. Jednak przy zastosowaniu modułu **SFM** (ang. *Switch Fabric Module*) nowsze moduły typu *fabric-enabled* mogą wykorzystać zalety nowej szyny.

Moduł SFM jest zasadniczo 16-portowym przełącznikiem, który łączy każdy moduł *fabric-enabled* poprzez szynę *fabric bus*. Ponieważ SFM przełącza do samego siebie, moduły komunikują się równocześnie, podobnie jak komputery komunikujące się poprzez przełącznik w sieci Ethernet. Dla kontrastu szyna 32 Gb/s działa w taki sposób, że wszystkie moduły otrzymują wszystkie pakiety, bez względu na to, jaki jest ich moduł docelowy (tak jak komputery komunikujące się poprzez hub w sieci Ethernet).

Ponieważ SFM kontroluje szynę *crossbar fabric bus*, może być umieszczany tylko w określonych slotach. Jedną z głównych wad tej konstrukcji jest to, że instalacje z wysokim poziomem redundancji wymagają dwóch slotów na moduły zarządzające i kolejnych dwóch na moduły SFM.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W przypadku obudowy z dziewięcioma slotami pozostawi to jedynie pięć slotów na karty rozszerzeń lub moduły usług.

Moduł zarządzający Supervisor-720 rozwiązuje problem dostępności slotów, ponieważ jest on zintegrowany z modułem SFM. Teraz wysoce elastyczne instalacje wymagają jedynie dwóch slotów dla modułów zarządzających. Jednak ponieważ Supervisor-720 posiada funkcjonalność SFM, musi zostać umieszczony w slotach przeznaczonych dla modułów SFM. Przykładowo w przypadku redundancji dla przełącznika 6509 moduły zarządzające Supervisor-2 umieszczone są w slotach 1 i 2, podczas gdy moduły SFM w slotach 5 i 6. Moduły Supervisor-720 powinny zostać zainstalowane w slotach 5 i 6, co zwalnia sloty 1 i 2 dla kart rozszerzeń lub modułów usług.

Szyny

W przełącznikach z serii 6500 płyta *backplane* składa się z czterech szyn opisanych poniżej:

D bus

Szyna danych (ang. *data bus*), czyli *D bus*, jest wykorzystywana przez chipset EARL do transmisji danych pomiędzy mudułami. Prędkość D bus to 32 Gb/s. Szyna danych jest współdzielona, tak jak standardowa sieć Ethernet, pod tym względem, że wszystkie moduły otrzymują wszystkie pakiety, które znalazły się na szynie. Kiedy ramki muszą zostać przekazane z portu jednego modułu na port drugiego modułu, przechodzą przez tę szynę, o ile szyna *crossbar fabric bus* nie jest zajęta i jest dostępna dla modułów.

R bus

Szyna wyników (ang. *results bus*), czyli *R bus*, obsługuje komunikację pomiędzy modułami oraz logikę przełączania dla modułów zarządzających. Prędkość *R bus* to 4 Gb/s.

C bus

Szyna sterująca (ang. *control bus*), czyli *C bus*, jest również czasem określana mianem **EOBC** (ang. *Ethernet Out-of-Band Channel*). Szyna sterująca wykorzystywana jest do komunikacji pomiędzy kartami rozszerzeń a procesorami zarządzania siecią w modułach zarządzających. Szyna sterująca jest właściwie siecią 100 Mb/s w półdupleksie. Kod kontroli linii pobierany jest do kart rozszerzeń podłączonych do tej szyny.

Crossbar fabric bus

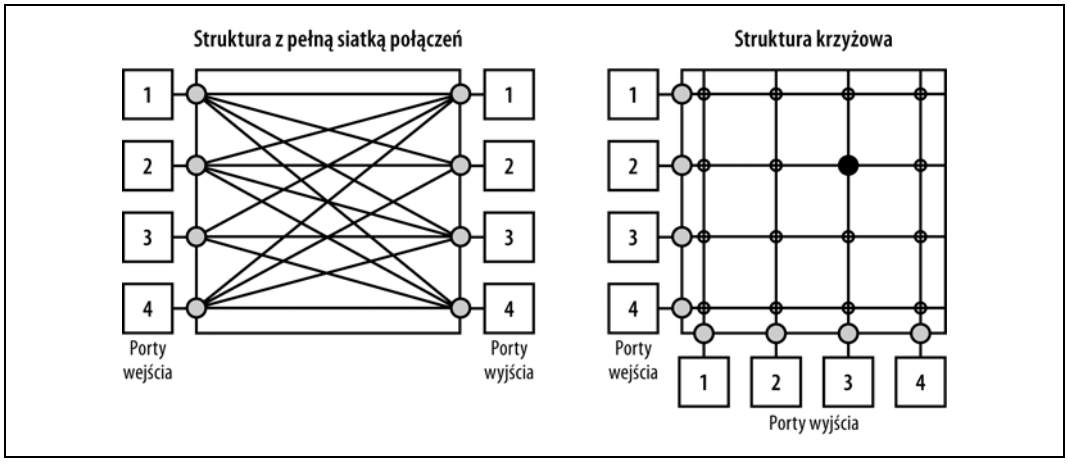
Przełączanie **krzyżowe** (ang. *crossbar*) to technologia przełączania, w której każdy z węzłów jest połączony z innymi węzłami przez krzyżujące się ścieżki. Alternatywną strukturą przełączania jest model obejmujący wszystkie możliwe połączenia, w którym każdy port jest bezpośrednio połączony z wszystkimi pozostałymi portami.

Rysunek 17.2 przedstawia schematy struktur przełączania. Termin **struktura** (ang. *fabric*) użyty jest do opisania siatki połączeń w takich schematach, ponieważ pod względem logicznym połączenia przypominają przeplatające się pasma tworzące strukturę siatki.

Struktura krzyżowa z rysunku 17.2 posiada jeden czarny punkt na przecięciu ścieżek. Jest to aktywne połączenie, podczas gdy reszta połączeń jest nieaktywna. Aktywne połączenie wskazuje, że port 2 komunikuje się z portem 3.

Szyna *crossbar fabric bus* w połączeniu z modułami Supervisor-2 oraz SFM osiąga przepustowość 256 Gb/s oraz 30 milionów pakietów na sekundę (Mp/s). W połączeniu z kartą dystrybucji danych (ang. *distributed forwarding card*) można osiągnąć nawet 210 Mp/s. Z modułem

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 17.2. Przykłady struktur przełączania

Supervisor-720 szyna ta obsługuje prędkość do 720 Gb/s. Jeśli wykorzystywane są moduły interfejsu dCEF (ang. *distributed Cisco Express Forwarding*), przełącznik 6500 wyposażony w Sup-720 może osiągać nawet 400 Mp/s.

Moduł SFM jest tym, co organizuje strukturę przełączania pomiędzy wszystkimi modułami typu *fabric-enabled* (pamiętaj, że funkcjonalność SFM jest zintegrowana w module Supervisor-720, więc w tym przypadku osobny moduł nie jest wymagany). Moduł jest właściwie sam w sobie przełącznikiem, który wykorzystuje szynę *fabric bus* na płycie *backplane* jako kanał komunikacyjny między modułami. Z tego powodu prędkość przełącznika 6500 może się zmienić w przypadku zastosowania nowszego modułu zarządzającego.

Rysunek 17.3 przedstawia model płyty *backplane* w obudowie przełącznika 6509. Patrząc na obudowę od frontu, zauważysz, że pokazane są złącza wszystkich slotów. W obudowie znajdują się dwa obwody płyty *backplane*, które są od siebie oddzielone. Po lewej stronie (patrząc z tyłu obudowy) znajdują się złącza zasilania oraz szyna *crossbar fabric bus*. Po prawej stronie szyny *D bus*, *R bus* oraz *C bus*. Obudowa przełącznika 6000 wygląda tak samo po prawej stronie, jednak po lewej nie ma szyny *crossbar fabric bus*.

Niektóre sloty przeznaczone są do określonych modułów, podczas gdy inne nie. Poniżej opisany jest rozkład slotów dla 6509:

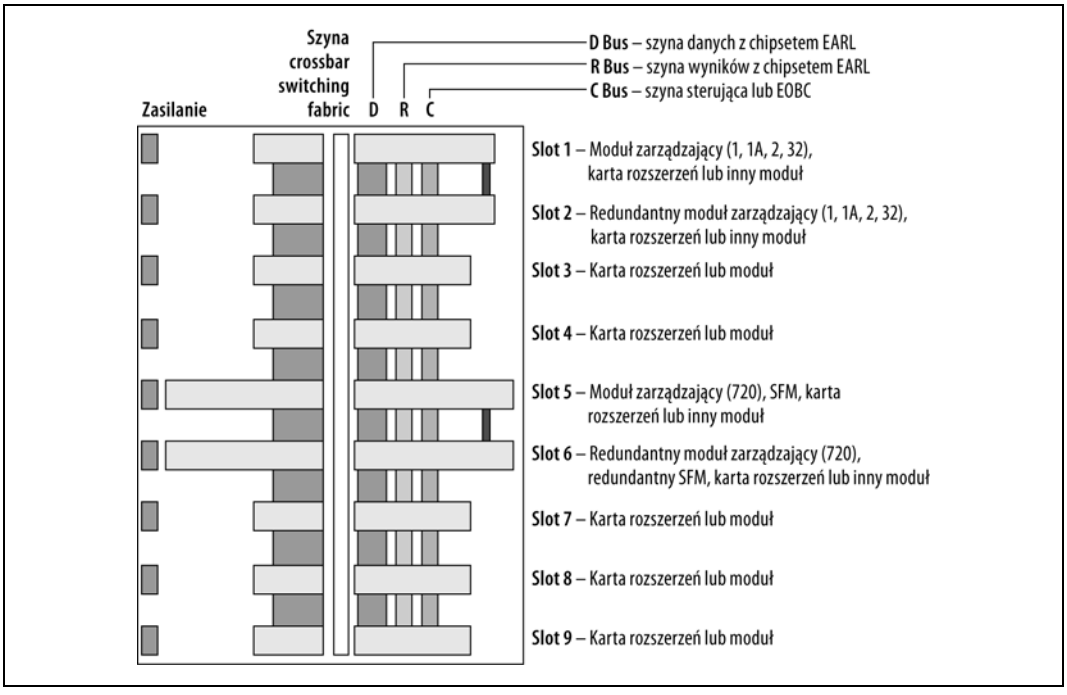
Slot 1

Slot 1 może być stosowany dla modułów zarządzających 1, 1A oraz 2, kart rozszerzeń lub modułów usług. Jeśli w obudowie znajduje się tylko jeden Sup-1, Sup-1A lub Sup-2, to musi być on umieszczony w slocie 1.

Slot 2

Slot 2 może być stosowany dla modułów zarządzających 1, 1A oraz 2, kart rozszerzeń lub modułów usług. Slot ten wykorzystywany jest dla redundantnego modułu zarządzającego, jeśli para modułów instalowana jest w trybie *failover*. Mimo że ten slot może posłużyć do zainstalowania pojedynczego modułu zarządzającego, to generalnie w takich przypadkach powinno się wybierać slot 1.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 17.3. Obwody płyty backplane przełącznika Cisco 6509

Slot 3

W tym slocie można zainstalować dowolną kartę rozszerzeń lub moduł, z wyjątkiem modułów zarządzających oraz modułów SFM.

Slot 4

W tym slocie można zainstalować dowolną kartę rozszerzeń lub moduł, z wyjątkiem modułów zarządzających oraz modułów SFM.

Slot 5

Slot 5 umożliwia zamontowanie modułu SFM lub modułu zarządzającego ze zintegrowanym modułem SFM, takiego jak Supervisor-720. Slot ten może również obsługiwać dowolne karty rozszerzeń lub moduły, z wyjątkiem modułu zarządzającego, który umieszczany jest w slocie 1 lub 2. W obudowie 6513 slot ten przeznaczony jest na karty rozszerzeń, a moduły zarządzające montowane są w slocie 7 i 8.

Slot 6

Slot 6 umożliwia zamontowanie modułu SFM lub modułu zarządzającego ze zintegrowanym modułem SFM, takiego jak Supervisor-720. Slot ten może również obsługiwać dowolne karty rozszerzeń lub moduły, z wyjątkiem modułu zarządzającego, który umieszczany jest w slocie 1 lub 2. Slot 6 przeznaczony jest dla redundantnego modułu SFM lub modułu zarządzającego w przypadku instalacji dwóch modułów w trybie *failover*. Mimo że pojedynczy moduł SFM lub moduł zarządzający może być instalowany w tym slocie, w tego typu instalacjach wykorzystuje się raczej slot 5. W obudowie 6513 slot ten przeznaczony jest na karty rozszerzeń, a moduły zarządzające montowane są w slocie 7 i 8.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Slot 7

Slot 7 służy do instalowania dowolnych kart rozszerzeń lub modułów, z wyjątkiem modułów zarządzających oraz modułów SFM. W obudowie 6513 slot ten przeznaczony jest dla modułów zarządzających i traktowany jest jak slot 5/6 w obudowie 6509.

Slot 8

Slot 8 służy do instalowania dowolnych kart rozszerzeń lub modułów, z wyjątkiem modułów zarządzających oraz modułów SFM. W obudowie 6513 slot ten przeznaczony jest dla modułów zarządzających i traktowany jest jak slot 5/6 w obudowie 6509.

Slot 9

Slot 9 służy do instalowania dowolnych kart rozszerzeń lub modułów, z wyjątkiem modułów zarządzających oraz modułów SFM.

W obudowie 6506 sloty są rozlokowane w ten sam sposób, z tą oczywistą różnicą, że nie ma slotów 7, 8 i 9. Sloty w obudowie 6513 są również rozmieszczone podobnie, jednak Sup-720 musi być instalowany w slotach 7 lub 8 zamiast w 5 lub 6. Dodatkowe sloty (10 – 13) mogą obsługiwać dowolne karty rozszerzeń lub moduły usług, z wyjątkiem modułów zarządzających oraz modułów SFM. Te ostatnie cztery sloty obudowy 6513 nie mogą obsługiwać określonych modułów *fabric-only blades*. Przy dobieraniu odpowiednich kart rozszerzeń dla tej obudowy zapoznaj się z dokumentacją Cisco, w której znajdziesz szczegółowe informacje na ten temat.

Ulepszone obudowy Enhanced Chassis

W ofercie dostępne są również ulepszone obudowy 6500, zwane *enhanced chassis*, które identyfikuje się literą **e** dodaną na końcu numeru modelu. Przykładem takiej obudowy jest 6500e. Obudowy *enhanced chassis* zaprojektowane są tak, aby umożliwić dostarczenie większej mocy dla kart rozszerzeń. Pojawienie się kart Power over Ethernet (PoE) dla aplikacji VoIP było jednym z kluczowych czynników wprowadzenia tej zmiany. *Enhanced chassis* posiadają wysokoobrotowe wentylatory umożliwiające schłodzenie tych energochłonnych modułów.

Obudowy *enhanced chassis* są wyposażone również w przeprojektowaną płytę *backplane*, która pozwala osiągnąć przepustowość 80 Gb/s na slot. Jest to teoretyczne podwojenie standardowej przepustowości dla 6500 *chassis* (40 Gb/s na slot). Jednak w tej chwili nie ma jeszcze dostępnych kart rozszerzeń lub modułów zarządzających obsługujących taką prędkość. Ta nowa architektura umożliwi umieszczenie na module ośmiu portów 10 Gb/s bez przekroczenia limitu emisji. Cisco produkuje obecnie tylko modele *enhanced chassis*, jednak standardowe modele są wciąż dostępne na rynku wtórnym.

Vertical Enhanced Chassis

Obecnie w ofercie dostępna jest specjalna wersja 6500 *enhanced chassis*. Model o nazwie 6509-V-E obsługuje przepustowość 80 Gb/s na slot i posiada funkcje charakterystyczne jak dotąd tylko dla przełączników NEBS *chassis*. Obudowa ta cechuje się przepływem powietrza z przodu na tył oraz pionowym usytuowaniem modułów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Moduły zarządzające

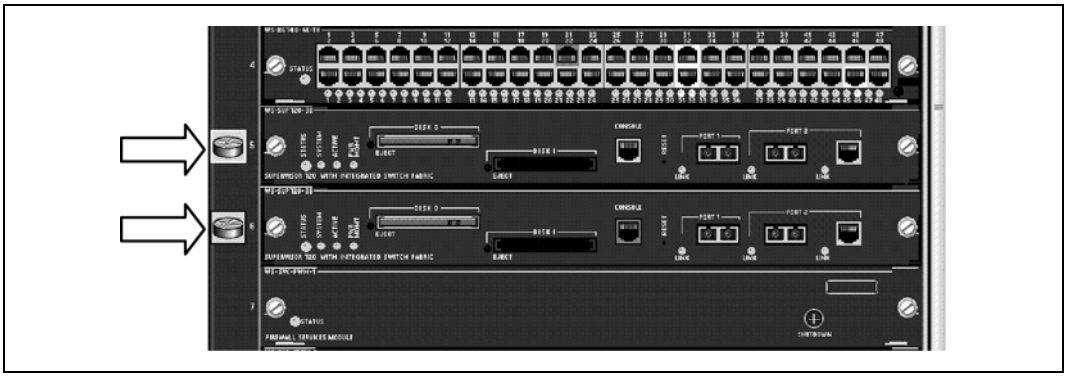
Przełączniki *chassis-based* nie posiadają zintegrowanych procesorów, tak jak ma to miejsce w przypadku mniejszych przełączników. Zamiast tego procesory montowane są w modułach, co umożliwia łatwą wymianę oraz unowocześnienie sprzętu. Procesor dla przełączników Cisco *chassis-based* zwany jest modułem zarządzającym (ang. *supervisor*). Stosuje się dla niego często skróconą nazwę **sup**.

Z biegiem lat wprowadzano kolejne modele modułów zarządzających, które oferowały większą prędkość oraz wszechstronność. Wzrost funkcjonalności osiągano również dzięki dodatkowym kartom rozszerzeń, które w późniejszych modelach modułów zarządzających były zintegrowane z tymi modułami.

MSFC

Moduły zarządzające oferują przetwarzanie w warstwie drugiej modelu ISO/OSI, podczas gdy karta-córka — zwana MSFC (ang. *Multilayer Switch Feature Card*) — obsługuje warstwę trzecią oraz posiada większą funkcjonalność. Modele 1 i 2 modułów zarządzających obsługują MSFC jako dodatkowe karty, podczas gdy późniejsze modele posiadają zintegrowany moduł MSFC.

W trybie hybrydowym IOS dla 6500 chassis moduł MSFC jest traktowany jako oddzielne urządzenie bez względu na zastosowany model modułu zarządzającego. W interfejsie graficznym CiscoView moduł MSFC reprezentowany jest przez małą ikonę routera na lewo od modułu zarządzającego, tam gdzie znajduje się kaseta z wentylatorami. Rysunek 17.4 przedstawia schemat CiscoView obrazujący moduł Supervisor-720 z modułem MSFC po lewej stronie.



Rysunek 17.4. Supervisor-720 oraz MSFC w CiscoView

Różne wersje modułu MSFC oznaczane są MSFC1, MSFC2 i MSFC3. MSFC2 jest stosowany w parze z modułem Supervisor-2, podczas gdy MSFC3 jest częścią Supervisor-720.

PFC

Karta PFC (ang. *policy feature card*) jest kartą rozszerzeń, która obsługuje sprzętowe funkcje QoS, zasadniczo zwiększając wydajność tam, gdzie potrzebny jest QoS. PFC nie wymaga żadnej bezpośredniej konfiguracji. Istnieją trzy generacje PFC, zwane PFC1, PFC2 i PFC3. PFC2 jest stosowana w parze z modułem Supervisor-2, podczas gdy PFC3 jest zintegrowana z modułem Supervisor-720.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Modele

Powszechnie spotykane dzisiaj modele modułów zarządzających to:

Supervisor-1A

Jest to najwolniejszy i najstarszy moduł zarządzający obsługujący przepustowość 32 Gb/s oraz 15 Mp/s (Supervisor-1A zastąpił pierwotny Supervisor Engine, zwany również Supervisor-1). Supervisor-1A nie jest już produkowany, jednak nadal można go spotkać w starszych instalacjach. W połączeniu z PFC oraz MSFC, Sup-1A umożliwia przekazywanie pakietów z warstwy drugiej do warstwy czwartej, oferując przy tym wyższy poziom bezpieczeństwa oraz QoS. Supervisor-1A był doskonałym rozwiązaniem dla węzłów dystrybucji okablowania dla sieci, które nie wymagały przepustowości oraz prędkości modułu Supervisor-2. Jeśli natrafisz obecnie na taki moduł wciąż pracujący w przełączniku, powinieneś go wymienić.

Supervisor-2

Model ten był standardem w przełączaniu szkieletowym oraz przy stronach internetowych e-commerce do czasu wprowadzenia modułu Supervisor-720. Supervisor-2 oferuje przepustowość 30 Mp/s przy prędkości 256 Gb/s, jeśli zastosowany jest w parze z modułem SFM. W połączeniu z PFC2 i MSFC2 zdolność przekazywania modułu Supervisor-2 wzrasta do 210 Mp/s. Umożliwia wtedy również przekazywanie z warstwy drugiej do warstwy czwartej, oferując przy tym lepszy poziom bezpieczeństwa oraz QoS. Ten moduł również jest obecnie dość przestarzały i jeśli to możliwe — powinien zostać wymieniony.

Supervisor-32

Jest to najnowszy zamiennik dla modułu Supervisor-1A. Wszystkie pracujące nadal moduły zarządzające Supervisor-1A powinny zostać zastąpione modułami Supervisor-32. Model różni się od innych modułów zarządzających tym, że posiada osiem portów Mini-GBIC 1 Gb/s lub 2 porty 10 Gb/s Ethernet XENPAK. Pozostałe moduły zarządzające oferują co najwyżej dwa porty 1 Gb/s.

Supervisor-720

Model ten jest znaczącym unowocześnieniem w stosunku do przestarzałej architektury Supervisor-2. Oferuje 400 Mp/s oraz aż 720 Gb/s i przeznaczony jest dla instalacji wymagających szerokiego pasma transmisji oraz krytycznych implementacji warstwy rdzenia. Supervisor-720 posiada zintegrowane moduły PFC3 oraz MSFC3, dostarczaja również nowe, przyspieszone funkcje CEF i dCEF. Moduły typu *fabric-only* w połączeniu z Sup-720 zapewniają przepustowość 40 Gb/s.

Supervisor-720-10G VSS

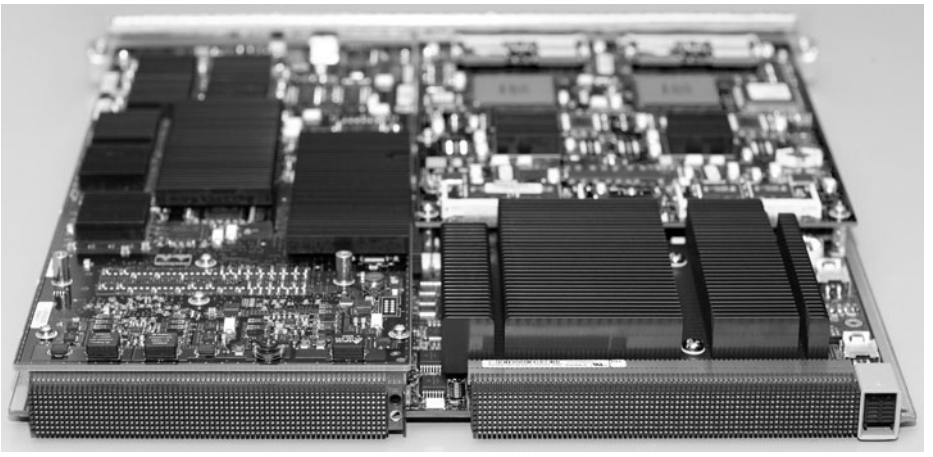
Ten moduł zarządzający jest zbliżony do Sup-720, posiada jednak dwa porty 10 Gb Ethernet. Obsługuje również nową funkcję, zwaną VSS (ang. *Virtual Switching System*). VSS umożliwia połączenie klastrem dwóch przełączników 6500, tak że tworzą one jeden logiczny switch. Podobne co do zasady do technologii łączenia w stos przełączników Cisco 3750, VSS umożliwia dzielenie kanałów EtherChannel pomiędzy dwa fizyczne przełączniki. Szczegóły na temat VSS znajdziesz w dalszej części rozdziału.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Moduły

Moduły dla 6500 *chasiss* mogą obsługiwać jeden lub obydwa obwody płyty *backplane*. Moduł, który nie obsługuje szyny *crossbar fabric*, jest nazywany modułem *nonfabric-enabled*. Moduł obsługujący szynę 32 Gb/s *D bus* oraz szynę *fabric bus* określany jest jako *fabric-enabled*. Jeśli moduł obsługuje tylko *fabric bus* i nie ma połączenia z *D bus*, uznawany jest za moduł typu *fabric-only*.

Moduły zarządzające nie posiadają takich samych złączy do połączenia z płytą *backplane* co moduły SFM. Moduły Supervisor-720 z funkcjonalnością SFM mają większe złącza, pasujące jedynie do gniazd w slotach 5 i 6. Złącza Sup-720 pokazane są na rysunku 17.5.



Rysunek 17.5. Złącza modułu zarządzającego Supervisor-720

Moduły *nonfabric-enabled* posiadają złącza tylko po jednej stronie, służące do połączenia z szyną *D bus*. Wszystkie moduły przełączników 6000 *chassis* są *nonfabric-enabled*, ponieważ seria 6000 nie posiada szyny *crossbar fabric bus*.

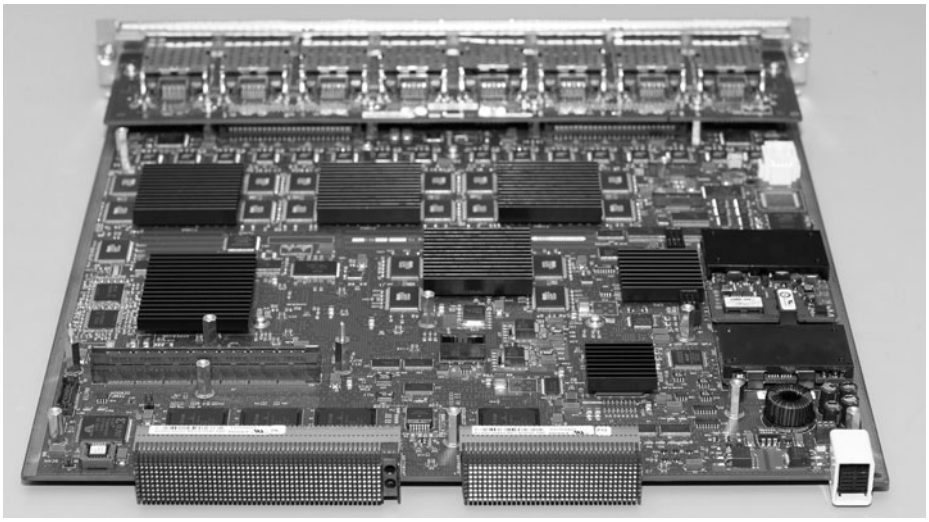
Moduł typu *fabric-enabled* posiada dwa złącza na tylnym panelu kasety: jedno dla *D bus*, a drugie dla *crossbar fabric bus*.

Przykład takiego podwójnego złącza (w tym przypadku 16-portowy gigabitowy moduł światłowodowy) przedstawia rysunek 17.6.

Moduły typu *fabric-only* posiadają tylko jedno złącze po stronie szyny *fabric bus*. Nie mają one złącza po stronie *D bus*.

Bądź ostrożny podczas montażu modułów w przełącznikach *chassis-based*, takich jak seria 6500. Większość komponentów znajdujących się na płycie modułu jest na tyle wysoka, że przy niewłaściwym lub siłowym montażu mogą one ulec zniszczeniu w wyniku uderzenia o ściany wewnątrz gniazd obudowy. Moduły zarządzające oraz moduły usług, takie jak CSM, ACE lub FWSM, są szczególnie podatne na uszkodzenie, ponieważ zawierają one znaczną liczbę komponentów na płycie urządzenia. Ceny detaliczne niektórych z tych modułów mogą sięgać nawet 150 000,00 zł, więc jak przypuszczam, nie chciałbyś być tym, który musi się przyznać do ich uszkodzenia.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 17.6. Złącza modułu fabric-enabled

Współpraca między modułami

Kiedy moduły *fabric-enabled* lub *fabric-only* umieszczone są w obudowie przełącznika łącznie z modułami *nonfabric-enabled*, moduł zarządzający musi pójść na kompromis, tak aby umożliwić współpracę pomiędzy różnymi szynami. W szczególnym przypadku, jeśli w obudowie znajduje się moduł *nonfabric-enabled*, Supervisor-720 nie będzie w stanie pracować z prędkością 720 Gb/s.

Poniżej przykład przełącznika 6509, na którym zainstalowano moduły typu *fabric-only* model 6748 Ethernet 10/100/1000 Mb oraz dwa moduły zarządzające Sup-720:

```
6509-1#sho mod
Mod Ports Card Type                                  Model              Serial No.
--- ----- ------------------------------------- ------------------ -----------
  1   48  CEF720 48 port 10/100/1000mb Ethernet  WS-X6748-GE-TX     SAL05340V5X
  2   48  CEF720 48 port 10/100/1000mb Ethernet  WS-X6748-GE-TX     SAL09347ZXK
  3   48  CEF720 48 port 10/100/1000mb Ethernet  WS-X6748-GE-TX     SAL05380V5Y
  4   48  CEF720 48 port 10/100/1000mb Ethernet  WS-X6748-GE-TX     SAL092644CJ
  5    2  Supervisor Engine 720 (Active)         WS-SUP720-3B       SAL05304AZV
  6    2  Supervisor Engine 720 (Hot)            WS-SUP720-3B       SAL09295RWB
  7   48  CEF720 48 port 10/100/1000mb Ethernet  WS-X6748-GE-TX     SAL05340Z9H
  8   48  CEF720 48 port 10/100/1000mb Ethernet  WS-X6748-GE-TX     SAL0938145M
  9   48  CEF720 48 port 10/100/1000mb Ethernet  WS-X6748-GE-TX     SAL053415EC
```

Polecenie `show fabric switching-mode` obrazuje, jak każdy z modułów komunikuje się z systemem. Na listingu można zauważyć, że wszystkie moduły wykorzystują szynę *crossbar switching bus*, a Sup-720 pracuje w trybie dCEF, który umożliwia przekazywanie pakietów z prędkością 720 Gb/s:

```
6509-1#sho fabric switching-mode
Fabric module is not required for system to operate
Modules are allowed to operate in bus mode
Truncated mode allowed, due to presence of aCEF720  module

Module Slot     Switching Mode
    1                 Crossbar
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
    2                Crossbar
    3                Crossbar
    4                Crossbar
    5                    dCEF
    6                Crossbar
    7                Crossbar
    8                Crossbar
    9                Crossbar
```

Każdy z modułów *fabric-only* posiada dwa połączenia 20 Gb/s do szyny *crossbar fabric bus*, co można zobaczyć, stosując komendę `show fabric status` lub `show fabric utilization`. Zauważ, że każdy z modułów zarządzających posiada tylko jedno połączenie 20 Gb/s do *fabric bus*:

```
6509-1#sho fabric util
 slot    channel      speed    Ingress %     Egress %
    1          0        20G            1            0
    1          1        20G            0            2
    2          0        20G            1            0
    2          1        20G            0            0
    3          0        20G            1            0
    3          1        20G            0            0
    4          0        20G            0            0
    4          1        20G            0            0
    5          0        20G            0            0
    6          0        20G            0            0
    7          0        20G            0            0
    7          1        20G            0            0
    8          0        20G            0            0
    8          1        20G            0            0
    9          0        20G            0            0
    9          1        20G            0            0
```

Dla porównania przedstawię przełącznik 6509 z dwoma modułami Supervisor-720, z jednym modułem *fabric-only*, dwoma modułami *fabric-enabled* oraz jednym modułem *nonfabric-enabled*:

```
6509-2#sho mod
Mod Ports Card Type                                Model              Serial No.
--- ----- ------------------------------------- ------------------ -----------
  1   48  CEF720 48 port 10/100/1000mb Ethernet  WS-X6748-GE-TX     SAL04654F2K
  4    8  Network Analysis Module                WS-SVC-NAM-2       SAD093002B6
  5    2  Supervisor Engine 720 (Active)         WS-SUP720-3B       SAL0485498A
  6    2  Supervisor Engine 720 (Hot)            WS-SUP720-3B       SAL09358NE6
  7    6  Firewall Module                        WS-SVC-FWM-1       SAD042408DF
  8    4  CSM with SSL                           WS-X6066-SLB-S-K9  SAD094107YN
  9    8  Intrusion Detection System             WS-SVC-IDSM-2      SAD048102CG
```

Moduł w slocie 1 jest taki sam jak moduły Ethernet w poprzednim przykładzie. Jest to moduł *fabric-only*. Moduły 4, 7 oraz 9 są *fabric-enabled*, a moduł 8 jest *nonfabric-enabled*. Listing z polecenia `show fabric switching-mode` pokazuje, że pojedynczy moduł *nonfabric-enabled* spowodował, iż moduł zarządzający przełączył się na niższą prędkość:

```
6509-2#sho fabric switching-mode
Global switching mode is Truncated
dCEF mode is not enforced for system to operate
Fabric module is not required for system to operate
Modules are allowed to operate in bus mode
Truncated mode is allowed, due to presence of aCEF720, Standby supervisor module

Module Slot     Switching Mode
    1                Crossbar
    4                Crossbar
    5                     Bus
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
6                Crossbar
7                Crossbar
8                     Bus
9                Crossbar
```

Moduł, o który chodzi w tym przykładzie, to CSM. Przed wprowadzeniem modułu ACE był to jeden z najdroższych dostępnych modułów. Pamiętaj, że cena nie jest równoznaczna z osiąganą prędkością. CSM był doskonałym urządzeniem i szczerze zalecałem jego stosowanie w sytuacjach, gdy rozkładanie obciążenia było konieczne. Jednak w instalacji wymagającej dużej przepustowości moduł ten mógł stać się wąskim gardłem. W przypadku architektury strony internetowej niezwykle rzadko zdarza się, żeby w kierunku interfejsu zewnętrznego (ang. *frontend*) płynął strumień danych większy niż 32 Gb/s. Wyższa przepustowość jest możliwa w przypadku równoważenia dużych farm serwerów aplikacji lub baz danych od strony wewnętrznej (ang. *backend*). Moduły ACE zastąpiły CSM w funkcji rozkładania obciążenia, ale przykład ten pokazuje, jak starsze i wolniejsze moduły wpływają na działanie całego przełącznika.

Zastosowanie na tym przełączniku komendy `show fabric status` pokazuje, że nie wszystkie moduły *fabric-enabled* są traktowane tak samo. Moduł *fabric-only* w slocie 1 posiada dwa kanały 20 Gb/s do szyny *fabric bus*. Moduł analizy sieci (ang. *Network Analysis Module* — NAM) w slocie 4 jest *fabric-enabled*, jednak łączy się tylko jednym kanałem 8 Gb/s podobnie jak moduły FWSM oraz IDS w slotach 7 i 9:

```
6509-2#sho fabric status
 slot    channel      speed    module              fabric
                               status              status
    1          0        20G        OK                  OK
    1          1        20G        OK                  OK
    4          0         8G        OK                  OK
    5          0        20G        OK                  OK
    6          0        20G        OK                  OK
    7          0         8G        OK                  OK
    9          0         8G        OK                  OK
```

Przykład ten ma na celu pokazanie, jak ważne jest zrozumienie sposobu, w jaki współpracują ze sobą moduły. Mimo że moduł może być „fabric blade", nie zawsze musi funkcjonować w taki sam sposób jak inny moduł *fabric-enabled*. Poznanie różnic w działaniu poszczególnych modułów może pomóc Ci zrozumieć bieżącą konfigurację oraz przydać się w przyszłości przy projektowaniu rozwiązań.

Rodzaje modułów

Moduły dla przełączników z serii 6500 dzielą się zasadniczo na karty rozszerzeń oraz moduły usług. Karty rozszerzeń oferują złącza takie jak miedziany lub światłowodowy Ethernet. Moduły usług oferują dodatkowe funkcjonalności. Przykładem modułów usług są FWSM oraz ACE.

Moduły usług dramatycznie zwiększają użyteczność przełącznika 6500. W jednej obudowie możesz zawrzeć kompletną architekturę serwera sieciowego łącznie z portami Ethernet, kanały internetowe DS3[1], firewalle, IDS oraz rozkład obciążenia. Wszystkie te urządzenia będą konfigurowane z poziomu pojedynczej obudowy, będą też zasilane z jednej obudowy. Aby zapewnić redundancję, można zastosować dwie identycznie skonfigurowane obudowy z funkcjonalnością *failover*, co wyeliminuje wszystkie pojedyncze punkty awarii.

[1] Jest to standard stosowany w Ameryce Północnej; jego europejskim odpowiednikiem jest sieć E3 — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Moduły Ethernet

Moduły Ethernet dostępne są w wielu odmianach i prędkościach. Niektóre oferują standardową łączność, podczas gdy inne zapewniają ekstremalne prędkości 40 Gb/s lub nawet 80 Gb/s dla połączeń z szyną *crossbar fabric bus*.

Opcje dostępnych złączy w modułach Ethernet obejmują RJ-45, GBIC, small-form-factor GBIC, XENPAK, SFP, a nawet złącza Amphenol dla bezpośrednich podłączeń do patch paneli. Zagęszczenie portów waha się od 4-portowych modułów 10 Gb/s XENPAK, przez 48-portowe moduły 1000 Mb/s RJ-45, a nawet 96-portowe moduły RJ-21 obsługujące 10/100 Mb/s. Opcjonalnie dostępne są funkcje PoE oraz dCEF.

FWSM

Moduły FWSM zapewniają usługi typu firewall podobnie jak samodzielne urządzenia PIX firewall. Obecnie FWSM to jedyne urządzenie wciąż dostępne w linii produktów PIX. Wszystkie pozostałe firewalle PIX zostały zastąpione modelami ASA oferującymi większą prędkość lub szerszy zestaw funkcji albo obie te rzeczy naraz. Ponieważ interfejsy w module FWSM są interfejsami wirtualnymi SVI, FWSM nie jest ograniczony do fizycznych połączeń, jak ma to miejsce w przypadku urządzeń samodzielnych. FWSM może posiadać kilkaset interfejsów, z których każdy przypisany jest do określonego VLAN-u na przełączniku. FWSM oferuje również przepustowość ponad 5 Gb/s w porównaniu z przepustowością 1,7 Gb/s urządzenia PIX 535, które było niegdyś jednym z najlepszych urządzeń z tej linii. Obecnie nie jest ono już tak imponujące, biorąc pod uwagę osiągi ASA 5580-40, które obsługuje 10 Gb/s.

FWSM obsługuje wiele kontekstów. Umożliwia to konfigurowanie wirtualnych firewalli, które mogą pełnić różne funkcje lub być obsługiwane przez różne podmioty. Mogą też spełniać oba te zadania równocześnie. Dobrym przykładem wykorzystania takiego rozwiązania jest sytuacja, w której usługodawca chce zapewnić indywidualne firewalle klientom, a posiada tylko jedno fizyczne urządzenie. Niegdyś była to cecha charakterystyczna jedynie dla modułów FWSM, jednak obecnie wiele urządzeń ASA również obsługuje konteksty.

FWSM jest wyodrębnionym urządzeniem w obudowie. Aby zarządzać FWSM, musisz się najpierw z nim połączyć. Poniżej konfiguracja połączenia z modułem FWSM znajdującym się w slocie 8:

```
Switch-IOS#session slot 8 proc 1
The default escape character is Ctrl-^, then x.
You can also type 'exit' at the remote prompt to end the session
Trying 127.0.0.81 ... Open

User Access Verification

Password:
Type help or '?' for a list of available commands.
Switch-IOS-FWSM > en
Password: ********
Switch-IOS-FWSM #
```

Jeśli FWSM pracuje w trybie pojedynczego kontekstu (ang. *single-context mode*), możesz stosować wszystkie komendy dostępne w urządzeniach PIX. Jeśli FWSM działa w trybie wielu kontekstów (ang. *multiple-context mode*), znajdujesz się domyślnie w kontekście systemowym (ang. *system context*). Musisz więc przełączyć się na odpowiedni kontekst, aby wprowadzić zmiany. Służy do tego komenda `changeto context`:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Switch-IOS-FWSM#sho context
Context Name      Class      Interfaces           URL
 admin            default                         disk:/admin.cfg
*EComm            default    vlan20,30            disk:/Ecomm.cfg
Switch-IOS-FWSM# changeto context EComm
Switch-IOS-FWSM/EComm#
```

W tym momencie znajdujesz się w kontekście EComm. Jeśli jesteś przyzwyczajony do firewalli PIX, wszystko powinno wyglądać znajomo:

```
Switch-IOS-FWSM/EComm#sho int
Interface Vlan20 "outside", is up, line protocol is up
        MAC address 0008.4cff.b403, MTU 1500
        IP address 10.1.1.1, subnet mask 255.255.255.0
                Received 90083941155 packets, 6909049206185 bytes
                Transmitted 3710031826 packets, 1371444635 bytes
                Dropped 156162887 packets
Interface Vlan30 "inside", is up, line protocol is up
        MAC address 0008.4cff.b403, MTU 1500
        IP address 10.10.10.1, subnet mask 255.255.255.0
                Received 156247369908 packets, 214566399699153 bytes
                Transmitted 2954364369 packets, 7023125736 bytes
                Dropped 14255735 packets
```

CSM

Moduły Cisco CSM są doskonałą alternatywą dla samodzielnych przełączników zawartości. CSM obsługuje przepustowość 4 Gb/s i jest dostępny w połączeniu z kartą rozszerzeń SSL akcelerator.

CSM został zastąpiony nowszym modułem ACE typu *fabric-enabled*.

Zintegrowanie CSM z przełącznikiem 6500 pracującym w trybie natywnym IOS jest bardzo proste. Wszystkie komendy CSM zawarte są w interfejsie poleceń (CLI) przełącznika. Polecenia dostępne są po wpisaniu komendy `module CSM` *`moduł_nr`*. Powoduje ona przejście w tryb konfiguracji `module contentswitchingmodule` *`muduł_nr`*:

```
Switch-IOS (config)#mod csm 9
Switch-IOS (config-module-csm)#
```

Jedną z większych wad modułów CSM jest to, że nie są one modułami *fabric-enabled*. Choć nie jest to problem, biorąc pod uwagę przepustowość dla samego modułu, stanowi to kłopot, kiedy przełącznik zawierający moduł CSM służy również do równoważenia obciążenia serwerów. CSM jest modułem 32 Gb/s. Umieszczenie go w przełączniku wykorzystującym obwód *fabric* na płycie *backplane* spowoduje, że moduł zarządzający powróci do pracy w trybie szyny zamiast szybszego trybu dCEF. Przełącznik z modułem Supervisor-720, modułami Ethernet *fabric-only* oraz modułem CSM nie będzie mógł pracować z prędkością 720 Gb/s, ponieważ moduł CSM ma ograniczone połączenia z płytą *backplane*.

Moduły CSM mogą pracować w trybie *failover* utrzymującym informacje o stanie aplikacji (ang. *stateful*). Para modułów CSM może synchronizować swoje konfiguracje, pod warunkiem, że posiadają wersję systemu 4.2(1) lub nowszą. Do synchronizacji służy komenda `hw-module csm` *`moduł_nr`* `standby config-sync`:

```
Switch-IOS#hw-module csm 9 standby config-sync
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Switch-IOS#
May  5 17:12:14: %CSM_SLB-6-REDUNDANCY_INFO: Module 9 FT info: Active: Bulk sync
started
May  5 17:12:17: %CSM_SLB-4-REDUNDANCY_WARN: Module 9 FT warning: FT configuration
might be out of sync.
May  5 17:12:24: %CSM_SLB-4-REDUNDANCY_WARN: Module 9 FT warning: FT configuration
back in sync
May  5 17:12:26: %CSM_SLB-6-REDUNDANCY_INFO: Module 9 FT info: Active: Manual bulk
sync completed
```

NAM

Moduł Cisco NAM jest zasadniczo sondą do zdalnego monitoringu (RMON) oraz urządzeniem przechwytującym pakiety. Umożliwia on monitorowanie dowolnego portu lub VLAN-u albo połączenie tych dwóch funkcji podobnie jak w przypadku zewnętrznego urządzenia do przechwytywania pakietów.

Moduł NAM jest sterowany za pomocą przeglądarki, co może być uciążliwe, gdy analizujesz duże pliki z przechwytywania pakietów. Zaletą takiego rozwiązania jest to, że nie jest wymagane dodatkowe oprogramowanie. Z NAM możesz skorzystać w dowolnym miejscu sieci, w którym możliwy jest dostęp do niego przez przeglądarkę.

Interfejs ekranu przechwytywania pakietów powinien być znajomy dla tych, którzy korzystali wcześniej z aplikacji typu Ethereal[2]. Każdy pakiet jest podzielony tak dalece, jak to możliwe, i nie ma dodatkowego okna pokazującego zawartość ASCII pakietu.

Jednym z ograniczeń przechwytywania pakietów jest brak inteligentnej sygnalizacji alarmowej, takiej, jaką można znaleźć w narzędziach przechwytywania pakietów z najwyższej półki. Wiele pozostałych funkcji jest dostępnych w module NAM, kiedy działa on jako sonda RMON.

NAM jest doskonałym narzędziem do diagnozowania problemów, a ponieważ jest zawsze pod ręką, może być nieoceniony podczas nagłych kryzysów (jest mało prawdopodobne, by ktoś pożyczył moduł z Twojego przełącznika 6509 działającego w sieci produkcyjnej, chociaż zdarzały się już dziwniejsze rzeczy). Możliwość przechwytywania więcej niż jednej sesji jednocześnie sprawia, że NAM jest idealnym dodatkiem do Twojego arsenału narzędzi administracyjnych. Dzięki zdolności przechwytywania ze źródeł RSPAN (patrz rozdział 18.) NAM może analizować ruch na dowolnym przełączniku w sieci.

Prosty zrzut ekranu z interfejsu NAM przedstawia rysunek 17.7.

Moduły IDSM

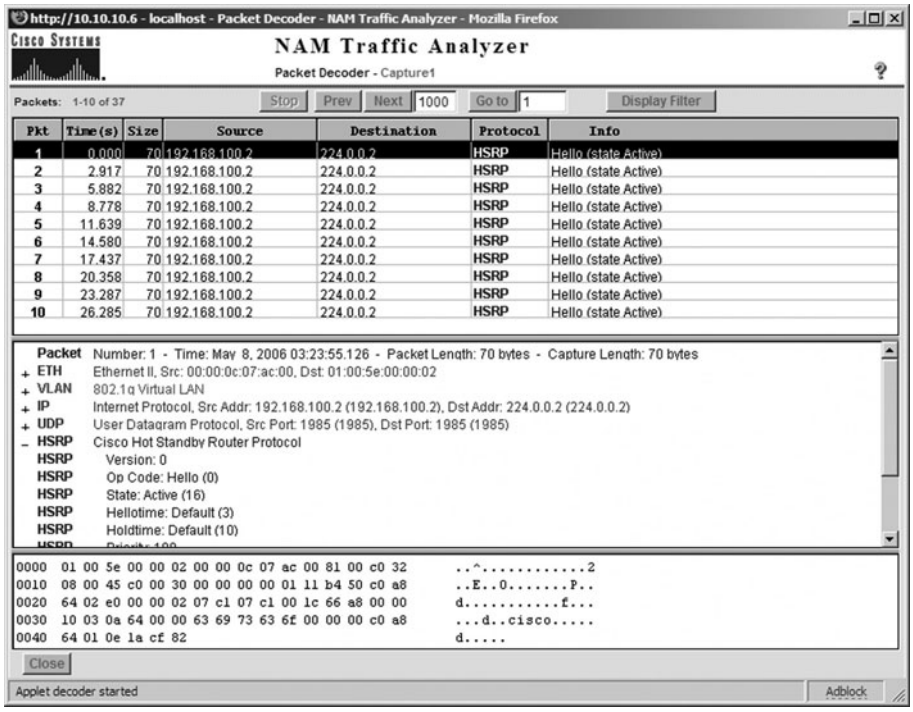
Możesz dodać funkcjonalność detekcji włamań do przełączników z serii 6500, montując moduł IDSM (ang. *intrusion detection system module*). Moduły te są właściwie prekonfigurowanymi serwerami Linuksowymi. Działają podobnie jak urządzenia IDS, ale posiadają dodatkowo możliwość próbkowania strumieni danych z prędkością uzyskiwaną na kablu, ponieważ podłączone są do szyny *crossbar fabric bus*.

Możesz zarządzać tymi modułami poprzez zintegrowany interfejs sieciowy pokazany na rysunku 17.8. Cisco zaleca jednak zarządzanie przez inną aplikację, taką jak VPN/Security Management Solution (VMS), Cisco Security Manager lub Cisco Security Monitoring Analysis and Response System (MARS).

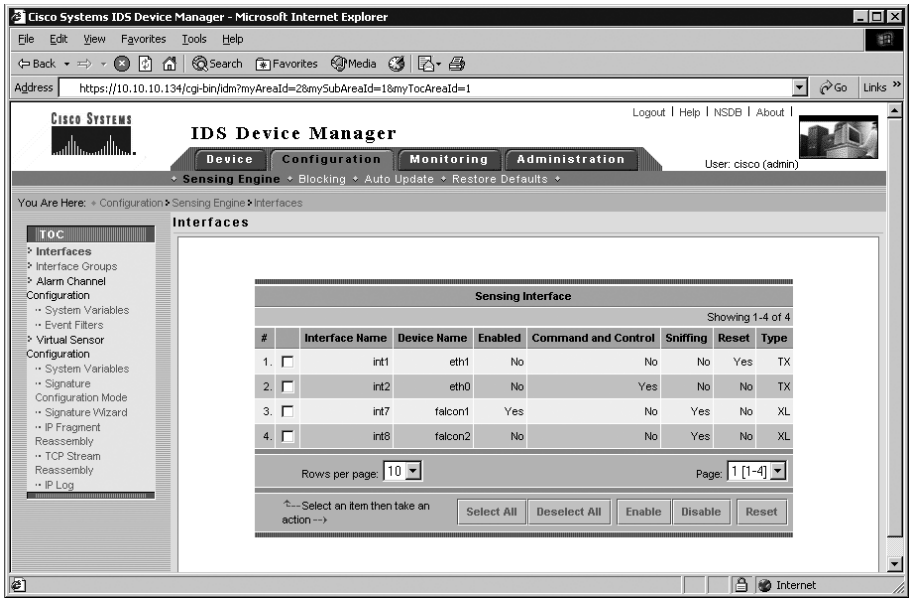
---

[2] Aplikacja Ethereal jest obecnie rozwijana pod nazwą Wireshark — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 17.7. Szczegóły przechwytywania pakietów w module NAM



Rysunek 17.8. Konfiguracja modułu IDS przez zintegrowany interfejs

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Podstawową konfigurację modułu możesz wykonać z poziomu samego przełącznika, łącząc się z modułem poleceniem `session slot` *`moduł_nr`* `processor` *`processor_nr`*. Zazwyczaj *`processor_nr`* to 1:

```
Switch-IOS-1#session slot 9 proc 1
The default escape character is Ctrl-^, then x.
You can also type 'exit' at the remote prompt to end the session
Trying 127.0.0.91 ... Open
login: cisco
Password:
***NOTICE***
This product contains cryptographic features and is subject to United States
and local country laws governing import, export, transfer and use. Delivery
of Cisco cryptographic products does not imply third-party authority to import,
export, distribute or use encryption. Importers, exporters, distributors and
users are responsible for compliance with U.S. and local country laws. By using
this product you agree to comply with applicable laws and regulations. If you
are unable to comply with U.S. and local laws, return this product immediately.
A summary of U.S. laws governing Cisco cryptographic products may be found at:
http://www.cisco.com/wwl/export/crypto
If you require further assistance please contact us by sending email to
export@cisco.com.
Switch-IOS-1-IDS#
```

Konfiguracja modułu IDSM różni się znacznie od konfiguracji pozostałych urządzeń i stanowi temat na osobną książkę.

Moduły FlexWAN

Moduły FlexWAN umożliwiają podłączenie linków WAN, takich jak T1, czy linków o wysokich prędkościach, czyli od DS3[3] do OC3[4].

Są dwa rodzaje modułów FlexWAN: FlexWAN i Enhanced FlexWAN. Podstawowe różnice pomiędzy tymi dwoma wersjami to prędkości CPU, wielkość pamięci oraz połączenie z szyną *crossbar fabric bus*.

Moduły Enhanced FlexWAN stosują te same adaptery portów WAN, które wykorzystywane są w routerach z serii Cisco 7600. Moduły te pracują w warstwie trzeciej i wymagają zastosowania modułu zarządzającego Supervisor-720 albo modułu Supervisor-2 w połączeniu z MSFC. W przypadku działania w trybie hybrydowym IOS interfejsy FlexWAN nie są widoczne w warstwie drugiej z poziomu CatOS.

Moduły CMM

Moduły mediów komunikacyjnych, czyli CMM (ang. *Communication Media Module*), zapewniają integrację telefonii w przełącznikach z serii 6500. Są to moduły *fabric-enabled* posiadające trzy sloty, które obsługują różne **adaptery portów** (ang. *port adapters*).

Adaptery portów dostępne dla CMM obejmują: moduły FXS (ang. *Foreign eXchange Service*) do podłączenia telefonów analogowych, modemów oraz faksów; moduły bramy T1/E1 CAS oraz

[3] T1 i DS3 to standardy stosowane w Ameryce Północnej; ich europejskimi odpowiednikami są E1 i E3 — *przyp. tłum.*

[4] Standard przepływności w sieciach światłowodowych stosowany w Ameryce Północnej; jego europejskim odpowiednikiem jest STM1 — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

PRI; adaptery konferencyjne i transkodujące, zapewniające usługi konferencyjne; a także moduły SRST (ang. *Unified Survivable Remote Site Telephony*), które zarządzają telefonami oraz połączeniami, kiedy system Unified Call Manager jest niedostępny.

Te adaptery portów mogą być mieszane i łączone w każdym zainstalowanym module CMM. Przełącznik z serii 6500 z modułem zarządzającym i wieloma modułami CMM zapewnia dużą gęstość portów dla łączności VoIP.

CatOS kontra IOS

Przełączniki Cisco Catalyst pierwotnie nie pracowały w systemie IOS — wczesne przełączniki *chassis-based* posiadały system operacyjny CatOS. Było to spowodowane tym, że technologia stosowana w tych przełącznikach pochodziła od innych firm, które zostały przez Cisco zakupione: Crescendo, Kalpana oraz Grand Junction.

CatOS może się wydawać niezgrabny dla tych, którzy korzystali dotąd tylko z IOS. Istnieją jednak pewne niezaprzeczalne zalety stosowania CatOS w środowisku przełączania. Jedna z tych zalet może być traktowana jak wada: jeśli Catalyst 6500 pracuje w systemie CatOS i posiada moduł MSFC, który dodaje funkcjonalność warstwy trzeciej, moduł ten będzie traktowany jako wyodrębnione urządzenie. Przełącznik wykorzystuje CatOS dla funkcjonalności warstwy drugiej, natomiast MSFC pracuje w IOS w celu zapewnienia funkcjonalności warstwy trzeciej i wyższych warstw. Podział ten może być łatwy do zrozumienia dla osób, które nie mają doświadczenia z przełącznikami IOS warstwy trzeciej. Jednak dla tych, którzy przyzwyczajeni są do przełączników IOS, takich jak Catalyst 3550 i 3750, konieczność przełączania pomiędzy systemami operacyjnymi może być uciążliwa i dezorientująca.

Ponieważ wszystkie nowe przełączniki Cisco warstwy trzeciej (jak 3550 i 3750) pracują tylko w systemie IOS, nauka logiki natywnego IOS może być dobrym pomysłem, tym bardziej że w sposób oczywisty Cisco dąży właśnie w tym kierunku. W pewnym momencie Cisco ogłosiło nawet plan wycofania CatOS, ale plany zostały wstrzymane z uwagi na ostry sprzeciw zagorzałych zwolenników tego systemu operacyjnego. Nadal spotykam CatOS w działających instalacjach, ale dzieje się to coraz rzadziej.

Kolejną przewagą CatOS nad IOS jest zwięzły sposób organizowania informacji. Doskonałym przykładem jest komenda `show port` w CatOS:

```
Switch-CatOS#sho port
Port  Name                 Status     Vlan       Duplex Speed Type
----- -------------------- ---------- ---------- ------ ----- ------------
 1/1  Trunk                connected  trunk        full  1000 1000BaseSX
 1/2  Trunk                connected  trunk        full  1000 1000BaseSX
 2/1  Trunk                connected  trunk        full  1000 1000BaseSX
 2/2  Trunk                connected  trunk        full  1000 1000BaseSX
 3/1  Web-1-E1             connected  20         a-full a-100 10/100BaseTX
 3/2  Web-2-E1             connected  20         a-full a-100 10/100BaseTX
 3/3  Web-3-E1             connected  20           full   100 10/100BaseTX
 3/4  Web-4-E1             connected  20           full   100 10/100BaseTX
 3/5  Web-5-E1             connected  20         a-full a-100 10/100BaseTX
 3/6  Web-6-E1             connected  20         a-full a-100 10/100BaseTX
 3/7  Web-7-E1             connected  20         a-full a-100 10/100BaseTX
 3/8  App-1-E1             connected  40         a-full a-100 10/100BaseTX
 3/9  App-2-E1             connected  40         a-full a-100 10/100BaseTX
 3/10 App-3-E1             connected  40         a-full a-100 10/100BaseTX
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
3/11 App-4-E1            connected  40         a-full a-100 10/100BaseTX
3/12                     notconnect              full   100 10/100BaseTX
3/13                     notconnect              full   100 10/100BaseTX
3/14 DB-1-E1             connected  50           full   100 10/100BaseTX
3/15 DB-2-E1             connected  50         a-full a-100 10/100BaseTX
3/16 DB-3-E1             connected  50         a-full a-100 10/100BaseTX
```

W tym zestawieniu znajdują się informacje o numerze portu, jego nazwie (jeśli jest zdefiniowana), statusie, przypisaniu do VLAN-u, prędkości i trybie dupleks, statusie autonegocjacji oraz typie portu. W pierwszej edycji tej książki zamieściłem następującą pochwałę komendy `show port` w CatOS:

> W IOS nie ma nic, co można by bezpośrednio porównać z tą komendą. Zamiast tego użytkownik musi zbierać częściowe informacje z wielu źródeł.

Ze wszystkich uwag, komentarzy i opinii, które otrzymałem, wynika, że ten fragment wzbudził największe poruszenie. Czytelnicy zwrócili mi uwagę, że komenda `show interface status` w IOS podaje informacje w sposób bardzo podobny do `show port` w CatOS. Kiedy otrzymałem pierwszy taki komentarz, odpaliłem swoje laboratorium i wpisałem tę komendę. Wyobraźcie sobie moje zdziwienie, gdy po 25 ciężkich latach w branży nauczyłem się tej cudownej komendy. Jest to nauczka, żeby nigdy nie pisać o rzeczach w kategoriach absolutnych, bo może się to odbić czkawką. Chciałbym podziękować wszystkim, którzy zwrócili mi na to uwagę:

```
3750-IOS#sho int status
Port      Name               Status       Vlan       Duplex  Speed Type
Gi1/0/1   Cozy-Server        notconnect   1            auto   auto 10/100/1000BaseTX
Gi1/0/2   Daisy-Server       notconnect   1            auto   auto 10/100/1000BaseTX
Gi1/0/3   Annie-Server       notconnect   1            auto   auto 10/100/1000BaseTX
Gi1/0/4   Guinness-Server    disabled     1            auto   auto 10/100/1000BaseTX
Gi1/0/5                      notconnect   1            auto   auto 10/100/1000BaseTX
Gi1/0/6                      notconnect   1            auto   auto 10/100/1000BaseTX
Gi1/0/7                      notconnect   1            auto   auto 10/100/1000BaseTX
Gi1/0/8                      notconnect   1            auto   auto 10/100/1000BaseTX
Gi1/0/9                      notconnect   1            auto   auto 10/100/1000BaseTX
Gi1/0/10                     notconnect   1            auto   auto 10/100/1000BaseTX
Gi1/0/11                     notconnect   1            auto   auto 10/100/1000BaseTX
 --More--
```

Mimo wszystko przedstawię jeszcze jedną z najczęściej stosowanych przeze mnie komend do zbierania informacji, ponieważ podaje ona adresy IP:

```
Switch-IOS#sho ip int brief
Interface                  IP-Address      OK? Method Status                Protocol
Vlan1                      unassigned      YES NVRAM  administratively down down
Vlan20                     10.10.20.2      YES manual up                    up
Vlan40                     10.10.40.2      YES manual up                    up
Vlan50                     10.10.50.2      YES manual up                    up
GigabitEthernet1/1         unassigned      YES unset  up                    up
GigabitEthernet1/2         unassigned      YES unset  up                    up
GigabitEthernet1/3         unassigned      YES unset  up                    up
GigabitEthernet1/4         unassigned      YES unset  up                    up
```

IOS ma tendencje do rozwlekłości. Za przykład może posłużyć listing z polecenia IOS `show interface` *`interfejs_nr`*, który przedstawia informacje dotyczące interfejsów:

```
Switch-IOS#sho int g3/1
GigabitEthernet3/1 is up, line protocol is up (connected)
  Hardware is C6k 1000Mb 802.3, address is 0015.6356.62bc (bia 0015.6356.62bc)
  Description: Web-1-E1
  MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 1/255
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  Encapsulation ARPA, loopback not set
  Full-duplex, 1000Mb/s
  input flow-control is off, output flow-control is on
  Clock mode is auto
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input never, output 00:00:47, output hang never
  Last clearing of "show interface" counters never
  Input queue: 0/2000/2/0 (size/max/drops/flushes); Total output drops: 2
  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  5 minute input rate 456000 bits/sec, 91 packets/sec
  5 minute output rate 110000 bits/sec, 81 packets/sec
     714351663 packets input, 405552413403 bytes, 0 no buffer
     Received 15294 broadcasts, 0 runts, 0 giants, 0 throttles
     2 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
     0 input packets with dribble condition detected
     656796418 packets output, 97781644875 bytes, 0 underruns
     2 output errors, 0 collisions, 2 interface resets
     0 babbles, 0 late collision, 0 deferred
     0 lost carrier, 0 no carrier
     0 output buffer failures, 0 output buffers swapped out
```

Listing z polecenia CatOS `show port` *`port_nr`* jest dużo bardziej przejrzysty, szczególnie kiedy w pośpiechu szukasz konkretnej informacji. Minusem jest, że polecenie CatOS wyświetla mniej danych niż jego odpowiednik w IOS:

```
Switch-CatOS: (enable)sho port 3/1
* = Configured MAC Address
Port  Name                 Status     Vlan       Duplex Speed Type
----- -------------------- ---------- ---------- ------ ----- ------------
 3/1  Web-1-E1             connected  20           auto  auto 10/100BaseTX

Port  AuxiliaryVlan AuxVlan-Status     InlinePowered     PowerAllocated
                                     Admin Oper   Detected mWatt mA @42V
----- ------------- -------------- ----- ------ -------- ----- --------
 3/1  none          none           -     -      -        -     -

Port  Security Violation Shutdown-Time Age-Time Max-Addr Trap     IfIndex
----- -------- --------- ------------- -------- -------- -------- -------
 3/1  disabled  shutdown             0        0        1 disabled       5

Port  Num-Addr Secure-Src-Addr     Age-Left Last-Src-Addr     Shutdown/Time-Left
----- -------- -----------------   -------- ----------------- ------------------
 3/1         0                 -          -                 -        -         -

Port  Flooding on Address Limit
----- -------------------------
 3/1                    Enabled
--More--
```

Wiele osób preferuje sposób wyświetlania informacji dla komend stosowany w przełącznikach CatOS, ale jako konsultant nie powinienem mieć specjalnych preferencji, gdyż muszę pracować na takim sprzęcie, jaki posiada klient.

Jedną z istotnych funkcji CatOS, która jeszcze do niedawna nie była dostępna dla IOS, jest `show top`. Wykonanie komendy `show top 5 util all back interval 60` spowoduje, że przełącznik rozpocznie w tle przygotowywanie raportu *Top-N* dla pięciu najbardziej wykorzystywanych portów i zachowa ten raport do przeglądu:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Switch-CatOS: (enable)show top 5 util all back interval 60
Switch-CatOS: (enable) 2006 May 07 12:47:00 EST +00:00 %MGMT-5-TOPN_START:Report 3
started by telnet/20.20.20.100/GAD
```

Zauważ, że argumentem `back` zdefiniowałem opcję pracy w tle (ang. *background*), mogę więc wykonywać inne czynności, w czasie kiedy przygotowywany jest raport:

```
Switch-CatOS: (enable)
Switch-CatOS: (enable)
Sec-6505-1-TOP: (enable)dir
-#- -length- -----date/time------ name
  2 10518855 May 2  2006 02:27:09 cat6000-supk8.7-7-9.bin
 15    82230 May 2  2006 08:21:55 switch.cfg

4604208 bytes available (11386576 bytes used)
Switch-CatOS: (enable)
Switch-CatOS: (enable)
Switch-CatOS: (enable) 2006 May 07 12:48:01 EST +00:00 %MGMT-5-TOPN_AVAILABLE:Report 3
available: (enable) 2006 May 07 12:48:01 EST +00:00 %MGMT-5-TOPN_AVAILABLE:Report 3
available //Raport TOPN jest dostępny jako raport 3
```

Kiedy przeglądałem katalog flash, mój raport został ukończony. Wyświetlona została informacja, że jest on dostępny jako `report #3`. Mogę go teraz przejrzeć, wpisując następującą komendę:

```
Switch-CatOS: (enable)sho top report 3
Start Time:     May 07 2006 12:47:00
End Time:       May 07 2006 12:48:01
PortType:       all
Metric:         util
Port  Band- Uti Bytes                Pkts       Bcst       Mcst       Error Over
      width  %  (Tx + Rx)            (Tx + Rx)  (Tx + Rx)  (Tx + Rx)  (Rx)  flow
----- ----- --- -------------------- ---------- ---------- ---------- ----- ----
 3/14   100   0               624014       1126         89         89     0    0
 3/15   100   0               105347        590          6         32     0    0
 3/16   100   0               889310       2319         89         99     0    0
 3/8    100   0               536246       3422         97         41     0    0
 3/9    100   0               315228       2094          0        405     0    0
```

Funkcja `show top` umożliwia również przygotowanie raportu *Top-N* dotyczącego występowania błędów. Jest to bardzo przydatne narzędzie, kiedy podejrzewasz problemy z autonegocjacją. Aby uruchomić w CatOS raport *Top-N* dotyczący błędów, zastosuj polecenie `show top error all back interval 60`.

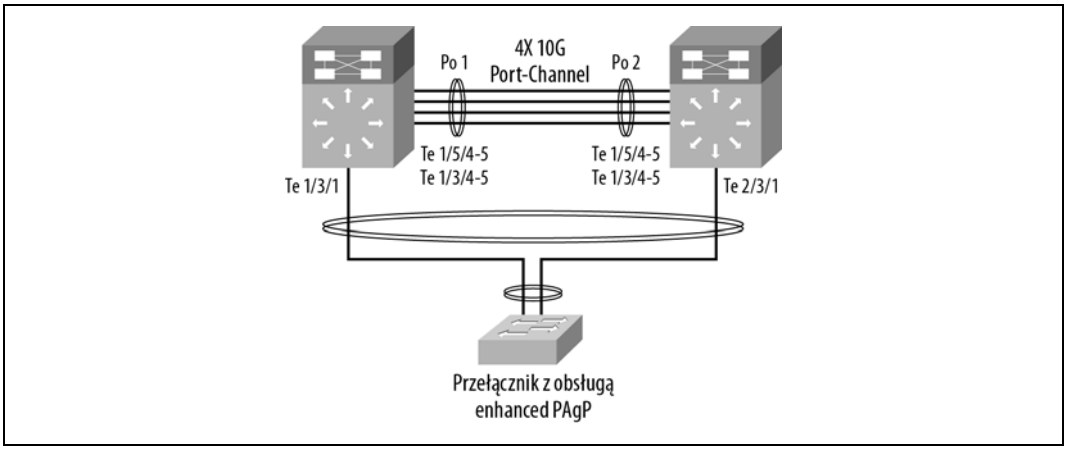
Wersje IOS 12.2(18)SXE oraz późniejsze dla przełączników Catalyst 6500 również umożliwiają generowanie raportów *Top-N*. Rezultaty są podobne jak w przypadku CatOS, inna jest jednak komenda. Do uruchomienia raportu *Top-N* w IOS stosuje się polecenie `collect top`.

Instalacja VSS

W tym podrozdziale omówię komendy służące do uruchamiania VSS dla pary samodzielnych przełączników 6509-V-E. Zbudujemy klaster, którego schemat przedstawiony jest na rysunku 17.9.

Na początku przełączniki są samodzielnymi urządzeniami, musimy więc zacząć od indywidualnej konfiguracji każdego z nich. Kiedy już to zrobimy, przedstawię odpowiednią konfigurację

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 17.9. Schemat VSS z wykorzystaniem przełącznika z obsługą PAgP

dla przełącznika znajdującego się po prawej stronie na rysunku. Na początek trzeba skonfigurować przełącznik w trybie *failover*, utrzymującym informacje o stanie aplikacji (ang. *stateful failover mode*). Robi się to następująco:

```
6509-VE-Left#conf t
Enter configuration commands, one per line.  End with CNTL/Z.
6509-VE-Left(config)#redundancy
6509-VE-Left(config-red)#mode ?
  rpr  Route Processor Redundancy
  sso  Stateful Switchover
6509-VE-Left(config-red)#mode sso
```

Stan redundancji możesz sprawdzić poleceniem `show redundancy states`:

```
6509-VE-Left#sho redundancy states
       my state = 13 -ACTIVE
     peer state = 1  -DISABLED
          Mode = Simplex
          Unit = Primary
        Unit ID = 5

Redundancy Mode (Operational) = sso
Redundancy Mode (Configured)  = sso
Redundancy State              = Non Redundant
     Maintenance Mode = Disabled
 Communications = Down      Reason: Simplex mode

   client count = 135
 client_notification_TMR = 30000 milliseconds
          keep_alive TMR = 9000 milliseconds
        keep_alive count = 0
    keep_alive threshold = 18
          RF debug mask = 0x0
```

Jeśli router pracuje w protokole OSPF, musisz dodać następujące komendy do konfiguracji OSPF (nie dotyczy nas to w tej sytuacji, więc nie muszę tego robić):

```
router ospf proces-id
   nsf
   exit
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Teraz musimy skonfigurować wirtualną domenę przełącznika (ang. *switch virtual domain*). Domeną jest liczba z zakresu 1 – 255:

```
6509-VE-Left(config)#switch virtual domain ?
  <1-255>  Virtual switch domain number
```

Domena musi być taka sama na obu przełącznikach. Teraz skonfiguruję domenę 100:

```
6509-VE-Left(config)#switch virtual domain 100
Domain ID 100 config will take effect only
after the exec command 'switch convert mode virtual' is issued
```

Przełącznikowi po lewej nadam numer 1, a przełącznikowi po prawej numer 2. Oczywiście możesz definiować swoje przełączniki wedle uznania:

```
6509-VE-Left(config-vs-domain)#switch 1
```

Możesz skonfigurować jeden z przełączników z wyższym priorytetem i wykorzystując wywłaszczanie, ustawić go tak, żeby był zawsze preferowany. Nie jestem jednak zwolennikiem takiego rozwiązania, ponieważ tryb *failover* zadziała tylko po przeładowaniu aktywnego przełącznika. Cisco nawet zniechęca do tego typu rozwiązań w swoim dokumencie dotyczącym zaleceń praktycznych (link do dokumentu znajdziesz na końcu rozdziału). System VSS nie funkcjonuje tak jak firewalle Cisco, gdzie aktywna jednostka może być zmieniona bez konieczności restartowania systemu. Jeśli chcesz skonfigurować wywłaszczanie, przypisz do przełącznika priorytet poleceniem `switch` *`switch_nr`* `priority` *`priorytet_wartość`*:

```
6509-VE-Left(config-vs-domain)#switch 1 priority ?
  <1-255>  Virtual switch priority value

6509-VE-Left(config-vs-domain)#switch 1 priority 105
```

Aby włączyć wywłaszczanie, zastosuj polecenie `switch` *`switch_nr`* `preempt`. Nawet przełącznik poinstruuje Cię, że nie jest to dobry pomysł:

```
6509-VE-Left(config-vs-domain)#switch 1 preempt ?
  <15-20>  Delay in minutes before standby takes over as active. Default is 15
           minutes //Opóźnienie w minutach zanim przełącznik w stanie gotowości zostanie przełącznikiem aktywnym.
                   // Domyślne opóźnienie to 15 minut
  <cr>
```

Domyślne opóźnienie to 15 minut:

```
6509-VE-Left(config-vs-domain)#switch 1 preempt

Please note that Preempt configuration will make the ACTIVE switch with lower
priority to reload forcefully when preempt timer expires. //Pamiętaj, że konfiguracja wywłaszczania
// spowoduje przymusowe przeładowanie AKTYWNEGO przełącznika z niższym priorytetem po wygaśnięciu czasu
// wywłaszczenia.

The default preempt timer is 15 minutes. It is set to the approximate time required
to bring up all linecards in a fully loaded chassis. Once the preempt timer starts,
the standby switch with higher priority will take over as active after 15 minutes. //Domyślny czas
// wywłaszczenia to 15 minut. Jest to przybliżony czas ponownego uruchomienia wszystkich kart rozszerzeń dla w pełni
// załadowanej obudowy. Przełącznik w stanie gotowości z wyższym priorytetem zostanie przełącznikiem aktywnym po 15
// minutach od rozpoczęcia odliczania czasu wywłaszczenia.
```

Jeśli się rozmyślisz, po prostu zaneguj tę komendę. W tym przypadku nie będzie już żadnych fajerwerków:

```
6509-VE-Left(config-vs-domain)#no switch 1 preempt
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Stosując VSS, możesz ręcznie przypisać adresy MAC, które będą wykorzystywane przez wirtualny przełącznik. Jednak łatwiejszym sposobem w większości instalacji jest ustawienie, aby przełącznik sam przypisał adresy MAC z predefiniowanego zakresu. Służy do tego polecenie `mac-address use-virtual`:

```
6509-VE-Left(config-vs-domain)#mac-address ?
  H.H.H        Manually set router mac address
  use-virtual  Use mac-address range reserved for Virtual Switch System

6509-VE-Left(config-vs-domain)#mac-address use-virtual
```

Kanał łączący oba przełączniki i zachowujący się jak kabel do łączenia w stos zwany jest **wirtualnym linkiem przełącznika**, czyli VSL (ang. *virtual switch link*). Kanał ten powinien posiadać tak wiele łączy wysokiej przepustowości, ile możesz poświęcić. Powinny one być rozdystrybuowane na różne moduły. Należy je również instalować po dwa, lub w wielokrotności dwóch, dla zapewnienia optymalnego dostępu. Kanał ten musi być niezawodny, należy więc uczynić go na tyle elastycznym, na ile to możliwe. W naszym przykładzie mamy 4 łącza 10 Gb Ethernet tworzące VSL.

VSL EtherChannel musi być **inny** dla każdego przełącznika, aby w przypadku gdy są one połączone, mogły rozróżnić, do którego z fizycznych przełączników należy dana strona linku. Jest to jedyna sytuacja, w której wymagana jest odmienna konfiguracja dla każdego port channel z VSS. Uważaj na to, ponieważ każdemu, kogo uczyłem, wydawało się to sprzeczne z intuicją.

Skonfigurowałem Po10 po lewej stronie i Po20 po prawej. Te kanały port channel muszą być przypisane do VSL za pomocą polecenia `switch virtual link` *`core-switch-nr`*. *`Core-switch-nr`* to numer, który przypisaliśmy przełącznikowi w sekcji dotyczącej wirtualnej domeny przełącznika:

```
6509-VE-Left(config)#int port-channel 10
6509-VE-Left(config-if)#switch virtual link 1
```

Dodajmy kilka fizycznych portów do port channel:

```
6509-VE-Left(config)#int range ten 3/4-5 , ten 5/4-5
6509-VE-Left(config-if-range)#channel-group 10 mode on
6509-VE-Left(config-if-range)#no shut
```

Możemy teraz aktywować VSS. Spowoduje to zrestartowanie przełącznika, który uruchomi się w trybie VSS. Od tej pory będziesz musiał określać porty, podając, gdzie przynależą — chassis/moduł/port:

```
6509-VE-Left#switch convert mode virtual
*Sep 22 19:07:10.968: %SYS-5-CONFIG_I: Configured from console by console
This command will convert all interface names
to naming convention "interface-type switch-number/slot/port",
save the running config to startup-config and
reload the switch. //Ta komenda zmieni nazwy wszystkich interfejsów, stosując konwencję „typ-interfejsu
// nr-przełącznika/slot/port". Bieżąca konfiguracja zostanie zapisana, a przełącznik zostanie uruchomiony ponownie.
Do you want to proceed? [yes/no]: yes //Czy chcesz kontynuować? [tak/nie]
Converting interface names
Building configuration...
[OK]
Saving converted configuration to bootflash: ...
Destination filename [startup-config.converted_vs-20090922-190952]
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
[usunięto nudny tekst bootowania przełącznika]

Initializing as Virtual Switch ACTIVE processor

[usunięto więcej nudnego tekstu bootowania]

*Sep 22 19:11:46.399: %VS_PARSE-4-NO_VSL_CONFIG: No VSL interface is configured
*Sep 22 19:11:45.836: %PFREDUN-6-ACTIVE: Initializing as ACTIVE processor for
this switch
*Sep 22 19:11:46.584: %VSLP-5-RRP_NO_PEER: No VSLP peer found. Resolving role
as Active
*Sep 22 19:11:47.023: %VSLP-2-VSL_DOWN:   VSL links down and not ready for any
traffic
*Sep 22 19:11:49.716: %OIR-SW1_SP-6-CONSOLE: Changing console ownership to
route processor
```

Po kolejnym nudnym bootowaniu przełącznik jest już online. W tym momencie można zauważyć, że wszystkie interfejsy zostały przenumerowane tak, aby dodać numer przełącznika, podobnie jak w przypadku stosu przełączników 3750:

```
6509-VE-Left#sho ip int brie
Interface              IP-Address      OK? Method Status                Protocol
Vlan1                  unassigned      YES NVRAM  administratively down down
GigabitEthernet1/1/1   unassigned      YES NVRAM  administratively down down
GigabitEthernet1/1/2   unassigned      YES NVRAM  administratively down down
GigabitEthernet1/1/3   unassigned      YES NVRAM  administratively down down
[fragment usunięty]
```

Listing z polecenia `show redundancy` wskazuje, że przełącznik pracuje w trybie *simplex*. Oznacza to, że mimo uruchomienia VSS przełącznik nie jest na razie połączony z drugim przełącznikiem. Ten problem zostanie rozwiązany, kiedy skonfigurujemy i zrestartujemy przełącznik po prawej stronie:

```
6509-VE-Left#sho redundancy
Redundant System Information :
------------------------------
       Available system uptime = 3 minutes
Switchovers system experienced = 0
              Standby failures = 0
        Last switchover reason = none
                 Hardware Mode = Simplex
    Configured Redundancy Mode = sso
     Operating Redundancy Mode = sso
              Maintenance Mode = Disabled
                Communications = Down      Reason: Simplex mode
[fragment usunięty]
```

Status klastra VSS możesz sprawdzić komendą `show switch virtual role`:

```
6509-VE-Left#sho switch virtual  role
Switch  Switch Status  Preempt    Priority  Role     Session ID
        Number         Oper(Conf) Oper(Conf)         Local  Remote
-----------------------------------------------------------------
LOCAL    1     UP      FALSE(N )   100(100)  ACTIVE   0      0

In dual-active recovery mode: No
```

Poniżej konfiguracja przełącznika z prawej strony schematu, stworzona na podstawie konfiguracji przełącznika znajdującego się po lewej stronie:

```
redundancy
 mode sso
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
switch virtual domain 100
 switch 2
 switch 2 priority 100
 mac-address use-virtual

int port-channel 20
 switch virtual link 2

int range ten 3/4-5 , ten 5/4-5
 channel-group 20 mode on
 no shut
```

Po dodaniu tej konfiguracji przełączniki będą mogły połączyć się w jeden logiczny przełącznik. Kiedy przełączniki przejdą w tryb VSS, będą musiały ponownie wykonać restart. Jeśli switche zostały skonfigurowane dla VSS, 10 Gb/s link VSL zostanie rozpoznany i uruchomiony, zanim uruchomione zostaną pozostałe moduły. Jest to istotne tylko w przypadku inicjującej konfiguracji i nie powinno być problemem podczas normalnej pracy.

Podczas bootowania klastra VSS możesz zobaczyć, że te łącza „podnoszą" się pierwsze, zanim zostaną uruchomione pozostałe elementy przełącznika. Oto fragment listingu z bootowania pary VSS:

```
*Sep 23 15:23:47.856: %PFREDUN-6-ACTIVE: Initializing as ACTIVE
processor for this switch
*Sep 23 15:23:53.912: %VSL_BRINGUP-6-MODULE_UP: VSL module in
slot 5 switch 1 brought up
*Sep 23 15:24:28.191: %VSLP-5-RRP_ROLE_RESOLVED: Role resolved
as STANDBY by VSLP
*Sep 23 15:24:28.195: %VSL-5-VSL_CNTRL_LINK:  New VSL Control
Link  5/5
*Sep 23 15:24:28.712: %VSLP-5-VSL_UP:  Ready for control traffic
*Sep 23 15:25:01.363: %OIR-SW1_SPSTBY-6-CONSOLE: Changing console
ownership to route processor
```

Aby sprawdzić, czy Twój link wirtualny jest uruchomiony, zastosuj komendę `show switch virtual link`. Ten listing pokazuje aktywny link wirtualny:

```
6509-VE-Right#sho switch virtual link
VSL Status : UP
VSL Uptime : 4 minutes
VSL SCP Ping : Pass
VSL ICC Ping : Pass
VSL Control Link : Te2/5/5
```

Możesz także użyć komendy `show switch virtual link detail`, aby uzyskać bardziej szczegółowe informacje. Listing dla tego polecenia jest dosyć obszerny, nie będę go więc tutaj zamieszczał.

Jeśli chcesz sprawdzić status klastra przełącznika wirtualnego, wprowadź polecenie `show switch virtual role`. W tym przykładzie zastosowałem tę komendę dla przełącznika po prawej:

```
6509-VE-Right#sho switch virtual role
Switch  Switch Status  Preempt    Priority  Role     Session ID
        Number         Oper(Conf) Oper(Conf)         Local  Remote
-----------------------------------------------------------------
LOCAL    2     UP      FALSE(N )   100(100)  ACTIVE   0      0
REMOTE   1     UP      FALSE(N )   100(100)  STANDBY  5555   5649
```

Kiedy system VSS jest online, a przełączniki staną się jednym przełącznikiem logicznym, będziesz mógł konfigurować system tylko z poziomu aktywnego modułu zarządzającego:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
6509-VE-Right-sdby>
Standby console disabled //Konsola przełącznika zapasowego jest wyłączona
```

W tym momencie najsensowniej jest nadać nazwę dla całego systemu logicznego, zamiast używać dezorientującej nazwy, która właściwie jest bezużyteczna (oba przełączniki nazywają się teraz 6509-VE-Right). Przeniosłem więc kabel konsoli z powrotem do przełącznika po lewej i jak widzisz, nazwa hosta jest taka sama jak dla przełącznika po prawej. Kiedy jesteś w trybie VSS, konfiguracja stosowana jest dla klastra logicznego, a nie dla fizycznego przełącznika:

```
6509-VE-Right(config)#hostname VSS
VSS(config)#exit
VSS#
```

Pozostałe polecane komendy VSS

Według dokumentacji Cisco komenda `mac-address-table synchronize` powinna być zastosowana dla klastra VSS. Więcej informacji znajdziesz w dokumencie Cisco zamieszczonym na stronie *http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/12.2SX/configuration/guide/vss.html#wp1131059*:

```
VSS(config)#mac-address-table synchronize
 % Current OOB activity time is [160] seconds
 % Recommended aging time for all vlans is at least three times the
  activity interval and global aging time will be changed automatically
  if required
```

VSS detekcja dual-active

Przełączniki dostępowe z odpowiednią wersją IOS mogą obsługiwać wiadomości enhanced PAgP, które umożliwiają wykrywanie problemu *dual-active*. Termin **dual-active**, określany również przez wiele osób jako **split-brain**, opisuje sytuację, w której dwa fizyczne przełączniki stają się przełącznikami aktywnymi. Stanowi to problem, ponieważ oba te przełączniki posiadają wtedy ten sam adres IP. To bardzo zły stan rzeczy i należy zrobić wszystko, by się przed tym bronić. Jest na to kilka sposobów. Jeśli posiadasz wyłącznie przełączniki Cisco, to Cisco zaleca zastosowanie enhanced PAgP. Jeśli masz przełączniki innych producentów, możesz zastosować odmienną metodę:

```
VSS#sho switch virtual dual-active ?
  bfd         Dual-active bfd summary
  fast-hello  Dual-active fast-hello summary
  pagp        Dual-active pagp summary
  summary     Dual-active config summary
```

Cisco oferuje enhanced PAgP do wykrywania problemu *dual-active*, kiedy trzeci przełącznik jest podłączony do klastra VSS. Aby można było wykorzystać tę funkcję, na trzecim przełączniku musi być zainstalowana odpowiednia wersja IOS. Minimalne wymagania dla IOS przedstawia tabela 17.1.

Dla pary przełączników VSS wykrywanie *dual-active* PAgP jest domyślnie włączone. Wiadomości wysyłane są jednak tylko na kanały port channel z włączonym trybem zaufanym (ang. *trust mode*). Aby uruchomić tryb zaufany, **musisz** zamknąć zaufany port channel, jeśli taki

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Tabela 17.1. Minimalne wymagania dla IOS

Seria urządzeń	Najniższa wymagana wersja IOS
Catalyst 3750	IOS 12.2(46)SE
Catalyst 4500	IOS 12.2(44)SE
Catalyst 6500	IOS 12.2(33)SXH
Catalyst 6500VSS	IOS 12.2(33)SXH1

istnieje. Kiedy interfejs port channel jest już wyłączony, trzeba skonfigurować wirtualną domenę przełącznika. Poniżej skonfigurowałem po30, kanał MEC łączący trzeci przełącznik z parą VSS tak, jak przedstawia rysunek 17.9:

```
VSS(config)#int po30
VSS(config-if)# shut
VSS(config)#switch virtual domain 100
VSS(config-vs-domain)#dual-active detection pagp
VSS(config-vs-domain)#dual-active detection pagp trust channel-group 30
VSS(config)#int po30
VSS(config-if)#no shut
```

Żeby sprawdzić status wykrywania *dual-active*, użyj polecenia `show switch virtual dual-active summary`:

```
VSS#sho switch virtual dual-active summary
Pagp dual-active detection enabled: Yes
Bfd dual-active detection enabled: Yes
Fast-hello dual-active detection enabled: Yes

No interfaces excluded from shutdown in recovery mode

In dual-active recovery mode: No
```

Możesz również sprawdzić status wykrywania *dual-active* dla linków poleceniem `show switch virtual dual-active pagp`:

```
VSS#sho switch virtual dual-active pagp
PAgP dual-active detection enabled: Yes
PAgP dual-active version: 1.1

Channel group 100 dual-active detect capability w/nbrs
Dual-Active trusted group: Yes
          Dual-Active     Partner              Partner   Partner
Port      Detect Capable  Name                 Port      Version
Gi1/1/1   No              None                 None      N/A
Gi2/1/1   No              None                 None      N/A
```

Komendy failover dla VSS

Jedynym sposobem na przeniesienie w klastrze VSS statusu aktywności na przełącznik zapasowy jest zrestartowanie przełącznika aktywnego. W systemie, który testowałem, zrestartowanie aktywnego przełącznika, bez ładowania aktualnej konfiguracji poza łączem VSL VSS zajmuje prawie 15 minut. Jednak zapasowy moduł zarządzający przejmuje funkcje modułu aktywnego w czasie poniżej jednej sekundy.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Pamiętaj, że tylko aktywny przełącznik będzie akceptował polecenia. Nagłówek wiersza poleceń w konsoli zapasowego przełącznika będzie miał końcówkę „-sdby" dodaną do nazwy hosta. Przy próbie konfiguracji przełącznik zapasowy wyświetli następującą informację:

```
VSS-sdby>en
Standby console disabled //Konsola przełącznika zapasowego jest wyłączona
```

Zawsze bądź ostrożny przy restartowaniu przełącznika VSS, ponieważ nawet jeśli chcesz zrestartować tylko jeden z przełączników, możesz przy tym zrestartować cały klaster. O ile wiem, nie jest to nic dobrego.

Jeśli komenda `reload` zostanie zastosowana na aktywnym przełączniku, cały klaster (oba przełączniki) zostanie zrestartowany. Wyświetlą się oczywiście pomocne informacje z ostrzeżeniem, jednak jak często każdy z nas czyta takie informacje?

Aby zrestartować przełącznik zapasowy z konsoli przełącznika aktywnego, zastosuj polecenie `redundancy reload peer`:

```
VSS# redundancy reload peer
Reload peer [confirm]
```

Żeby zrestartować tylko aktywny przełącznik i wymusić tym samym uruchomienie trybu awaryjnego (*failover*), skorzystaj z komendy `redundancy reload shelf`:

```
VSS# redundancy reload shelf
Reload this shelf[confirm]
```

Możesz też zrestartować wybrany przełącznik, dodając numer przełącznika (`shelf`) w tej samej komendzie:

```
VSS# redundancy reload shelf 2
Reload the entire remote shelf[confirm]
```

Jeśli chcesz wymusić uruchomienie trybu *failover*, zastosuj polecenie `redundancy force-swit ↪chover`. Aktywny przełącznik zostanie zrestartowany, co spowoduje przejęcie jego funkcji przez przełącznik zapasowy. Nie znam innej metody na wymuszenie zmiany przełączników niż restart przełącznika aktywnego, więc nie robię tego lekką ręką — pamiętaj, że taki restart może zająć nawet 15 minut:

```
VSS# redundancy force-switchover
This will reload the active unit and force switchover to standby[confirm]
```

Polecenie `show redundancy switchover` służy do sprawdzenia, kiedy miało miejsce ostatnie uruchomienie trybu *failover*:

```
VSS#sho redundancy switchover
Switchovers this system has experienced          : 2
                              //Liczba przypadków zamiany przełączników dla tego systemu
Last switchover reason                           : active unit removed
                              //Powód ostatniej zamiany przełączników: jednostka aktywna została usunięta
Uptime since this supervisor switched to active  : 2 hours, 14 minutes
                              //Czas pracy od momentu, kiedy obecny moduł zarządzający stał się modułem aktywnym
Total system uptime from reload                  : 3 hours, 32 minutes
                              //Całkowity czas pracy systemu od ostatniego restartu
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Inne komendy VSS

Komendy odnoszące się do sprzętu są co jakiś czas modyfikowane, aby umożliwić użytkownikowi przegląd wyposażenia każdego z przełączników. Przykładowo komenda `show module` posiada obecnie dodatkowy argument `switch`:

```
VSS#sho module switch ?
  <1-2>    Virtual switch number //Numer przełącznika wirtualnego
  all      show all linecard module information //Pokaż informacje o wszystkich kartach rozszerzeń
  version  show all linecard version information //Pokaż informacje o wersji dla wszystkich kart rozszerzeń
  |        Output modifiers //Modyfikatory
  <cr>
```

Zwróć na to uwagę, ponieważ listing dla komendy `show module` zawiera tylko informację o przełączniku aktywnym i nie wskazuje, którego przełącznika dotyczą wyświetlone informacje. Aby wyświetlić moduły obu przełączników, skorzystaj z polecenia `show module switch all`.

Analogicznie polecenie `show environment` zostało zmodyfikowane argumentem `switch`, jednak domyślnie pokazuje ono informacje o obydwu przełącznikach, również bez dodawania nowego argumentu:

```
VSS#sho env switch ?
  <1-2>  Virtual switch number
```

VSS w praktyce

Cisco opublikowało dokument z zaleceniami praktycznymi dotyczącymi VSS. Powinieneś się z nim zapoznać, jeśli planujesz zainstalowanie VSS lub będziesz administrował już istniejącym środowiskiem VSS. Dokument ten znajdziesz pod adresem *http://www.cisco.com/en/US/products/ps9336/products_tech_note09186a0080a7c837.shtml*.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 18.

Cisco Nexus

W chwili kiedy piszę ten rozdział, linia produktów Cisco Nexus składa się z czterech poziomów urządzeń: 7000, 5000, 2000 oraz 1000. Seria 1000 to z reguły urządzenia typu *server appliance* (1010) lub wirtualne przełączniki (1000v). W tym rozdziale skupię się głównie na urządzeniach 7000 – 2000. Podobnie jak w przypadku innych produktów Cisco urządzenia Nexus są tym większe i tym bardziej funkcjonalne, im większa liczba stosowana jest w nazwie modelu.

Jedną z przyczyn popularności platformy Nexus jest coś, co nazywa się **unified fabric**. *Unified fabric* umożliwia połączenie na jednym złączu standardowej sieci z siecią SAN (ang. *Storage Area Network*). Może to być całkiem istotne, ponieważ pozwala zastąpić w centrum danych parę przełączników Catalyst 6500 oraz parę przełączników typu *fabric* MDS 9000 jedną parą przełączników Nexus 7000. W tym rozdziale skupię się na sieciowych funkcjach przełączników Nexus, pomijając raczej kwestię magazynowania danych. Jestem pewien, że są przedsiębiorstwa korzystające z tego rozwiązania, jednak od kilku lat, podczas których pracuję z urządzeniami Nexus, nie widziałem tego w akcji. Jak dotąd zbudowałem trzy rozległe centra danych i pracowaliśmy z Cisco nad zrozumieniem protokołu FCoE (ang. *Fibre Channel over Ethernet*), ale wciąż nie udało nam się namówić kierownictwa na jakieś konkretne zakupy.

W przełącznikach Nexus zastosowano system operacyjny zwany NX-OS. NX-OS wywodzi się od przełączników SAN z serii MDS, ale wygląda i zachowuje się podobnie jak IOS. Napisałem „podobnie", ponieważ istnieją subtelne różnice, które mogą Cię zmylić, kiedy będziesz uczył się tego systemu. Ta książka zawiera mnóstwo takich przykładów, więc staraj się odnaleźć wszystkie różnice, dobrze się przy tym bawiąc.

Nexus — dostępny sprzęt

Urządzenia sieciowe Nexus dostępne są w czterech seriach: 7000, 5000, 2000 oraz 1000. Seria 7000 to duże przełączniki typu *chassis* o podobnym przeznaczeniu co Catalyst 6500 (oraz przełączniki SAN MDS-9000). Seria 5000 to rozwiązanie do montażu na górnym poziomie szafy rackowej, zapewniające dużą liczbę interfejsów 10 Gb. Urządzenia FEX 2000 zostały zaprojektowane jako rozszerzenia dla przełączników 7000 oraz 5000.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Nexus 7000

Przełączniki Nexus 7000 są obecnie dostępne w dwóch modelach: 7010 oraz 7018. Podobnie jak w przypadku innych przełączników Cisco *chassis* numer modelu odzwierciedla liczbę slotów w obudowie. Przełączniki te pokazane są na rysunku 18.1. Przygotowywane są kolejne modele zawierające cztery lub sześć slotów. W tym momencie nie są one jeszcze dostępne.

Rysunek 18.1. Nexus 7018 oraz 7010. Zdjęcie zamieszczono dzięki uprzejmości Cisco Systems, Inc. Nieuprawnione użycie zabronione

Jedną istotną rzeczą, o której należy pamiętać przy projektowaniu centrów danych z wykorzystaniem tych przełączników, jest sposób cyrkulacji powietrza. Nexus 7010, podobnie jak reszta urządzeń z linii Nexus, charakteryzuje się tym, że powietrze przepływa z przodu na tył. Umożliwia to montaż w taki sposób, że front obudowy nie jest wystawiony na ciepłe powietrze.

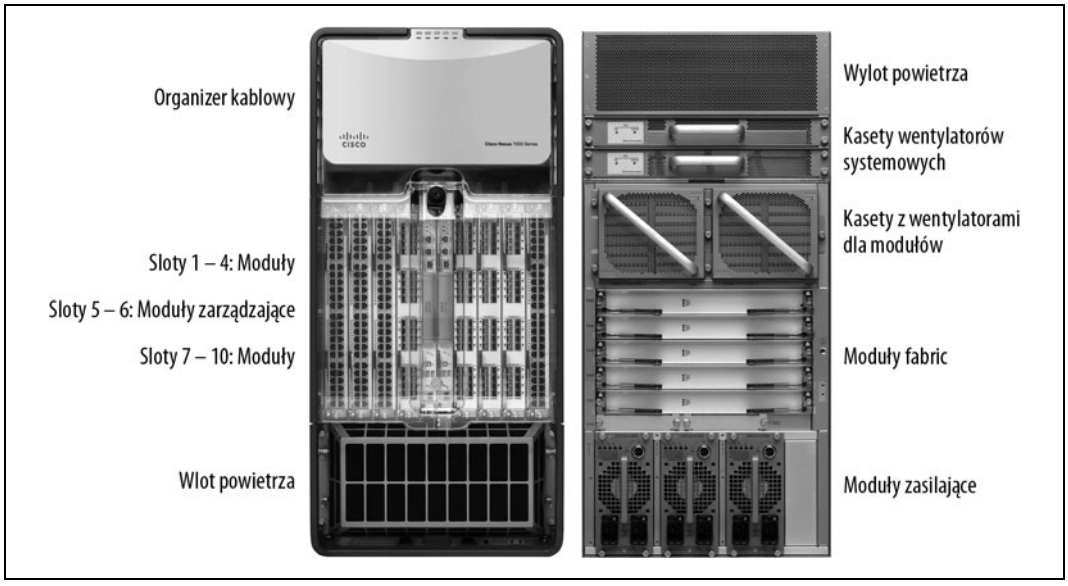
Dla odmiany w przełącznikach 7018 powietrze przepływa przez boczne ściany obudowy. Pytałem o to wiele osób pracujących w Cisco, jednak ich odpowiedź była zawsze taka sama jak uzasadnienie zawarte w dokumentacji urządzenia. 7018 jest zaprojektowany do montażu na końcu rzędu racków i najwyraźniej ma być umieszczany bokiem. Nigdy nie widziałem przełącznika zamontowanego w ten sposób, jestem jednak pewien, że ktoś już tego próbował. Może okazać się to całkiem ważne, ponieważ spotkałem się z sytuacją, gdy nie brano pod uwagę jakiegoś sprzętu tylko dlatego, że przepływ powietrza odbywał się w tym urządzeniu przez ściany boczne. Pierwotne modele 6509 oraz 6509E również charakteryzują się takim przepływem powietrza. Dlatego też niektórzy wykorzystaliby jedynie model 6509-V-E lub starszy

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

model NEBS, ponieważ zastosowano tam „właściwą” cyrkulację powietrza. Jeśli zamierzasz montować te monstra w szafach rackowych, pomyśl nad zaaranżowaniem odpowiedniej ilości pustej przestrzeni po każdej ze stron, aby zapewnić właściwą cyrkulację powietrza i odpowiednie chłodzenie. Z tego, co zdążyłem zauważyć, nowe modele cztero- lub sześcioslotowe przełącznika Nexus 7000 będą również posiadały cyrkulację powietrza przez ściany boczne.

Przełączniki Nexus 7000 są zbudowane z wykorzystaniem tak zwanej architektury **midplane**. Z kolei w przełączniku takim jak 6509 zastosowano architekturę *backplane*. **Backplane** oznacza, że wszystkie moduły, adaptery zasilania oraz wentylatory są zainstalowane od frontu i łączą się wewnątrz z tym panelem przełącznika. Przełączniki 7000 są tak zaprojektowane, że zasilacze oraz wentylatory instalowane są z tyłu, natomiast moduły na froncie. Z tyłu są również instalowane moduły typu *fabric*. Złącza dla tych wszystkich modułów znajdują się na środku przełącznika — stąd **midplane**.

Seria Nexus 7000 jest zaprojektowana tak, aby zminimalizować awaryjność. Jak możesz zobaczyć na rysunku 18.2, w tych przełącznikach znajdują się trzy adaptery zasilania, pięć modułów typu *fabric*, dwie szuflady z wentylatorami do modułów i dwie szuflady z wentylatorami dla systemu. Z zamontowanymi modułami zarządzającymi te przełączniki całkiem dobrze nadają się dla sieci wymagającej wysokiej dostępności.



Rysunek 18.2. Przód i tył przełącznika Nexus 7010. Zdjęcie zamieszczono dzięki uprzejmości Cisco Systems, Inc.; wykorzystanie bez zgody zabronione

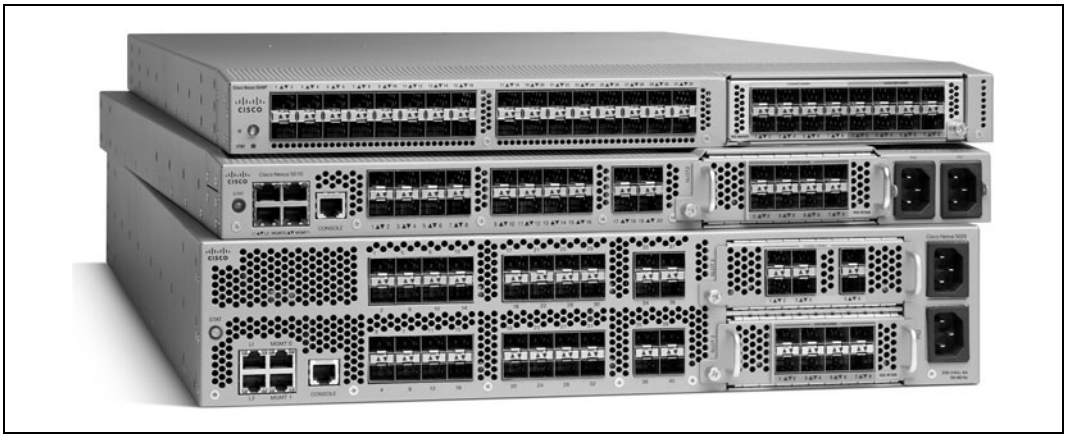
Może wydawać się dosyć dziwne, że moduły *fabric* są odseparowane, skoro wcześniej za dość istotne uznaliśmy zintegrowanie ich z modułami zarządzającymi dla przełączników 6500. W 6500 są tylko dwa moduły *fabric* (po jednym na każdy moduł zarządzający). W przypadku 7010 możemy mieć ich nawet pięć. Czemu jest to takie ważne? Ponieważ każdy moduł fabric zwiększa możliwości przełączania, a w przypadku pięciu slotów oferują one redundancję typu N+1. Oznacza to, że choć potrzebujemy jedynie czterech modułów, to możemy zainstalować ich pięć. Kiedy jeden zawiedzie, nie ma to wpływu na zdolność przełączania. Ponadto możemy unowocześniać przełącznik, instalując bardziej wydajne moduły bez konieczności wyłączania

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

całego przełącznika. W sieciach wymagających prawdziwie wysokiej dostępności tego typu architektura jest koniecznością. Sprzęt sieciowy rzadko oferuje takie funkcje, więc faceci od telefonów od dawna się z nas śmieją, gdyż systemy telefoniczne od dziesięcioleci stosują takie rozwiązania.

Nexus 5000

Obecnie dostępne są trzy modele przełączników z serii Nexus 5000. Są to: 5010, 5020 oraz 5548. Model 5010 jest urządzeniem rackowym o wysokości 1 U oferującym dwadzieścia portów 10 Gb/s i jeden port rozszerzeń. 5020 ma wysokość 2 U i posiada czterdzieści portów 10 Gb/s oraz dwa porty rozszerzeń. Model 5548 ma wysokość 1 U i posiada trzydzieści dwa porty 10 Gb/s oraz jeden port rozszerzeń. Porty rozszerzeń w przypadku 5010 i 5020 są takie same, natomiast port rozszerzeń dla 5548 jest inny. Przełącznik 5548 może posiadać maksymalnie czterdzieści osiem portów 10 Gb/s z przepustowością 960 Gb/s, co odpowiada prędkości transmisji kablowej dla wszystkich czterdziestu ośmiu interfejsów 10 Gb/s. Rodzinę Nexus 5000 przedstawia rysunek 18.3.



Rysunek 18.3. Nexus 5500, 5010 oraz 5020. Zdjęcie zamieszczono dzięki uprzejmości Cisco Systems, Inc. Nieuprawnione użycie zabronione

Przełączniki 5010 i 5020 posiadają po dwa moduły zasilające, mają jednak złącza zasilania umieszczone na tym samym panelu co porty danych. Moduły zasilające z funkcją *hot-swap* znajdują się na przeciwległym panelu przełącznika.

Przełączniki te charakteryzują się cyrkulacją powietrza z frontu na tył. Jednak wiele osób jest zdziwionych, że w prawidłowo zaprojektowanych pod względem cyrkulacji ciepłego i zimnego powietrza centrach danych wychodzi na to, że przełączniki te muszą być instalowane tył na przód. Mam na myśli to, że panel przełącznika, na którym znajdują się porty, jest w rzeczywistości tylnym panelem. Pamiętaj, że te switche są przeznaczone do montażu na górnych poziomach serwerowych szaf rackowych, tak więc porty będą po tej samej stronie co porty serwerów. Przy montażu tych przełączników na własnych uchwytach rackowych (metoda *center-of-row*) wiele osób próbuje je instalować niewłaściwą stroną.

Kolejną rzeczą, na którą trzeba zwracać uwagę w przypadku tych przełączników, jest konieczność ich montażu w taki sposób, aby wystawały one kilka centymetrów poza tylne szyny. Oznacza to, że będą wystawać poza przestrzeń przeznaczoną teoretycznie dla szafy rackowej. Powo-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

dem tego jest ponownie kwestia związana z cyrkulacją powietrza. Wierzchnia pokrywa w tylnej część przełącznika 5010 jest nieco nachylona. Ten fragment górnej pokrywy posiada otwory wentylacyjne, które umożliwiają wydostawanie się powietrza z obudowy, tak jak to pokazano na rysunku 18.4. Przedstawione urządzenie to w rzeczywistości FEX 2148, ale przełącznik 5010 jest zbudowany w ten sam sposób.

Rysunek 18.4. FEX 2148 — widok z boku. Zwróć uwagę na otwory wentylacyjne w pokrywie

Staraj się nie umieszczać kabli w przestrzeniach pomiędzy zainstalowanymi w szafie przełącznikami, ponieważ utrudni to właściwy obieg powietrza oraz chłodzenie. Nie musisz się tym przejmować w przypadku przełączników 5020, gdyż większa część otworów wentylacyjnych jest w tym modelu umieszczona na tylnym panelu.

Mimo iż przełączniki z serii Nexus 5000 zostały zaprojektowane do stosowania w centrach danych, wiele z nich ma, jak się wydaje, całkiem znaczące ograniczenia. Przykładowo przełączniki 5010 i 5020 z wersją systemu operacyjnego NX-OS 5.0 nie obsługują interfejsów SVI oraz routingu i mogą posiadać tylko 512 VLAN-ów. Ograniczenia te mogą słusznie wydawać się irytujące. Przełączniki te zostały jednak zaprojektowane do jednego celu, którym jest podłączenie dużej liczby serwerów do struktury przełączania (i magazynowania danych) w centrach danych. Tę rolę spełniają całkiem dobrze. Nie myśl jednak, że możesz wykorzystać parę przełączników 5000 w roli rdzenia w sieci korporacyjnej LAN. Nie są one po prostu stworzone do tego celu. Nowszy przełącznik 5548 może obsługiwać warstwę trzecią dzięki dodatkowej karcie rozszerzeń typu *field-upgrade*, która posiada sloty na kolejne moduły. Obsługuje on również 4000 VLAN-ów, 16 portów w jednym port channel i szereg innych udogodnień. Jeśli interesują Cię przełączniki Nexus z serii 5000, to radzę wybrać nowsze modele.

Seria przełączników 5000 znacznie różni się od standardowych switchy 1 U w linii produktowej Cisco. Muszą one być instalowane w serwerowych szafach rackowych. Chociaż poprzednie zdjęcie dobrze tego nie pokazuje, przełączniki te są bardzo długie i zajmują przy tym całą dostępną głębokość szafy serwerowej. Przełączniki 5000 po prostu nie mogą być instalowane w standardowych rackach telekomunikacyjnych o dwusłupowym szkielecie. Pamiętaj, że zostały one zaprojektowane do stosowania w centrach danych i tam jest ich miejsce. Rysunek 18.5 przedstawia Nexusa 5010, na którym dla porównania umieszczono przełącznik Catalyst 2960. Zauważ, jaka jest długość 5010. Jest on ponad trzy razy dłuższy niż 2960.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Rysunek 18.5. Catalyst 2960 umieszczony dla porównania na przełączniku Nexus 5010

W przełącznikach 5000 z lewej strony tylnego panelu z portami znajdują się cztery porty RJ-45. Oznaczone są następująco: Mgmt-0, Mgmt-1, L1 oraz L2. Obecnie jedynym konfigurowalnym portem jest interfejs Mgmt-0. Pozostałe trzy interfejsy zostaną wykorzystane jako porty rozszerzeń dopiero w przyszłości. Znajdujący się obok port konsoli (ang. *console port*) działa w standardowy sposób. W przypadku modelu 5548 te cztery porty umieszczone są dla odmiany na przednim panelu przełącznika.

Przy projektowaniu linii Nexus za cel postawiono sobie osiągnięcie wysokiej dostępności przełącznika. Z tego względu moduły zasilające oraz kasety z wentylatorami są *hot-swappable*, czyli mogą być wymieniane bez konieczności wyłączania przełącznika. Rysunek 18.6 pokazuje switch Nexus 5510 po usunięciu modułów zasilających oraz kaset z wentylatorami.

Rysunek 18.6. Nexus 5010 po usunięciu modułów zasilających oraz kaset z wentylatorami

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

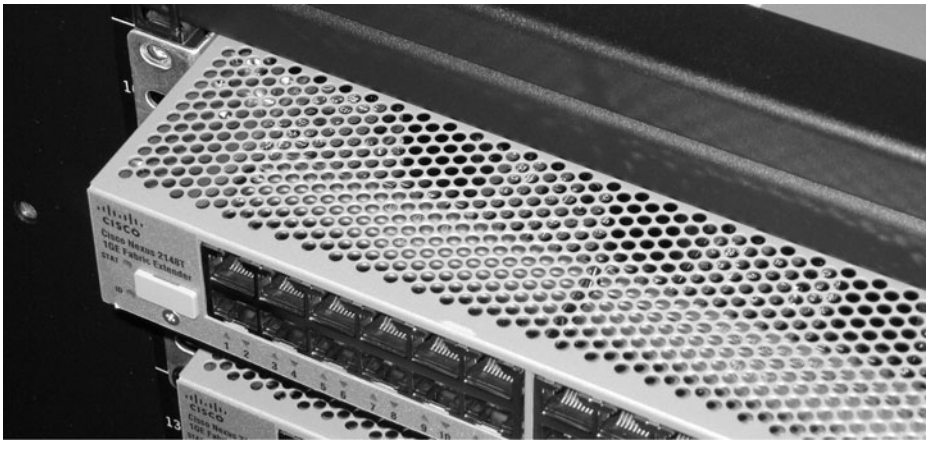
Nexus 2000

Urządzenia Nexus 2000 nie są właściwie przełącznikami, a raczej rozszerzeniami strukturalnymi (ang. *fabric extenders* — FEX). Urządzenia FEX mogą być traktowane jak zewnętrzne moduły dla przełączników Nexus. W przypadku Nexus 5000 możesz podłączyć do 12 urządzeń FEX. To tak, jakby Nexus 5020 posiadał 14-slotową obudowę, przy czym „sloty" nie muszą być zamontowane w tych samych rackach co „moduł zarządzający". Możesz połączyć ze sobą dwa przełączniki Nexus 5000, co umożliwi podłączenie 24 urządzeń FEX w przypadku dwuadapterowych systemów serwerowych. Takie rozwiązanie może być niesamowicie elastyczne. Jednak nie jest ono pozbawione wad, o czym przekonasz się w podrozdziale poświęconym konfiguracji. Nowsze modele Nexus i nowsze wersje NX-OS mogą obsługiwać nawet większą liczbę urządzeń FEX.

Ponieważ FEX nie są przełącznikami, nie posiadają własnej inteligencji. Nie mają nawet portu konsoli. Aby skorzystać z FEX, musisz podłączyć go do przełącznika Nexus wyższego poziomu, takiego jak 5000 lub 7000. Jeśli potraktujesz urządzenia FEX jako zewnętrzne moduły rozszerzające możliwości przełącznika, łatwiej będzie Ci zrozumieć sposób ich działania.

Uważaj na pierwotne modele FEX 2148 48-portów 1 Gb/s, gdyż mają one kilka istotnych ograniczeń. Najbardziej irytujący jest brak obsługi kanałów EtherChannel w obrębie pojedynczego FEX-u. Nie mniej uciążliwe jest, że ich gigabitowe porty nie obsługują prędkości 100 Mb/s oraz 10 Mb/s. Cisco wycofało już te urządzenia i oferuje znacznie ulepszony model FEX 2248, w którym rozwiązano te problemy. Pamiętaj tylko, że aby skorzystać z 2248 dla przełączników 5000, musisz mieć na tych przełącznikach zainstalowaną wersję NX-OS o kodzie 4.2(1)N1(1) lub nowszą.

Urządzenia z serii 2000 mają podobne wymagania odnośnie do cyrkulacji powietrza w szafie rackowej co seria 5000. FEX zamontowany w szafie z kontrolą przepływu powietrza pokazany jest na rysunku 18.7.



Rysunek 18.7. FEX 2148 zamontowany w szafie, z barierą przepływu powietrza znajdującą się nad nim

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

FEX-y nie są przełącznikami i nie należy ich w ten sposób traktować. Ich podstawowym celem jest podłączenie hostów do sieci. To rozróżnienie dobrze objaśnia natura interfejsów FEX, które można nazwać **portami hosta** (ang. *host ports*) lub **portami fabric** (ang. *fabric ports*):

Porty hosta
: Porty hosta są przeznaczone do podłączania do sieci serwerów (hostów). Z tego względu porty hosta mają włączoną funkcję Spanning Tree BPDU Guard odpowiedzialną za wyłączanie portu, do którego został podłączony kolejny przełącznik. Ważne jest, byś to zrozumiał, ponieważ BPDU Guard nie może być w FEX-ie wyłączony.

Porty fabric
: Jedynym celem portów fabric jest podłączenie przełącznika 5000 lub 7000. Taki port nie jest bezpośrednio konfigurowalny, może być jednak wykorzystany, kiedy na podłączonym porcie switch 5000 ustawiony jest tryb fabric. Opiszę to bardziej szczegółowo w jednym z następnych podrozdziałów. Porty fabric obsługują jedynie prędkość 10 Gb/s i nie mogą być wykorzystane dla hostów.

Obecnie FEX-y dostępne są w trzech wersjach[1]:

- Nexus 2148T GE Fabric Extender,
- Nexus 2232PP 10 GE Fabric Extender,
- Nexus 2248TP GE Fabric Extender.

Jak już wcześniej wspominałem, wszystkie urządzenia Nexus zostały zaprojektowane z przeznaczeniem do stosowania w centrach danych. Z tego powodu posiadają podwójne moduły zasilające i kasetę wentylatorów z obsługą *hot-swap*.

Rysunek 18.8 przedstawia front urządzenia FEX, z którego wysunięto dwa moduły zasilające i kasetę z wentylatorami.

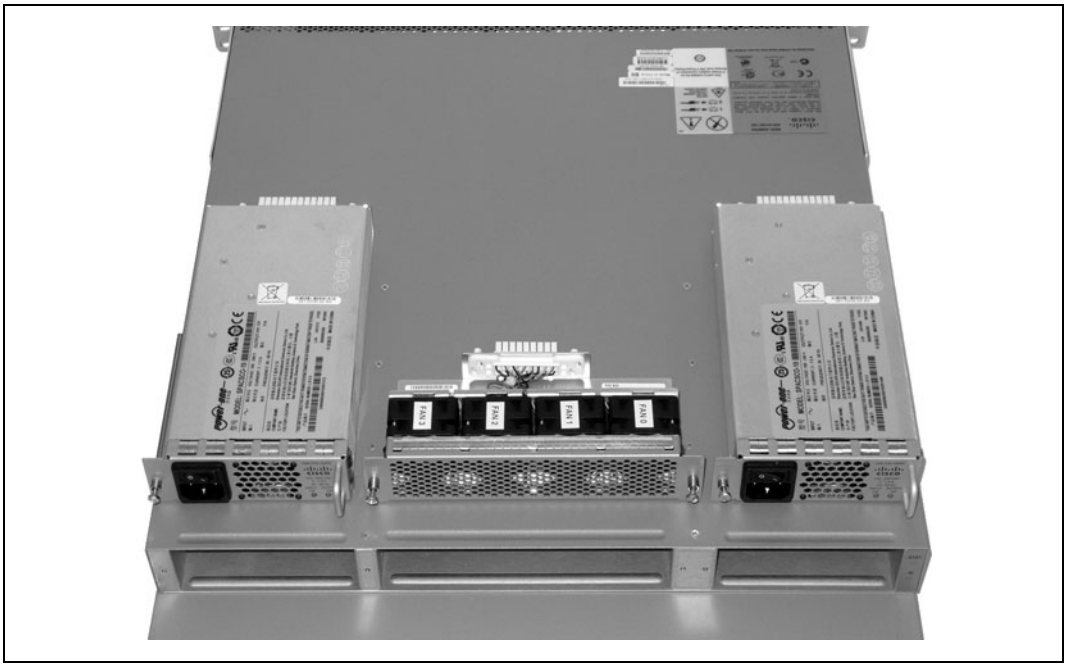
Seria Nexus 1000

Urządzenia z serii Nexus 1000 umiejscowić można właściwie na poziomie serwerów. Mamy do wyboru dwa podstawowe modele: przełącznik wirtualny 1000V do stosowania z oprogramowaniem VMware oraz urządzenie 1010V do zastosowań serwerowych.

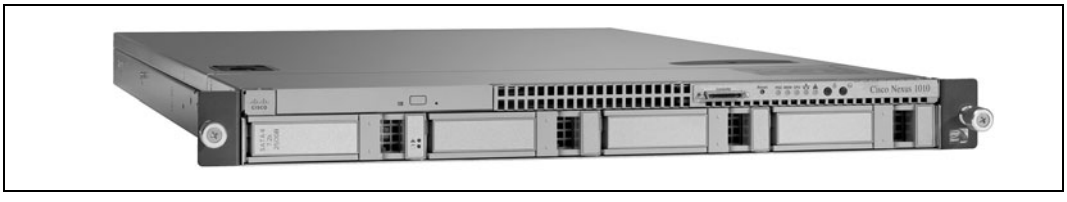
Nexus 1000V jest przełącznikiem programowym przeznaczonym do pracy z programem VMware, w szczególności jako wsparcie dla architektury chmury. 1010 Virtual Services Appliance (VSA) jest praktycznie VSM-em, który kontroluje wirtualne przełączniki VMware w systemie UCS (ang. *Unified Computing System*) lub serwery innych producentów. Rysunek 18.9 przedstawia urządzenie Cisco Nexus 1010 VSA. W tym rozdziale nie będę już opisywał tych urządzeń bardziej szczegółowo.

[1] W momencie wydania książki w Polsce firma Cisco oferuje większą liczbę modeli urządzeń tego typu — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 18.8. Nexus 2148 FEX z wyjętymi modułami zasilającymi oraz kasetą z wentylatorami



Rysunek 18.9. Cisco Nexus 1010 VSA. Zdjęcie zamieszczono dzięki uprzejmości Cisco Systems, Inc. Nieuprawnione użycie zabronione

NX-OS

NX-OS to system operacyjny przełączników Nexus. NX-OS został napisany od podstaw jako modularny, skalowalny oraz elastyczny system operacyjny dla następnej generacji przełączników stosowanych w centrach danych.

NX-OS wygląda podobnie do IOS. Jest na tyle zbliżony do IOS, abyś poczuł się dość komfortowo, ale jednocześnie na tyle odmienny, że jest ciekawy, ale i całkowicie frustrujący, dopóki się do niego nie przyzwyczaisz. Wiele poleceń jest takich samych, jednak informacje, których dostarczają, i sposób ich przedstawienia nie jest taki, jakiego byś się spodziewał, mając pewne doświadczenie z IOS. Znajdą się też komendy, które różnią się całkowicie od znanych Ci z IOS, a w niektórych przypadkach są nawet sprzeczne z tym, czego nauczyłeś się z innych urządzeń Cisco.

Istnieje wiele stron internetowych poświęconych licznym różnicom pomiędzy IOS oraz NX-OS i mógłbym prawdopodobnie zapełnić całą książkę, wyliczając te różnice. Skoncentruję się jednak na najbardziej powszechnych sytuacjach, które mogą wyprowadzać na manowce osoby dopiero

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

uczące się systemu NX-OS. Niektóre z tych przypadków stanowią podstawę moich ulubionych pytań z serii „A więc wydaje Ci się, że masz pojęcie o przełącznikach Nexus?”, które zwykłem zadawać podczas rozmów kwalifikacyjnych. Jeśli więc dostaniesz pracę, możesz wysłać mi podziękowania.

NX-OS kontra IOS

NX-OS jest oparty na Linuksie. Dzięki temu korzysta z wielu funkcji, którymi cieszą się użytkownicy Linuksa, w tym nie tylko ze stabilności i modularności. Pierwszą rzeczą, która zazwyczaj jest powodem frustracji wielu osób, jest właśnie modularność NX-OS. NX-OS został zaprojektowany tak, aby był jak najbardziej wydajny, z tego względu system operacyjny nie instaluje niepotrzebnych modułów. Dla przykładu, jeśli chcesz uruchomić protokół OSPF, musisz najpierw zainstalować moduł OSPF. Jest to odejście od IOS, gdzie wszystkie usługi są instalowane bez względu na potrzeby. Jak możesz sobie pewnie wyobrazić, ładowanie pełnego kodu dla OSPF, BGP i niezliczonej ilości innych funkcji, których nie potrzebujesz, może być nieefektywne. Jeśli kiedykolwiek kompilowałeś samodzielnie jądro Linuksa pod kątem szybkości i niezawodności, z pewnością dostrzeżesz korzyści, jakie przynosi ta metoda.

W NX-OS moduły systemu zwane są **funkcjami** (ang. *features*). Aby załadować moduł, zastosuj polecenie `feature`. Komenda `show feature` pokaże Ci, jakie moduły są dostępne. Poniżej lista funkcji dla Nexus 5000:

```
NX-5K-1# sho feature
Feature Name          Instance  State
--------------------  --------  --------
tacacs                1         disabled
lacp                  1         enabled
interface-vlan        1         enabled
private-vlan          1         disabled
udld                  1         disabled
vpc                   1         enabled
fcoe                  1         disabled
fex                   1         enabled
```

Pracując nad tym podrozdziałem, spróbowałem wprowadzić komendę w niewłaściwym trybie konfiguracyjnym. Nexus jest na tyle inteligentny, żeby wykryć, jaką popełniłem głupotę, a następnie skomentować to, przedstawiając mi jednocześnie informacje, których szukałem. Poniżej przykład próby zastosowania polecenia `feature` w trybie konfiguracyjnym interfejsu:

```
NX-7K-1(config-if)# feature ?
*** No matching command found in current mode, matching in (config) mode ***
      //Niewłaściwa komenda w bieżącym trybie konfiguracyjnym. Komenda powinna być wprowadzana w trybie (config).
  bgp             Enable/Disable Border Gateway Protocol (BGP)
  cts             Enable/Disable CTS
  dhcp            Enable/Disable DHCP Snooping
[fragment usunięty]
```

Oto lista funkcji dla Nexusa 7000. Jak widzisz, większy przełącznik posiada bardziej rozbudowany zestaw funkcji:

```
NX-7K-1(config)# feature ?
  bgp             Enable/Disable Border Gateway Protocol (BGP)
  cts             Enable/Disable CTS
  dhcp            Enable/Disable DHCP Snooping
  dot1x           Enable/Disable dot1x
  eigrp           Enable/Disable Enhanced Interior Gateway Routing Protocol
                  (EIGRP)
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
eou             Enable/Disable eou(l2nac)
glbp            Enable/Disable Gateway Load Balancing Protocol (GLBP)
hsrp            Enable/Disable Hot Standby Router Protocol (HSRP)
interface-vlan  Enable/Disable interface vlan
isis            Enable/Disable IS-IS Unicast Routing Protocol (IS-IS)
lacp            Enable/Disable LACP
msdp            Enable/Disable Multicast Source Discovery Protocol (MSDP)
netflow         Enable/Disable NetFlow
ospf            Enable/Disable Open Shortest Path First Protocol (OSPF)
ospfv3          Enable/Disable Open Shortest Path First Version 3 Protocol
                (OSPFv3)
pbr             Enable/Disable Policy Based Routing(PBR)
pim             Enable/Disable Protocol Independent Multicast (PIM)
pim6            Enable/Disable Protocol Independent Multicast (PIM) for IPv6
port-security   Enable/Disable port-security
private-vlan    Enable/Disable private-vlan
rip             Enable/Disable Routing Information Protocol (RIP)
```

Są takie funkcje, które mogą Cię zaskoczyć. Przykładowo, jeśli chcesz użyć interfejsów SVI, musisz włączyć tę funkcję poleceniem `feature interface-vlan`. Elementy, które mogłyby wydawać się oczywiste, jak HSRP, są w przypadku Nexusa funkcjami. Ma to sens, przy założeniu, że SVI niezbędne do działania HSRP jest również funkcją. Pierwsza rzecz, którą mówię młodszym inżynierom, kiedy pracują na urządzeniach Nexus, to: jeśli to, nad czym pracujesz, nie chce działać, sprawdź, czy wszystkie niezbędne funkcje zostały załadowane. Taka sytuacja przytrafia się każdemu (mnie również) przynajmniej kilka razy, zanim włączanie wymaganych funkcji stanie się rutyną.

Koniec z nazwami interfejsów opisującymi ich prędkość

Każdy interfejs Ethernet nazywa się w NX-OS **Ethernet**. Koniec z FastEthernet, GigabitEthernet czy TenGigabitEthernet. Bez względu na prędkość, każdy interfejs to teraz po prostu **Ethernet**. Jest to irytujące dla tych, którzy używali IOS od wersji 9. Jednak po raz kolejny Cisco nie poprosiło o moje wsparcie przy projektowaniu swojego nowego systemu operacyjnego. Trochę to trwało, ale wreszcie nie widzę setek 10-megabitowych interfejsów, kiedy przeglądam konfigurację na przełączniku Nexus.

Z pewnością nie będzie to problemem dla młodszych, wyluzowanych inżynierów, którzy nigdy nie pracowali z routerami Cisco 2501. Do licha, ja miałem do czynienia nawet z routerami AGS posiadającymi wewnętrzne sloty rozszerzeń jak w PC. Routery te nie miały również parsera wiersza poleceń. Były to jednak czasy przed 100-megabitowym Ethernetem, tak więc możecie mi już odpuścić.

W NX-OS nie ma ukrytej konfiguracji

Kiedy zmienia się funkcja interfejsu (np. z warstwy drugiej na trzecią), zdarza się, że IOS przechowuje oryginalną konfigurację, która zostaje przywrócona wraz z przywróceniem pierwotnej funkcji. NX-OS tego nie robi.

Status interfejsu przedstawia jego stan wraz z podaniem przyczyny

Bardzo lubię tę funkcję. Komenda `show interface` wyświetla status interfejsu oraz gdy interfejs nie jest aktywny, przedstawia raport z podaniem przyczyny:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-7K-1-Cozy(config-if-range)# sho int e7/3
Ethernet7/3 is down (SFP not inserted) //Interfejs Ethernet7/3 nie jest aktywny (brak modułu SFP)
  Hardware: 10000 Ethernet, address: 0026.9807.95c2 (bia 5475.d09a.81de)
  MTU 1500 bytes, BW 10000000 Kbit, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 1/255
[fragment usunięty]
```

Komenda show interface brief

Nie wiem, jak u Ciebie, ale w moim przypadku komenda `show ip int brief` prawdopodobnie znalazłaby się na liście 10 komend, których używam codziennie. Byłem zbity z tropu, kiedy nie zadziałała ona tak, jak tego oczekiwałem. Patrzyłem więc tępo w ekran przez dłuższy czas i przewijałem tekst strzałkami w górę i w dół, mając nadzieję, że listing się jakoś zmieni. Kiedy już się pozbierałem, zrobiłem coś niesamowitego: zajrzałem do instrukcji. Tam odkryłem polecenie `show interface brief`:

```
NX-7K-1-Daisy# sho int brief
--------------------------------------------------------------------------------
Ethernet      VLAN   Type Mode   Status  Reason                   Speed     Port
Interface                                                                    Ch #
--------------------------------------------------------------------------------
Eth3/25       --     eth  routed up      none                       1000(D) --
Eth3/26       --     eth  routed up      none                       1000(D) --
Eth3/27       --     eth  routed down    Link not connected         auto(D) --
Eth3/28       --     eth  routed down    Administratively down      auto(D) --
Eth3/29       --     eth  routed down    Administratively down      auto(D) --
Eth3/30       1      eth  access up      none                       1000(D) 10
Eth3/31       --     eth  routed down    Administratively down      auto(D) --
Eth3/32       --     eth  routed down    Administratively down      auto(D) --
Eth3/33       --     eth  routed down    Administratively down      auto(D) --
Eth3/34       --     eth  routed down    Administratively down      auto(D) --
Eth3/35       --     eth  routed down    Administratively down      auto(D) --
Eth3/36       --     eth  routed down    Administratively down      auto(D) -
```

Jest to bardzo sprytna komenda. Nie pokazuje opisów, ale za to informuje o zastosowanych VLAN-ach, statusie up/down, prędkości oraz numerze port channel, do którego należy dany interfejs. W powyższym przykładzie możesz zauważyć, że e3/30 należy do port channel 10.

Brak komendy write memory

Cisco od dłuższego czasu straszyło usunięciem komendy `write memory`, jednak ci z nas, którzy od wielu lat pracują z tym sprzętem, zdążyli się do niej przyzwyczaić. Niestety urządzenia Nexus nie obsługują już tej komendy:

```
NX-5K-1# write ?
  erase  Destroys the configuration on persistent media //Parametr erase usuwa konfigurację z nośnika trwałego
```

Teraz musisz zastosować polecenie `copy running startup`, aby zapisać konfigurację. Pojawił się pasek postępu, który zapełnia się w miarę zapisywania konfiguracji, co jest dość miłe:

```
NX-5K-1# copy running-config startup-config
[#######################################] 100%
```

Brak komendy do

Komenda `do` nie jest już potrzebna. NX-OS dopuszcza wykonanie każdej komendy niekonfiguracyjnej w trybie konfiguracyjnym:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-5K-2(config-if)# sho int po 100 switchport
Name: port-channel100
  Switchport: Enabled
  Switchport Monitor: Not enabled
  Operational Mode: access
  Access Mode VLAN: 1 (default)
 [fragment usunięty]
```

Brak komendy show arp

Komenda `show arp` została zastąpiona komendą `show ip arp`. Ma to sens, biorąc pod uwagę, że ARP jest protokołem odnoszącym się do IP. Niektóre przełączniki obsługują obie te komendy, ale nie Nexus. Dodatkowo może zajść konieczność zdefiniowania VRF (więcej informacji znajdziesz w podrozdziale „Virtual Routing and Forwarding" na stronie 279):

```
NX-7K-1-Daisy# sho ip arp vrf Earth
IP ARP Table for context Earth
Total number of entries: 1
Address         Age       MAC Address     Interface
10.0.0.2        00:09:39  0026.981b.f043  Ethernet3/25
```

Brak komendy interface range

Co? Uwielbiałem tę komendę! Jak mogli usunąć jedną z najbardziej praktycznych komend!? Spokojnie, komendy może już nie ma, ale jej funkcjonalność pozostała. I jest nawet lepsza niż przedtem. Funkcjonalność ta nie wymaga teraz oddzielnej komendy. Jak to możliwe? Po prostu podaj zakres, wpisując komendę `interface`, a przełącznik sam przejdzie w tryb konfiguracji zakresu interfejsów (ang. *interface-range mode*):

```
NX-7K-1-Cozy(config-vlan)# int e3/1 - 2
NX-7K-1-Cozy(config-if-range)#
```

Obsługa zapisu z ukośnikiem dla wszystkich komend IPv4 oraz IPv6

Myślę, że to bardzo dobrze. Ubolewam jednak nad tym, że inżynierowie, którzy już teraz nie potrafią dokonać w pamięci podziału na podsieci, będą mieli jeszcze jeden argument, żeby się nie uczyć. Ach, ten postęp. Czy nie mówiłem już, byście się mnie nie czepiali?

```
NX-7K-1-Cozy(config)# int 3/10
NX-7K-1-Cozy(config-int)# ip address 10.255.255.2/24
```

Dwa poziomy pomocy w wierszu poleceń

W wierszu poleceń (CLI) NX-OS wciąż możesz użyć znaku zapytania, aby wyświetlić dostępne komendy na danym poziomie. Poniżej krótki przykład, który wyda się znajomy osobom posiadającym doświadczenie z powłoką bash w Uniksie. Na początek zastosuję coś, co w Uniksie nazywa się **dopełnienie tabulatorem** (ang. *tab completion*). Wpisuję pierwszych kilka liter z nazwy komendy i wciskam *Tab*, aby zobaczyć, jakie są opcje:

```
NX-7K-1-Cozy(config)# tacacs<TAB>
tacacs+         tacacs-server
```

Jest to miłe, bo szybkie i proste. Jeśli wolisz tradycyjne opisy z IOS dla każdej dostępnej komendy, po prostu użyj jak zawsze znaku zapytania:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-7K-1-Cozy(config)# tacacs?
  tacacs+        Enable tacacs+
  tacacs-server  Configure TACACS+ server related parameters
```

Konfiguracja routingu głównie z poziomu interfejsów

Procesy routera są nadal uruchamiane globalnie:

```
NX-7K-1-Cozy(config)# router ospf 100
```

Jednak znaczna część poleceń wprowadzana jest teraz w trybie konfiguracji interfejsu:

```
NX-7K-1-Cozy(config-router)# int e3/10
NX-7K-1-Cozy(config-if)# ip router ospf 100 area 0
NX-7K-1-Cozy(config-if)# ip ospf hello-interval 5
```

Komendy BGP nadal znajdują się w bloku BGP, ponieważ są one związane z sąsiadowaniem urządzeń.

Więcej argumentów dla komend show

NX-OS posiada więcej użytecznych argumentów dostępnych dla poleceń `show`, niż znajdziesz w IOS. Oto lista z Nexusa 7000:

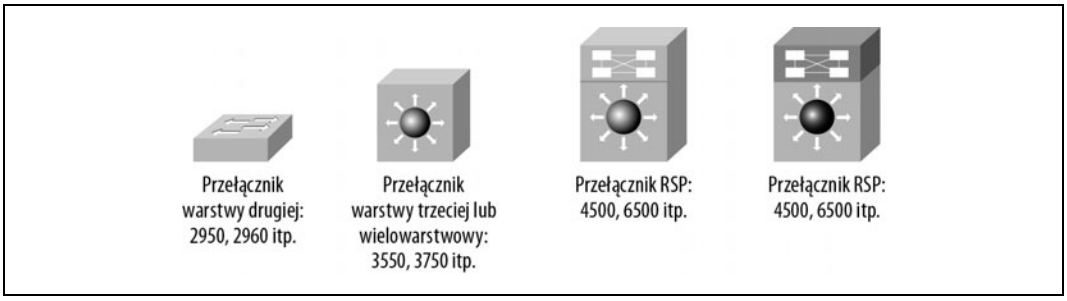
```
Daisy-Cozy# sho run | ?
  cut      Print selected parts of lines.
  diff     Show difference between current and previous invocation (creates
           temp files: remove them with 'diff-clean' command and dont use it
           on commands with big outputs, like 'show tech'!)
  egrep    Egrep - print lines matching a pattern
  grep     Grep - print lines matching a pattern
  head     Display first lines
  human    Output in human format (if permanently set to xml, else it will turn
           on xml for next command)
  last     Display last lines
  less     Filter for paging
  no-more  Turn-off pagination for command output
  section  Show lines that include the pattern as well as the subsequent lines
           that are more indented than matching line
  sed      Stream Editor
  sort     Stream Sorter
  tr       Translate, squeeze, and/or delete characters
  uniq     Discard all but one of successive identical lines
  vsh      The shell that understands cli command
  wc       Count words, lines, characters
  xml      Output in xml format (according to .xsd definitions)
  begin    Begin with the line that matches
  count    Count number of lines
  end      End with the line that matches
  exclude  Exclude lines that match
  include  Include lines that match
```

Doskonałym nowym poleceniem jest `last`, które jest podobne do `tail` w Uniksie. Kolejnym jest `egrep`, które z pewnością dostarczy miłośnikom Uniksa powodów do radości. Znajdziesz też polecenie `no-more` umożliwiające przechwytywanie stron listingu bez konieczności ustawiania `term length 0`. Wystarczy, że pobawisz się trochę tymi funkcjami, a z pewnością poczujesz frustrację, wracając do IOS i jego nudnych, ograniczonych opcji wyświetlania informacji.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

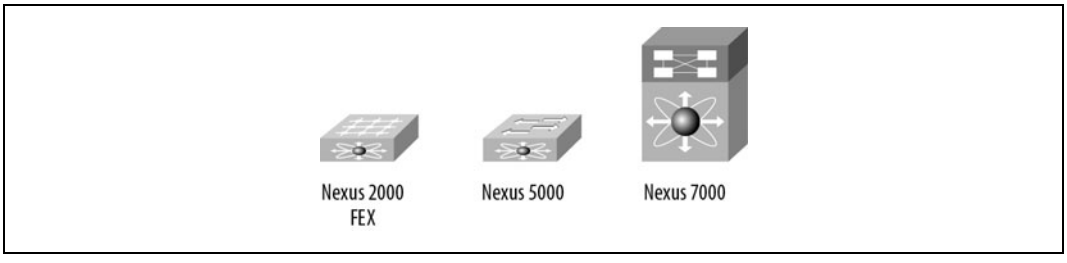
Ikonografia Nexusa

Zanim zagłębię się w szczegóły dotyczące podstaw architektury urządzeń Nexus, muszę wprowadzić kilka nowych ikon przedstawiających te urządzenia. Na rysunku 18.10 znajdują się ikony, których używałem dotychczas w swoich diagramach do przedstawiania tradycyjnych przełączników.



Rysunek 18.10. Tradycyjne ikony urządzeń sieciowych w stylu Cisco

Cisco wprowadziło nową ikonografię specjalnie dla przełączników Nexus. Nowe ikony pokazane są na rysunku 18.11. Zauważ, że wszystkie ikony Nexusa posiadają ten sam symbol przypominający model budowy atomu, z dodanymi strzałkami w czterech kierunkach.



Rysunek 18.11. Nowe ikony Nexusa

Cechy architektury Nexusa

Przełączniki Nexus zostały od podstaw zaprojektowane w taki sposób, aby zapewnić dużą elastyczność oraz skalowalność. Architektura Nexusa charakteryzuje się następującymi cechami:

Płaszczyzna zarządzania i kontroli jest niezależna od płaszczyzny danych

W skrócie oznacza to, że przekazywanie pakietów jest procesem oddzielnym od zarządzania i obsługi systemu. Kiedy jesteś podłączony do przełącznika sesją SSH i wprowadzasz komendy konfiguracyjne, sprzęt może nadal przekazywać pakiety bez konieczności wykorzystywania CPU. Umożliwia to takie działania jak wykonanie aktualizacji systemu operacyjnego przełącznika bez przerywania procesu przekazywania pakietów.

Obsługa zarządzania out-of-band

Seria 5000 posiada dedykowany interfejs zarządzający, który w modelach nieobsługujących warstwy trzeciej jest jedynym portem z możliwością przypisania adresu IP. Ten port posiada własną instancję VRF (więcej informacji znajdziesz w podrozdziale „Virtual Routing and Forwarding” na stronie 279).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zarządzanie lights-out w serii 7000

Seria Nexus 7000 posiada specjalny port z dedykowanym procesorem, który pozostaje włączony nawet po wyłączeniu zasilania przełącznika (przypomina to funkcję, jaką serwery posiadają od lat). Jest to port CMP (ang. *Connectivity Management Port*), który umożliwia kontrolowanie przełącznika dzięki całkowicie niezależnej od NX-OS funkcjonalności **lights-out**.

Duża prędkość płyty backplane w serii 7000

Oryginalne moduły *fabric* przeznaczone dla przełączników z serii Nexus 7000 zapewniają prędkość do 230 Gb/s na slot oraz redundancję N+1.

Redundancja N+1 w serii 7000

Redundancja N+1 oznacza, że istnieje dodatkowy moduł, który jest aktywowany w przypadku awarii modułu podstawowego. Tak więc, jeśli zasadniczo potrzebujesz dwóch modułów *fabric*, możesz zainstalować w sumie trzy moduły — ten trzeci będzie modułem redundantnym.

Sprzętowy silnik przekazywania pakietów (ang. ***Hardware forwarding engine*****)**
na każdym module we/wy (seria 7000)

Każdy moduł posiada silnik dystrybucji przekazywania pakietów, który umożliwia przełącznikowi przekazywanie pakietów bez wykorzystywania CPU, pod warunkiem że istnieją tabele przekazywania.

Skalowalność

Obudowy przełączników Nexus 7000 zostały zaprojektowane w taki sposób, że nowsze i bardziej wydajne moduły *fabric* mogą być instalowane podczas pracy przełącznika. Takie moduły mogą zwiększyć dostępną przepustowość na slotach oraz wprowadzić obsługę dodatkowych funkcji.

Wirtualizacja

Nexus obsługuje instancje VRF, wirtualne urządzenia VDC (ang. *virtual device contexts*) oraz vPC (ang. *virtual port channels*). Wszystkie te funkcje są opisane w niniejszym rozdziale.

Elastyczność

W odpowiednich warunkach przełączniki Nexus mogą być unowocześniane sprzętowo bez wpływu na przepływ pakietów. Procesy są umiejscowione w wydzielonych obszarach pamięci chronionej (ang. *protected-memory space*). Płaszczyzna kontroli jest niezależna od płaszczyzny danych. Wszystkie te funkcje oraz jeszcze wiele innych czynią ten przełącznik bardziej odpornym na awarie. Widziałem przełączniki Nexus 7000 pracujące na poziomie 100 procent wykorzystania procesora, co jednak nie miało wpływu na przepływające przez nie pakiety. Jest to możliwe, ponieważ przetwarzane przez procesor tabele routingu są następnie tłumaczone na tabele przekazywania (ang. *forwarding tables*), a te są z kolei wysyłane do odpowiednich modułów. Dopóki tabele przekazywania się nie zmieniają, CPU nie jest wykorzystywany w procesie przekazywania pakietów.

Integracja routingu warstwy drugiej, przełączania w warstwie trzeciej oraz przełączania SAN

Chociaż temat ten nie jest omówiony w niniejszej książce, warto wspomnieć, że Nexus został zaprojektowany tak, aby oferować obsługę sieci pamięci masowych (ang. *storage networking*) łącznie z protokołem FCoE.

Obsługa procesorów wielordzeniowych i wielowątkowych

Architektura procesorów wielordzeniowych zapewnia znaczący wzrost mocy obliczeniowej.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Virtual Routing and Forwarding

Instancje VRF są odrębnymi i różnymi procesami routingu w obrębie jednego routera. W skrócie, router może posiadać wiele tabeli routingu, z których każda ma własne protokoły i zestaw tras. Nie jest to cecha charakterystyczna wyłącznie dla Nexusa. IOS w wersji 15 również obsługuje VRF.

IOS w wersji 15? Myślałem, że jesteśmy przy wersji 12.x. Co się stało z wersjami 13 i 14? Najwyraźniej trzynastka jest pechowa podobnie jak w niektórych kulturach czternastka. Cisco przeskoczyło więc z 12 od razu do 15. Możemy chyba teraz wrócić do rozważań nad materią tego rozdziału.

Nexus 5000 posiada domyślnie dwie instancje VRF: VRF domyślny (ang. *default VRF*) oraz VRF zarządzający (ang. *management VRF*). Poniżej przypiszemy trasę domyślną do VRF-u zarządzającego na przełączniku 5000. Zauważ, że adres IP przypisany do interfejsu zarządzającego znajdzie się również w VRF-ie zarządzającym:

```
vrf context management
  ip route 0.0.0.0/0 10.1.1.1
```

Bądź ostrożny z instancjami VRF, szczególnie w przypadku serii 5000. Trzeba zacząć od tego, że jak dotąd przełączniki z tej serii właściwie nie obsługują routingu (najnowszy model 5500 będzie obsługiwał routing dzięki dodatkowej karcie rozszerzeń). Ich obecne iteracje preferują umieszczanie adresu zarządzającego w zarządzającym VRF-ie. Prawdopodobnie sparzysz się na tym niejednokrotnie, ponieważ VRF zarządzający nie jest domyślnym VRF-em.

Oto doskonały przykład. Każdy lubi komendę `show ip interface brief`, wypróbujmy ją więc na skonfigurowanym przełączniku Nexus 5000:

```
NX-5K-1# sho ip int brie
IP Interface Status for VRF "default" //Status interfejsu IP dla VRF: „domyślny"
Interface            IP Address      Interface Status
```

Hmm, to dziwne. Wiem, że skonfigurowałem interfejs zarządzający, czemu więc go nie widać? Ponieważ domyślny VRF został nazwany „domyślny" (sprytne, co?), ale interfejs zarządzający znajduje się w VRF-ie zarządzającym. Aby go zobaczyć, musisz dodać argument `vrf management` do komendy `show ip interface brief`:

```
NX-5K-1# sho ip int brie vrf management
IP Interface Status for VRF "management" //Status interfejsu IP dla VRF: „zarządzający"
Interface            IP Address      Interface Status
mgmt0                10.10.10.1      protocol-up/link-up/admin-up
```

Nie jesteś sam, jeśli pomyślałeś: „To głupie". Pamiętaj jednak, że VRF zarządzający został zaprojektowany, aby być „poza pasmem" (ang. *out-of-band*) przełącznika, dlatego też znajduje się właśnie tutaj. Ma to sens, szczególnie jeśli używasz przełącznika 7000 zamiast 5000. Może to być jednak frustrujące, gdyż jest rozwiązaniem niestandardowym. Jeśli Twoja komenda nie działa, sprawdź VRF-y. To powinno pomóc. Nawet polecenie `ping` wymaga określenia VRF-u:

```
NX-5K-1# ping 10.10.10.2 vrf management
PING 10.10.10.2 (10.10.10.2): 56 data bytes
64 bytes from 10.10.10.2: icmp_seq=0 ttl=254 time=4.762 ms
64 bytes from 10.10.10.2: icmp_seq=1 ttl=254 time=8.052 ms
64 bytes from 10.10.10.2: icmp_seq=2 ttl=254 time=8.078 ms
64 bytes from 10.10.10.2: icmp_seq=3 ttl=254 time=8.168 ms
64 bytes from 10.10.10.2: icmp_seq=4 ttl=254 time=8.178 ms
```

Na szczęście możesz zmienić bieżący kontekst routingu poleceniem `routing-context vrf` *`vrf_nazwa`*:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-5K-1# routing-context vrf management
NX-5K-1%management#
```

Teraz każda kolejna komenda będzie stosowana dla VRF-u zarządzającego:

```
NX-5K-1%management # ping 10.10.10.2
PING 10.10.10.2 (10.10.10.2): 56 data bytes
64 bytes from 10.10.10.2: icmp_seq=0 ttl=254 time=4.762 ms
64 bytes from 10.10.10.2: icmp_seq=1 ttl=254 time=8.052 ms
64 bytes from 10.10.10.2: icmp_seq=2 ttl=254 time=8.078 ms
64 bytes from 10.10.10.2: icmp_seq=3 ttl=254 time=8.168 ms
64 bytes from 10.10.10.2: icmp_seq=4 ttl=254 time=8.178 ms
```

Minusem jest dość brzydki nagłówek wiersza poleceń, ale jest to cena, którą trzeba zapłacić za lenistwo przy wpisywaniu komend. Aby przywrócić poprzedni nagłówek, użyj polecenia `routing-context vrf default`:

```
NX-5K-1%management# routing-context vrf default
NX-5K-1#
```

Instancje VRF mogą być bardzo pożyteczne do tworzenia takich funkcjonalności jak wielodzierżawność w centrach danych, która polega na współdzieleniu przez wielu użytkowników końcowych tej samej usługi i zasobu. Możliwość podłączenia na tym samym przełączniku 10 klientów, z których każdy wykorzystuje tę samą przestrzeń adresową 10.10.10.0/24, jest całkiem niezłą korzyścią.

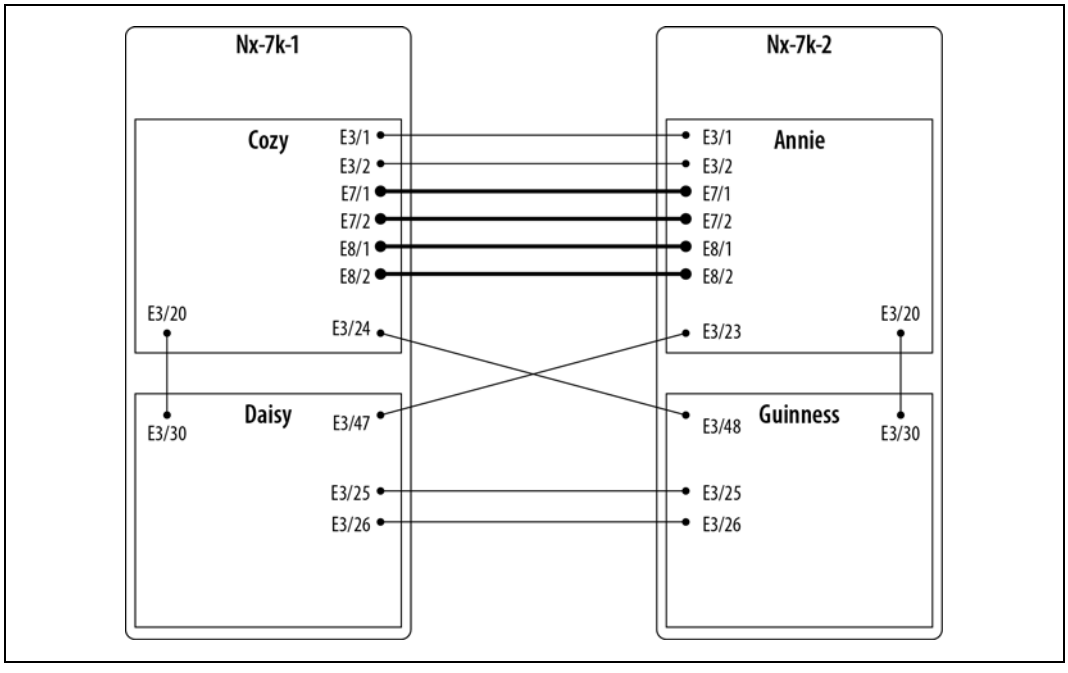
Virtual Device Contexts

Przełączniki Nexus 7000 umożliwiają podział na kilka przełączników logicznych. Jeśli jesteś do mnie podobny, z pewnością ucieszy Cię to, że możesz posiadać stos przełączników wirtualnych w pojedynczej obudowie. Oczywiście założyłem, że urządzenia VDC są podobne do kontekstów firewalli w ASA lub modułach FWSM, gdzie dostępnych jest do 50 kontekstów. Naturalnie, pomyliłem się.

Aktualnie tylko seria Nexus 7000 obsługuje VDC. Co więcej, na razie dostępne są w tych potężnych urządzeniach jedynie cztery konteksty VDC. Pomyślałem, że to nieco bezsensowna liczba kontekstów, spytałem więc ludzi z Cisco, skąd to pozornie głupie ograniczenie. Odpowiedzieli, że chcieli upewnić się, iż każdy kontekst będzie posiadał pełną moc i możliwości przełącznika Nexus. Argumentowali, że przy 50 kontekstach — z których każdy obsługuje dziesiątki tysięcy tras oraz posiada niezliczone możliwości przełącznika Nexus — znacznie ucierpiałaby wydajność w obrębie kontekstu. Moim zdaniem lepiej byłoby zapewnić większą liczbę kontekstów przy jednoczesnym wyłączeniu określonych funkcji — tak jak w przypadku firewalli ASA. Być może takie rozwiązanie pojawi się w przyszłości, ale na razie utknęliśmy na maksymalnie czterech kontekstach.

Do naszego laboratorium zabrałem dwa Nexusy 7000. Na każdym zdefiniowałem dwa konteksty VDC i połączyłem je między sobą na różne sposoby. Przełączniki nazwałem NX-7k-1 oraz NX-7k-2. Dwa konteksty VDC na NX-7k-1 to Cozy i Daisy, a kolejne dwa na NX-7k-2 to Annie i Guinness. Te nazwy to imiona moich nowofunlandów. Nie spędziłeś z tymi psami tylu lat co ja, ale diagram powinien pomóc Ci zrozumieć relacje między nimi. Rysunek 18.12 przedstawia schemat kontekstów VDC i ich wzajemne połączenia.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 18.12. Połączenia między kontekstami VDC na przełącznikach Nexus 7000

Możemy zaczynać. Skonfigurowałem już konteksty VDC na NX-7k-1, skonfigurujmy więc teraz konteksty Annie i Guinness na NX-7k-2.

Najpierw zdefiniuję VDC komendą `vdc`. Muszę podać nazwę i numer VDC. VDC1 jest kontekstem systemowym, należy więc rozpocząć od numeru 2:

```
NX-7K-2(config)# vdc Annie id 2
Note:  Creating VDC, one moment please ... //Tworzenie VDC, proszę czekać...
2010 May 27 00:05:57 NX-7K-2 %$ VDC-1 %$ %VDC_MGR-2-VDC_ONLINE: vdc 2 has come online
                                                          //vdc 2 jest teraz online
```

Czemu nie musimy najpierw włączyć funkcji VDC? Nie jestem co prawda autorem kodu, ale sądzę, że funkcja VDC jest integralną częścią systemu operacyjnego. Nawet podczas uruchamiania przełącznika 5010, który nie obsługuje kontekstów VDC, otrzymasz raport o bootowaniu VDC#1.

W ten prosty sposób mamy już VDC. Jednak czy to router, czy przełącznik, wirtualny, czy fizyczny, jest bezużyteczny bez przypisania interfejsów — zróbmy więc to. Przypiszę teraz pierwsze 24 porty z 48-portowego modułu znajdującego się w slocie 3. Możesz oczywiście przypisać inne porty, ale musisz robić to grupami należącymi do obwodów ASIC. Jak przekonasz się w dalszej części rozdziału, grupy te mogą wydawać się dość dziwne. Na razie postarajmy się nie komplikować:

```
NX-7K-2(config-vdc)# allocate interface Ethernet3/1-24
Moving ports will cause all config associated to them in source vdc to be removed.
                                //Przeniesienie portów spowoduje usunięcie całej ich konfiguracji w źródłowym vdc
Are you sure you want to move the ports (y/n)?  [yes] //Czy nadal chcesz przenieść te porty (tak/nie)? [tak]
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

To ostrzeżenie informuje, że wybrane porty zostaną usunięte z kontekstu systemowego. Jeśli te interfejsy były wcześniej skonfigurowane w kontekście systemowym, przeniesienie ich do innego VDC może mieć nieciekawe konsekwencje. Można by to porównać z wyszarpaniem połowy modułu z wirtualnego urządzenia NX-7k-2 i przeniesieniem go do wirtualnego urządzenia Annie. Nie uszkodzi to przełącznika, ale wszystko, co podłączysz do tych portów, będzie offline.

Następnym krokiem będzie przypisanie wszystkich 32 portów modułu 32 GB:

```
NX-7K-2(config-vdc)# allocate interface Ethernet8/1-32
Moving ports will cause all config associated to them in source vdc to be removed.
Are you sure you want to move the ports (y/n)?  [yes] y
```

Szybkie sprawdzenie upewni nas, że mamy teraz dwa konteksty VDC — oryginalny kontekst systemowy oraz kontekst Annie. Możemy podejrzeć konfigurację VDC poleceniem `show vdc`:

```
NX-7K-2(config-vdc)# sho vdc
vdc_id  vdc_name                        state               mac
------  --------                        -----               ----------
1       NX-7K-2                         active              00:26:98:1b:f0:41
2       Annie                           active              00:26:98:1b:f0:42
```

Stwórzmy teraz kontekst VDC o nazwie Guinness. Nadamy mu identyfikator 3:

```
NX-7K-2(config-vdc)# vdc Guinness id 3
Note:  Creating VDC, one moment please ...
2010 May 27 00:08:41 NX-7K-2 %$ VDC-1 %$ %VDC_MGR-2-VDC_ONLINE: vdc 3 has come online
```

Teraz przypiszę interfejsy. Znajdzie się tutaj druga połowa portów 48-portowego modułu. W tym kontekście VDC nie będzie portów 10 Gb/s:

```
NX-7K-2(config-vdc)#   allocate interface Ethernet3/25-48
Moving ports will cause all config associated to them in source vdc to be removed.
Are you sure you want to move the ports (y/n)?  [yes] y
```

W tym momencie wstępna konfiguracja kontekstów VDC jest zakończona. Zdefiniowane konteksty pokaże nam polecenie `show vdc`:

```
NX-7K-2(config-vdc)# sho vdc
vdc_id  vdc_name                        state               mac
------  --------                        -----               ----------
1       NX-7K-2                         active              00:26:98:1b:f0:41
2       Annie                           active              00:26:98:1b:f0:42
3       Guinness                        active              00:26:98:1b:f0:43
```

Aby zobaczyć szczegółowy raport alokacji portów pomiędzy kontekstami VDC, mogę zastosować polecenie `show vdc membership`. Wolałbym jednak użyć komendy `show run vdc`, która pozwoli sprawdzić, jakie polecenia zostały zastosowane do każdego VDC. Wszystkie komendy `limit-resource` mają wartości domyślne. Ich zastosowanie nie mieści się w ramach tej i tak już obszernej książki:

```
NX-7K-2(config)# sho run vdc

!Command: show running-config vdc
!Time: Thu May 27 00:54:11 2010

version 4.2(4)
vdc NX-7K-2 id 1
  limit-resource vlan minimum 16 maximum 4094
  limit-resource monitor-session minimum 0 maximum 2
  limit-resource vrf minimum 16 maximum 200
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  limit-resource port-channel minimum 0 maximum 768
  limit-resource u4route-mem minimum 32 maximum 32
  limit-resource u6route-mem minimum 16 maximum 16
  limit-resource m4route-mem minimum 58 maximum 58
  limit-resource m6route-mem minimum 8 maximum 8
vdc Annie id 2
  allocate interface Ethernet3/1-24
  allocate interface Ethernet8/1-32
  boot-order 1
  limit-resource vlan minimum 16 maximum 4094
  limit-resource monitor-session minimum 0 maximum 2
  limit-resource vrf minimum 16 maximum 200
  limit-resource port-channel minimum 0 maximum 768
  limit-resource u4route-mem minimum 8 maximum 8
  limit-resource u6route-mem minimum 4 maximum 4
  limit-resource m4route-mem minimum 8 maximum 8
  limit-resource m6route-mem minimum 2 maximum 2
vdc Guinness id 3
  allocate interface Ethernet3/25-48
  boot-order 1
  limit-resource vlan minimum 16 maximum 4094
  limit-resource monitor-session minimum 0 maximum 2
  limit-resource vrf minimum 16 maximum 200
  limit-resource port-channel minimum 0 maximum 768
  limit-resource u4route-mem minimum 8 maximum 8
  limit-resource u6route-mem minimum 4 maximum 4
  limit-resource m4route-mem minimum 8 maximum 8
  limit-resource m6route-mem minimum 2 maximum 2
```

Zwróć uwagę na komendę `boot-order`. Dzięki niej mogę zdefiniować kolejność uruchamiania kontekstów VDC podczas inicjalizacji systemu.

Teraz, kiedy konteksty VDC są już utworzone, możemy ich w końcu użyć. Aby uzyskać dostęp do określonego VDC z poziomu kontekstu systemowego, zastosuj polecenie `switchto vdc` *`vdc_nazwa`*:

```
NX-7K-2(config-vdc)# switchto vdc Annie
         ---- System Admin Account Setup ----
Do you want to enforce secure password standard (yes/no) [y]:

  Enter the password for "admin":
  Confirm the password for "admin":

         ---- Basic System Configuration Dialog VDC: 2 ----
This setup utility will guide you through the basic configuration of
the system. Setup configures only enough connectivity for management
of the system.
[fragment usunięty]
```

Jeśli przyzwyczaiłeś się do konfiguracji kontekstów w firewallach ASA, nieustannie złościć Cię będzie to, że Cisco zdecydowało się zastosować w Nexusie komendę `switchto` w miejsce komendy `changeto` w ASA.

Po przełączeniu do wybranego kontekstu VDC mamy wrażenie, jakby przełącznik Nexus został ponownie uruchomiony, co właściwie ma miejsce. Domyślnie nagłówek wiersza poleceń to nazwa hosta dla kontekstu systemowego oraz nazwa VDC, przedzielone myślnikiem. Można to zmienić poleceniem `vdc combined-hostname`:

```
NX-7K-2-Annie#
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli chciałbyś teraz przełączyć się ponownie do kontekstu systemowego, wprowadź komendę `switchback`:

```
NX-7K-2-Annie# switchback
```

Skonfigurowałem już kilka interfejsów w kontekstach na obu przełącznikach. Komenda `show cdp neighbors` uruchomiona z poziomu VDC Annie pokazuje, że „widzi" on inne konteksty jako urządzenia, ale tylko dlatego, że łączą je fizyczne kable Ethernet. Nie istnieje komunikacja między kontekstami VDC w obrębie przełącznika. Konteksty „nie wiedzą", że są kontekstami ani że inne konteksty są urządzeniami wirtualnymi:

```
NX-7K-2-Annie(config-if-range)# sho cdp neighbors
Capability Codes: R - Router, T - Trans-Bridge, B - Source-Route-Bridge
                  S - Switch, H - Host, I - IGMP, r - Repeater,
                  V - VoIP-Phone, D - Remotely-Managed-Device,
                  s - Supports-STP-Dispute

Device-ID             Local Intrfce Hldtme Capability  Platform      Port ID
NX-7K-1-Cozy(JAF1408BERL)
                       Eth3/1        140    R S I s     N7K-C7010     Eth3/1
NX-7K-1-Cozy(JAF1408BERL)
                       Eth3/2        140    R S I s     N7K-C7010     Eth3/2
NX-7K-1-Daisy(JAF1408BERL)
                       Eth3/23       172    R S I s     N7K-C7010     Eth3/47
NX-7K-1-Cozy(JAF1408BERL)
                       Eth7/1        179    R S I s     N7K-C7010     Eth7/1
NX-7K-1-Cozy(JAF1408BERL)
                       Eth7/2        179    R S I s     N7K-C7010     Eth7/2
NX-7K-1-Cozy(JAF1408BERL)
                       Eth8/2        123    R S I s     N7K-C7010     Eth8/2
```

Tryb przepustowości współdzielonej oraz dedykowanej

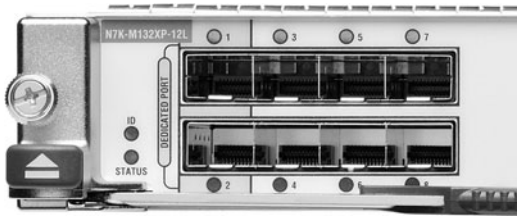
Niektóre moduły 10 Gb dla Nexusa obsługują coś, co nazywa się **trybem portu dedykowanego** (ang. *dedicated port mode*). W tym trybie mogę skonfigurować, aby przełącznik dopuszczał maksymalną przepustowość na wybranym porcie kosztem pozostałych trzech portów. Takie funkcjonowanie związane jest z budową modułów. Porty Ethernet kontrolowane są przez układy scalone, zwane ASIC (ang. *Application-Specific Integrated Circuits*). Układy ASIC odpowiedzialne są za przekazywanie pakietów dla określonej liczby portów. W 32-portowym module 10 Gb/s na przełączniku Nexus 7000 każdy ASIC kontroluje cztery porty. W praktyce oznacza to, że otrzymujesz tylko 10 Gb/s przepustowości dla każdej czteroportowej grupy w module (nadsubskrypcja 4:1). Być może brzmi to strasznie, jest jednak powszechnym rozwiązaniem dla modułów tego typu.

W swojej praktyce napotkałem interesujące zagadnienia związane z awarią układów ASIC. Mieliśmy raz na przełączniku 6509 bardzo dziwny problem, który zdawał się dotyczyć losowej grupy serwerów. Administrator zorientował się, że po przepięciu serwera na inny port przełącznika problem znikał. Kiedy zacząłem przekopywać się przez fizyczne podłączenia, odkryłem, że serwery, których dotyczył problem, były wszystkie podpięte do przylegających ze sobą sześciu portów przełącznika. Porty te były kontrolowane przez jeden układ ASIC, który uległ awarii. Przenieśliśmy więc serwery na inne porty i zorganizowaliśmy wymianę modułu.

Nexus 7000 daje możliwość przeznaczenia wszystkich zasobów układu ASIC na jeden port z danej grupy, co nazywa się **trybem dedykowanej przepustowości** (ang. *dedicated rate-mode*).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W takim przypadku pozostałe porty kontrolowane przez ten ASIC są wyłączane. Porty w pojedynczej grupie nie są numerowane kolejno. Innymi słowy, jeden ASIC kontroluje porty e1/1, 3, 5 oraz 7, podczas gdy następny ASIC porty e1/2, 4, 6 oraz 8. Może Ci się to wydawać dziwne, dopóki nie spojrzysz na frontowy panel modułu. Jak widać na rysunku 18.13, schemat przydziału portów do grupy odpowiada fizycznemu układowi portów w module. Porty 1, 3, 5 i 7 są fizycznie przylegające i dzielą to samo łącze sprzętowe (więc i ten sam ASIC) w obrębie modułu. Na przedstawionym module są nawet zaznaczone porty, które mogą być skonfigurowane jako porty dedykowane.



Rysunek 18.13. 32-portowy moduł 10 Gb/s Nexus z oznaczonymi portami dedykowanymi. Zdjęcie zamieszczono dzięki uprzejmości Cisco Systems, Inc. Nieuprawnione użycie zabronione

Dlaczego więc mielibyśmy dedykować całą przepustowość do jednego portu? Większość urządzeń z reguły nie wykorzystuje pełnych możliwości interfejsu, co jest powodem współdzielenia jednego układu ASIC przez kilka portów. Są jednak urządzenia wymagające wyższych przepustowości. Jednym z przykładów są serwery kopii zapasowych, do których skierowany jest ruch sieciowy o dużym natężeniu, pochodzący z wielu źródeł. Kolejnym takim urządzeniem będzie serwer baz danych. Jeśli Twoja sieć jest rozległa, możesz być zmuszony do skonfigurowania dedykowanych portów uplink, tak aby jak najlepiej wykorzystać dostępną prędkość łączy.

Aby skonfigurować port w trybie dedykowanym, zastosuj polecenie interfejsu `rate-mode`. Opcje tego polecenia to `dedicated` (dedykowany) oraz `shared` (współdzielony):

```
NX-7K-1-Cozy(config-if)# rate-mode ?
  dedicated  Rate mode is dedicated
  shared     Rate mode is shared
```

Kiedy ustawiasz port w trybie dedykowanym, wszystkie porty z danej grupy muszą posiadać status *admin-down*. W innym przypadku otrzymasz komunikat błędu:

```
NX-7K-1-Cozy(config-if)# rate-mode dedicated
ERROR: Ethernet7/1: all ports in the port-group must be in admin shut state
                                        //Wszystkie porty w tej grupie muszą mieć status „admin shut"
```

W poniższej sekwencji poleceń wyłączyłem (ang. *shut down*) wszystkie porty w pierwszej grupie na module 7, a następnie przypisałem zasoby ASIC do pierwszego portu:

```
NX-7K-1-Cozy(config-if)# int e7/1-4
NX-7K-1-Cozy(config-if-range)# shut
NX-7K-1-Cozy(config-if-range)# int e7/1
NX-7K-1-Cozy(config-if)# rate-mode dedicated
```

W tym momencie przełącznik nie pozwoli mi już na powrót włączyć pozostałych portów z grupy:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-7K-1-Cozy(config-if)# int e7/3
NX-7K-1-Cozy(config-if)# no shut
ERROR: Ethernet7/3: Config not allowed, as first port in the port-grp is dedicated
                                    //Konfiguracja niedozwolona, ponieważ pierwszy port z grupy jest portem dedykowanym
```

Na listingu z polecenia `show int brief` port dedykowany będzie oznaczony flagą (D):

```
NX-7K-1-Cozy(config-if-range)# sho int brie  | beg Eth7
--------------------------------------------------------------------------------
Ethernet      VLAN   Type Mode   Status  Reason                   Speed     Port
Interface                                                                    Ch #
--------------------------------------------------------------------------------
Eth7/1        --     eth  routed up      none                       10G(D) --
Eth7/2        --     eth  routed up      none                       10G(S) --
Eth7/3        --     eth  routed down    SFP not inserted          auto(S) --
Eth7/4        --     eth  routed down    SFP not inserted          auto(S) -
[fragment usunięty]
```

Interfejs będzie teraz zgłaszał się w trybie przepustowości dedykowanej:

```
X-7K-1-Cozy# sho int e7/1
Ethernet7/1 is up
  Hardware: 10000 Ethernet, address: 0026.9807.95c2 (bia 5475.d09a.81dc)
  MTU 1500 bytes, BW 10000000 Kbit, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA
  full-duplex, 10 Gb/s, media type is 10g
  Beacon is turned off
  Auto-Negotiation is turned off
  Input flow-control is off, output flow-control is off
  Rate mode is dedicated
  Switchport monitor is off
  Last link flapped 00:03:48
[fragment usunięty]
```

Jak możesz się spodziewać po przyjrzeniu się frontowemu panelowi modułu, tylko pierwszy port z grupy może być skonfigurowany w trybie dedykowanej przepustowości:

```
NX-7K-1-Cozy(config-if)# int e7/1, e7/3, e7/5, e7/7
NX-7K-1-Cozy(config-if-range)# rate-mode dedicated
ERROR: Config not applied on all ports. //BŁĄD: Konfiguracja nie została zastosowana do wszystkich portów
       Config allowed only on the first port in each port group.
                                    //Konfiguracja dozwolona tylko na pierwszym porcie w grupie
```

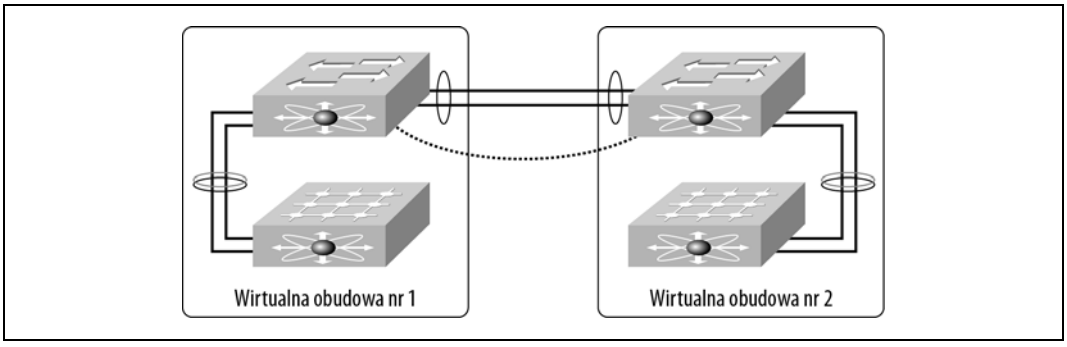
W przełączniku Nexus 5548, w którym wszystkie porty osiągają prędkość połączenia kablowego, każdy z portów jest w trybie dedykowanym i nie można tego zmienić. Poniżej znajduje się listing z polecenia `show int brief` wykonanego na tym przełączniku. Każdy port oznaczony jest flagą (D):

```
NX-5548# sho int brie
--------------------------------------------------------------------------------
Ethernet      VLAN   Type Mode   Status  Reason                   Speed     Port
Interface                                                                    Ch #
--------------------------------------------------------------------------------
Eth1/1        1      eth  access down    SFP validation failed       10G(D) 10
Eth1/2        1      eth  access down    SFP not inserted            10G(D) 10
Eth1/3        1      eth  access down    SFP not inserted            10G(D) 10
Eth1/4        1      eth  access down    SFP not inserted            10G(D) 10
Eth1/5        1      eth  access down    SFP not inserted            10G(D) 10
Eth1/6        1      eth  access down    SFP not inserted            10G(D) 10
Eth1/7        1      eth  access down    SFP not inserted            10G(D) 10
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

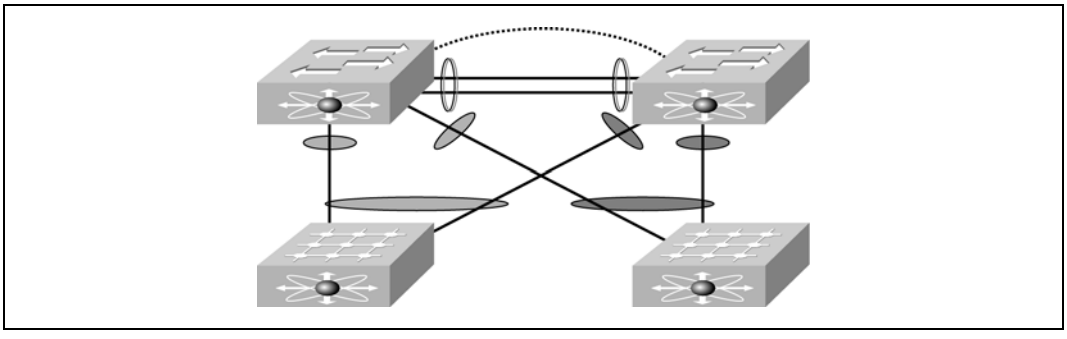
Konfiguracja urządzeń FEX

Istnieją dwa sposoby podłączania urządzeń FEX do głównych przełączników. Pierwsza metoda, zwana trybem **passthru**, przedstawiona jest na rysunku 18.14. W tej topologii każdy FEX powinien być traktowany jak moduł przełącznika, do którego jest podłączony.



Rysunek 18.14. FEX-y podłączone w trybie passthru

Drugi typ topologii podłączania urządzeń FEX, zwany trybem **crossover**, został pokazany na rysunku 18.15. Schemat ten jest podobny do schematu tradycyjnej sieci wieloprzełącznikowej, pamiętaj jednak, że FEX-y nie są przełącznikami, tylko rozszerzeniami strukturalnymi. Taka topologia ma swoje zalety, ponieważ każdy FEX jest teraz podłączony do obydwu głównych przełączników. Jednak moim zdaniem ryzyko takiej konfiguracji przewyższa ewentualne korzyści. Jedynym przypadkiem, w którym zalecam konfigurację w trybie *crossover*, jest sytuacja, gdy serwery podłączone do FEX-ów nie należą do dwóch sieci.



Rysunek 18.15. FEX-y podłączone w trybie crossover

Uważaj na starsze modele urządzeń FEX. Nexus FEX 2148 nie umożliwia tworzenia kanałów EtherChannel w obrębie pojedynczego FEX-u.

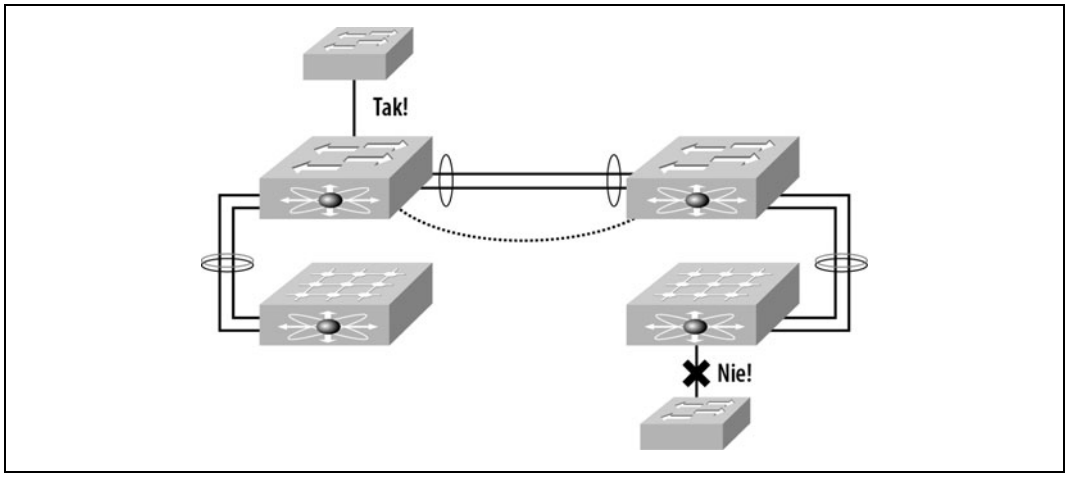
Wadą schematu *crossover* jest to, że każdy z głównych przełączników będzie traktował dany FEX jako rozszerzenia swojej własnej struktury (właściwie tak jest). Bardzo łatwo jest się pomylić i skonfigurować ten sam fizyczny FEX odmiennie na każdym głównym przełączniku. Może to mieć katastrofalne skutki, nie wspominając o trudnościach ze zdiagnozowaniem przyczyny

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

problemu. Jeśli taka topologia jest uzasadniona dla jakiejś sieci, zdecydowanie zalecam skorzystanie z funkcji konfiguracji synchronicznej (ang. *config-sync*) w Nexusie, którą opisałem w dalszej części rozdziału. Nie zapominaj jednak, że konfiguracja synchroniczna jest dostępna tylko dla NX-OS w wersji 5.0(2) i późniejszej.

Musisz pamiętać, że FEX-y nie przełączają same. Za przełączanie odpowiedzialne są główne przełączniki — w naszym przykładzie para przełączników Nexus 5000. Nawet jeśli skomunikować muszą się dwa hosty podłączone do tego samego FEX-u, ramki przejdą przez uplink, zostaną przełączone na switchu 5000 i z powrotem przesłane uplinkiem do FEX-u. Zamiast myśleć o urządzeniach FEX jak o przełącznikach, lepiej potraktować je jako moduły przełącznika typu *chassis*. Kiedy już uznasz, że przełącznik 5000 działa jak moduł zarządzający, a FEX jak zwykły moduł, topologia *passthru* nabierze więcej sensu. Nie podłączyłbyś przecież w topologii *crossover* modułu przełącznika *chassis* do modułu zarządzającego drugiego przełącznika z pary, prawda?

Do urządzenia FEX nie możesz podłączyć kolejnego przełącznika. Na wszystkich portach FEX-u skonfigurowana jest domyślnie funkcja BPDU Guard, której nie można wyłączyć. Możesz jednak podłączyć kolejny przełącznik do Nexusa 5000, tak jak pokazuje to rysunek 18.16.



Rysunek 18.16. Podłączanie kolejnych przełączników w sieci z FEX-ami

OK, możesz podłączyć przełącznik do urządzenia FEX. Musisz jednak uniemożliwić przełącznikowi wysyłanie ramek BPDU. Oszczędzę Ci przemowy z cyklu „to bardzo kiepski pomysł".

Konfigurowanie urządzeń FEX jest bardzo proste, zakładając, że rozumiesz, w jaki sposób one działają. Na początek musisz włączyć funkcję FEX:

```
NX-5K-1(config)# feature fex
```

Opisy dla urządzeń FEX

W przypadku Nexus 5000 po włączeniu funkcji FEX należy zdefiniować opis dla każdego urządzenia FEX. Do wyboru masz liczby z przedziału 100 – 199. Przypisana liczba ma znaczenie tylko lokalnie i nie jest dostępna dla innych urządzeń poza przełącznikiem, na którym została zdefi-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

niowana. Jeśli podłączasz jeden FEX do dwóch przełączników 5000, możesz je nazwać inaczej w każdym z nich. Uważaj jednak, gdyż może to być mylące:

```
NX-5K-1(config)# fex ?
  <100-199>  FEX number
```

Oto fragment listingu ze skonfigurowanego urządzenia FEX:

```
fex 100
  description FEX0100
  type "Nexus 2148T"
```

W tym przypadku opis został automatycznie dodany przez przełącznik. Można go jednak zmienić. Typ (Nexus 2148T) również został dodany automatycznie, dzięki funkcji wykrywania urządzeń FEX na przełączniku. Typu FEX nie można zmienić.

Fizyczne podłączenie

Podłączając port lub porty przełącznika 5K do portów *fabric* urządzenia FEX, musisz skonfigurować porty przełącznika również jako porty *fabric*. Do tego celu służy polecenie `switchport mode fex-fabric`. Następnie porty te należy przypisać do odpowiedniego urządzenia FEX komendą `fex associate` *`fex_nr`*. Jeśli wykorzystywany jest więcej niż jeden uplink, porty mogą być zgrupowane w kanał EtherChannel:

```
interface Ethernet1/1
  switchport mode fex-fabric
  fex associate 100
  channel-group 10
```

Komenda `fex associate` kojarzy urządzenie FEX podłączone na danym porcie z jego konfiguracją. W tym przypadku będzie to FEX 100.

Kiedy FEX zostanie rozpoznany po raz pierwszy, prawdopodobnie będzie musiał pobrać odpowiedni mikrokod z przełącznika 5000 lub 7000, do którego jest podłączony. Proces ten może zająć 10 – 15 minut i wiąże się z koniecznością zrestartowania urządzenia FEX. Może nie być żadnego wskaźnika wykonywania tego procesu, zrób więc sobie przerwę na obiad lub próbuj obsesyjnie wpisywać raz po raz komendę `show fex`, aby sprawdzić, co się dzieje. Możesz właściwie zrobić obie te rzeczy, postaraj się jednak nie pozostawić okruchów jedzenia w centrum danych.

Przypinanie

Standardowo dla 48-portowego przełącznika 1 Gb/s, przy założeniu, że mamy uplink EtherChannel 4×10 Gb/s dający w sumie 40 Gb/s, możliwy współczynnik nadsubskrypcji dla tego uplinku to 4,8:1. Przypinanie (ang. *pinning*) umożliwia przypisanie banków portów do indywidualnych uplinków, co zastępuje współdzielenie jednego EtherChannel przez wszystkie porty. Czemu mielibyśmy to robić? Czytaj dalej.

Dlaczego współczynnik wynosi 4,8:1, a nie 1,2:1? W końcu mamy uplink 40 Gb/s oraz 48 portów 1 Gb/s. Rozbieżność wynika z natury rozkładania obciążenia dla EtherChannel. Pamiętaj, że EtherChannel nie jest właściwie szybszym łączem, ale raczej kilkoma łączami o tej samej prędkości obciążanymi zgodnie z predefiniowanym algorytmem. Więcej informacji na temat haszowania EtherChannel znajdziesz w rozdziale 7.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przy czterech uplinkach 10 Gb/s możemy przypisać po 12 portów do każdego uplinku, co daje nadsubskrypcję 1,2:1. Jest to potencjalne zwiększenie prędkości uplinku w stosunku cztery do jednego. Wydaje się, że to dobry pomysł. Minusem tej metody jest wyłączenie wszystkich 12 portów w przypadku awarii jednego uplinku. Nie możemy liczyć na dynamiczne usuwanie problemu, jak ma to miejsce dla EtherChannel przy przypinaniu portów *fabric*.

Przypinanie może być konfigurowane z jednym, dwoma lub czterema linkami. Rozkład dla 48-portowego urządzenia FEX przedstawia się następująco:

Przypinanie pojedyncze (ang. *pinning one*)
: Wszystkie porty hosta wykorzystują jeden uplink. W tej metodzie stosuje się uplink EtherChannel.

Przypinanie podwójne (ang. *pinning two*)
: Pierwsze 24 porty wykorzystują jeden uplink, a kolejne 24 porty drugi uplink.

Przypinanie poczwórne (ang. *pinning four*)
: Każda 12-portowa grupa portów posiada własny uplink.

Kiedy mamy FEX, w którym wszystkie cztery uplinki są zgrupowane w jeden kanał, musimy zastosować przypinanie pojedyncze, definiowane w bloku konfiguracji FEX:

```
fex 101
  pinning max-links 1
  description FEX0101
  type "Nexus 2148T"
```

Jeśli chcemy zrezygnować z koncepcji stosowania EtherChannel i zamiast tego lepiej wykorzystać szerokość pasma, powinniśmy zastosować przypinanie poczwórne:

```
fex 101
  pinning max-links 4
  description FEX0101
  type "Nexus 2148T"
```

W przypadku próby konfiguracji uplinków do postaci kanału dostaniesz informację o błędzie, jeśli przypinanie będzie ustawione na poziomie wyższym niż pojedyncze:

```
NX-5K-1(config-if-range)# channel-group 10
command failed: port not compatible [FEX pinning max-links not one]
```
//Polecenie nie mogło być wykonane: port niekompatybilny [przypinanie FEX nie jest pojedyncze]

Uplinki fabric dla FEX z wykorzystaniem vPC

Aby skonfigurować pojedynczy FEX zmostkowany z dwoma przełącznikami Nexus 5000 (schemat podłączenia *cross-connect*), musisz zastosować vPC do połączenia tych linków w jedno łącze logiczne:

NX-5K-1:

```
interface port-channel 20
  switchport mode fex-fabric
  vpc 20
  fex associate 101
  speed 10000
```

NX-5K-2:

```
interface port-channel 20
  switchport mode fex-fabric
```

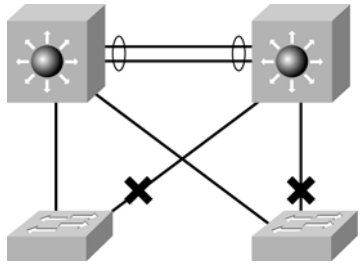

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
vpc 20
fex associate 101
speed 10000
```

Jeśli nie jest to do końca jasne, przeczytaj następny podrozdział, w którym opisałem kanały vPC. Pamiętaj, że konfiguracja FEX powinna być identyczna na obu przełącznikach 5000. W przeciwnym razie mogą pojawić się nieprzewidziane rezultaty.

Virtual Port Channel

Częstym wymaganiem dla urządzeń sieciowych jest możliwość jednoczesnego podłączenia do dwóch przełączników (ang. *dual-home*). Przełączniki podłączone do kolejnych dwóch przełączników tworzą pętlę w warstwie drugiej, co może mieć katastrofalne skutki dla sieci. W takich przypadkach rozwiązaniem jest przerwanie pętli. Standardowo do odnajdywania pętli i blokowania niewłaściwych połączeń stosuje się Spanning Tree. Zapobiega to takim problemom jak burza rozgłoszeniowa. Protokół Spanning Tree (STP) sprawdza się dobrze, jest jednak dość powolny i uniemożliwia wykorzystanie połowy z dostępnej przepustowości dla przełącznika w konfiguracji *dual-home*. Rysunek 18.17 przedstawia tradycyjny schemat konfiguracji *dual-home* z wykorzystaniem STP. Zauważ, że niedostępne są dwa linki, z których każdy prowadzi do jednego z przełączników znajdujących się na dole rysunku. Protokół Spanning Tree zablokował te linki.



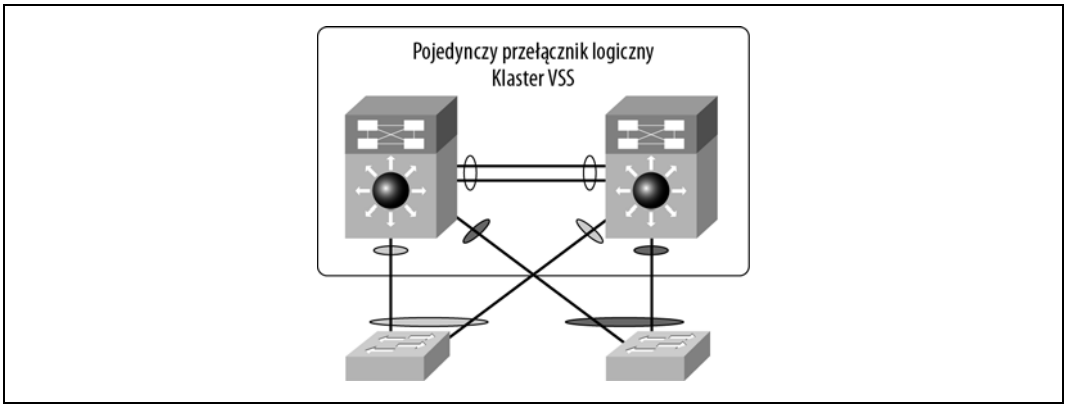
Rysunek 18.17. Tradycyjny schemat Spanning Tree

W środowisku przełączników 6500, posiadających moduły zarządzające VSS, możemy stworzyć jeden logiczny przełącznik z dwóch przełączników fizycznych. Powstały w ten sposób przełącznik logiczny zwany jest **klastrem VSS** (ang. *VSS cluster*) — więcej informacji na ten temat znajdziesz w rozdziale 17. Ponieważ obydwa linki z dolnych przełączników prowadzą teraz do jednego logicznego przełącznika, protokół Spaning Tree nie jest potrzebny. Schemat klastra VSS przedstawiony jest na rysunku 18.18. Dolne przełączniki postrzegają klaster VSS jako jeden punkt podłączenia, więc mogą ustanowić pojedynczy port channel do tego urządzenia. Cisco opisuje tę sytuację jako dwa przełączniki z jedną magistralą sterowania.

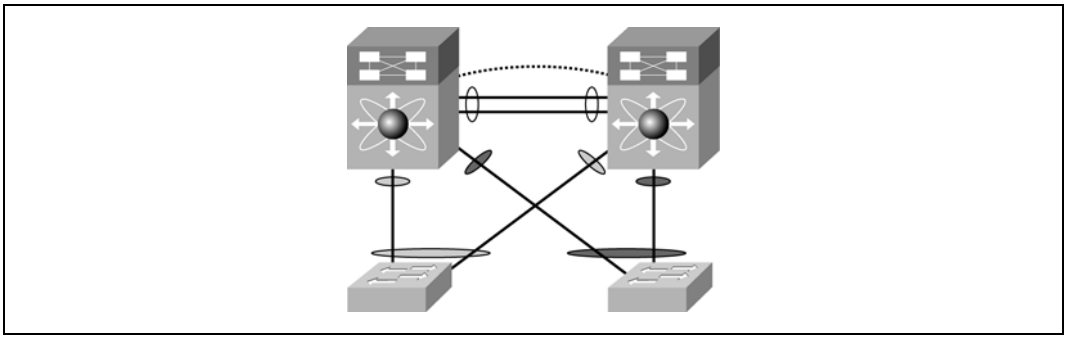
Jest to znacznie „czystsze" rozwiązanie, ponieważ nie wymaga zastosowania STP. Każde łącze może być teraz w pełni wykorzystane, co pozwala podwoić prędkość uplinków z dolnych przełączników.

Nexus nie obsługuje technologii VSS, korzystając w zamian z vPC. W przypadku vPC górne przełączniki nie tworzą logicznej całości, ale przełączniki na dole nadal postrzegają je jako pojedynczy punkt podłączenia podczas tworzenia kanałów port channel. Rysunek 18.19 przedstawia

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 18.18. Schemat 6500 VSS



Rysunek 18.19. Schemat vPC

dwa przełączniki podłączone do dwóch Nexusów 7000, które stosują vPC dla wszystkich połączeń. Cisco opisuje taką sytuację jako dwa przełączniki z jedną magistralą danych i dwoma magistralami sterowania.

Przebrnijmy przez etapy konfiguracji vPC.

Najpierw musisz włączyć funkcję vPC:

```
NX-5k-1(config)# feature vpc
```

Przed przystąpieniem do dalszej konfiguracji musisz zdefiniować domenę vPC. Numer domeny powinien być taki sam dla wszystkich przełączników, na których skonfigurowane zostaną linki vPC:

```
NX-5K-1(config-if)# vpc domain 1
```

Chociaż nie jest to wymagane, skonfigurowałem jeden z przełączników z pary, nadając mu priorytet nadrzędności (ang. *master*) dzięki poleceniu `vpc role priority`. Priorytet nadrzędności powinien mieć przełącznik, który byłby mostkiem głównym dla Spanning Tree. O nadrzędności decyduje niższa wartość priorytetu. Domyślna wartość to 32768:

```
NX-5K-1(config-if)# role priority 1000
```

Kanał vPC wymaga pomiędzy parą przełączników linku, który nazywa się **peer keepalive link**. Jest to dodatkowy link w stosunku do port channel między tymi przełącznikami i zawiera on trunki dla VLAN-ów. *Peer keepalive link* musi być łączem w warstwie trzeciej, ponieważ po

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

obu stronach wykorzystywany jest adres IP. Link ten może wykorzystywać istniejącą sieć IP. Jak sugeruje Cisco, powinien on być domyślnie ustanowiony na interfejsach zarządzania. Nie jestem zwolennikiem umieszczania tego, co można nazwać połączeniem krytycznym (przynajmniej z punktu widzenia obserwatora), w sieci zarządzającej, jednak ten przykład tak to przedstawia. VRF zarządzający jest zastosowany jako domyślny w Nexusie 5000:

```
NX-5K-1(config-vpc)# peer-keepalive destination 10.10.10.2
Note:
 --------:: Management VRF will be used as the default VRF ::--------
                                                    //VRF zarządzający zostanie zastosowany jako VRF domyślny
```

W Nexusie 7000 link ten może być ustanowiony na dowolnym interfejsie, ponieważ wszystkie interfejsy są domyślnie ustawione w trybie routingu. Poniżej skonfigurowałem interfejs e3/1 jako *peer keepalive link*:

```
NX-7K-1(config)# int e3/1
NX-7K-1(config-if)# ip address 10.255.255.1/24
NX-7K-1 (config)# vpc domain 1
NX-7K-1 (config-vpc-domain)# peer-keepalive destination 10.255.255.2
```

Teraz skonfigurujmy drugą stronę linku na 7K-2:

```
NX-7K-2(config)# int e3/1
NX-7K-2(config-if)# ip address 10.255.255.2/24
```

I w końcu vPC *peer keepalive link*:

```
NX-7K-2 (config)# vpc domain 1
NX-7K-2 (config-vpc-domain)# peer-keepalive destination 10.255.255.1
```

Kanał vPC wymaga, aby ten link między parą przełączników — najlepiej trunk — został zdefiniowany jako **peer link**. Będzie on zastosowany do przenoszenia ruchu, co jest niezbędne do wykorzystania vPC. Ogólnie rzecz biorąc, link ten jest głównym łączem pomiędzy tymi dwoma przełącznikami. Według dokumentacji vPC *peer link* powinien być kanałem EtherChannel składającym się z wielu linków 10 Gb/s. W warunkach laboratoryjnych możesz jednak wykorzystać do testów pojedynczy link. Bez względu na liczbę linków interfejs musi być skonfigurowany jako port channel. To polecenie jest identyczne dla przełączników 5000 i 7000. Poniżej port channel 1 został wykorzystany jako *peer link*:

```
NX-5K-1(config-if)# int po 1
NX-5K-1(config-if)# vpc peer-link
Please note that spanning tree port type is changed to "network" port type on vPC
peer-link. This will enable spanning tree Bridge Assurance on vPC peer-link provided
the STP Bridge Assurance (which is enabled by default) is not disabled. //Zauważ, że typ portu Spanning
// Tree został zmieniony na „network" dla vPC peer-link. Uruchomi to funkcję Bridge Assurance Spanning Tree dla vPC
// peer-link, pod warunkiem że STP Bridge Assurance (które jest domyślnie włączone) nie zostało wyłączone.
```

Kiedy zostanie to skonfigurowane, musisz oznakować na każdym przełączniku 5K kanały EtherChannel, które mają być połączone pojedynczym vPC. Dobrą praktyką jest znakowanie EtherChannel tym samym numerem, jaki posiada port channel:

```
interface port-channel 10
  vpc 10
```

Przydzielaj vPC zawsze ten sam numer, jaki ma port channel, do którego vPC został przypisany. Debugowanie vPC przy niezgodnej numeracji jest łamigłówką. Wolałbym jej nie rozwiązywać z dyrektorem zarządzającym na karku, ponaglającym mnie, abym przywrócił przełącznik do działania. Staraj się możliwie upraszczać konfigurację.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W każdym przypadku, gdy vPC nie będzie działać, będziesz się zastanawiał, co się u licha dzieje. Na szczęście Cisco raczyło zaimplementować polecenie lub dwa, które mogą pomóc rozwiązać problem.

Dla vPC konfiguracja po obu stronach linku musi być identyczna podobnie jak w przypadku interfejsów składowych dla EtherChannel. Różnica polega na tym, że w przypadku vPC konfiguracja dwóch osobnych przełączników musi pasować, a wymagania te dotyczą czegoś więcej niż tylko interfejsów.

Istnieją dwa rodzaje parametrów kompatybilności dla vPC: typ 1 oraz typ 2. Typ 1 to parametry, które muszą pasować, aby vPC został utworzony (lub jeśli został utworzony, to żeby nie został zerwany). Typ 2 to parametry, które powinny pasować, jednak vPC nie zostanie zerwany, jeśli nie będą.

Pierwszym etapem postępowania w przypadku niedziałającego vPC jest zastosowanie komendy `show vpc`. Poniżej zamieściłem przykład prawidłowo funkcjonującego vPC. Jeśli posiadałbyś 30 kanałów vPC, każdy z nich znalazłby się na listingu tego polecenia. Zauważ, że pokazany jest również status *peer link*.

Listing dla tej komendy może się różnić w zależności od modelu przełącznika, wersji systemu operacyjnego oraz uruchomionych funkcji.

```
NX-7K-2-Annie(config-if)# sho vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id                     : 100
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : Suspended (Destination IP not reachable)
Configuration consistency status: success
Type-2 consistency status         : success
vPC role                          : secondary
Number of vPCs configured         : 1
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ -------------------------------------------------
1    Po1    up     1
vPC status
----------------------------------------------------------------------
id   Port   Status Consistency Reason                     Active vlans
--   ----   ------ ----------- -------------------------- ------------
10   Po10   up     success     success                    1
```

Kolejną komendą, którą powinieneś polubić podczas diagnozowania problemów z vPC, jest `show vpc consistency-parameters`. Kiedy określisz, z którym kanałem vPC musisz się zmierzyć, otrzymasz listę szczegółowych informacji dotyczących ustawień do dopasowania. Oto listing z tego polecenia dla działającego vPC:

```
NX-7K-2-Annie(config-if)# sho vpc consistency-parameters vpc 10
    Legend:
        Type 1 : vPC will be suspended in case of mismatch
Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
STP Port Type               1     Default                Default
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
STP Port Guard              1     None                  None
STP MST Simulate PVST       1     Default               Default
mode                        1     on                    on
Speed                       1     1000 Mb/s             1000 Mb/s
Duplex                      1     full                  full
Port Mode                   1     access                access
Native Vlan                 1     1                     1
MTU                         1     1500                  1500
Allowed VLANs               -     1                     1
Local suspended VLANs       -     -                     -
```

Teraz zmieńmy coś na jednym z przełączników i zobaczmy, co się stanie:

```
NX-7K-2-Annie(config-if)# int e3/23
NX-7K-2-Annie(config-if)# speed 100
```

Listing z komendy `show vpc` pokaże, że kanał vPC został zerwany:

```
NX-7K-2-Annie(config-if)# sho vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id                     : 100
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : Suspended (Destination IP not reachable)
Configuration consistency status: failed
Configuration consistency reason: vPC type-1 configuration incompatible - STP gl
obal port type inconsistent
Type-2 consistency status         : success
vPC role                          : secondary
Number of vPCs configured         : 1
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ -----------------------------------------------------
1    Po1    up     -
vPC status
----------------------------------------------------------------------
id   Port   Status Consistency Reason                     Active vlans
--   ----   ------ ----------- -------------------------- ------------
10   Po10   down*  failed      Global compat check failed -
```

Szczegółowy raport informujący o obecnych problemach zgodności otrzymasz, wpisując polecenie `show vpc consistency-parameters`:

```
NX-7K-2-Annie(config)# sho vpc consistency-parameters vpc 10
    Legend:
        Type 1 : vPC will be suspended in case of mismatch
Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
STP Port Type               1     Default                Default
STP Port Guard              1     None                   None
STP MST Simulate PVST       1     Default                Default
mode                        1     on                     on
Speed                       1     100 Mb/s               1000 Mb/s
Duplex                      1     half                   full
Port Mode                   1     access                 access
Native Vlan                 1     1                      1
MTU                         1     1500                   1500
Allowed VLANs               -     1                      1
Local suspended VLANs       -     1                      -
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kanały vPC wymagają czegoś więcej niż tylko zgodności interfejsów. Istnieje cała długa lista parametrów, które muszą być zgodne na obu przełącznikach. Aby sprawdzić tę listę, dodaj argument `global` do polecenia `show vpc consistency-parameters`. Jeśli którykolwiek z parametrów typu 1 nie będzie zgodny, vPC nie zadziała.

W swojej praktyce napotkałem problemy dotyczące vPC, które nie wynikały z niezgodności błędnie podanych na listingu polecenia `show vpc`. Jeśli coś takiego ma miejsce, usuń interfejs port channel oraz całą konfigurację grupy kanału dla fizycznego interfejsu. Po ponownym skonfigurowaniu tych opcji vPC prawdopodobnie zadziała. Takie zachowanie może dotyczyć wersji kodu 4.x oraz 5.x dla przełączników 5000.

```
NX-7K-2-Annie(config)# sho vpc consistency-parameters global
    Legend:
        Type 1 : vPC will be suspended in case of mismatch
Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
STP Mode                    1     Rapid-PVST             Rapid-PVST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type, Edge         1     Normal, Disabled,      Normal, Disabled,
BPDUFilter, Edge BPDUGuard        Disabled               Disabled
STP MST Simulate PVST       1     Enabled                Enabled
Interface-vlan admin up     2                            -
Allowed VLANs               -     1,10                   1,10
Local suspended VLANs       -     -                      -
```

Synchronizacja konfiguracji

Wersja 5.0(2)N1(1) NX-OS dla przełączników 5000 wprowadziła nową funkcję, zwaną **synchronizacją konfiguracji** (ang. *config-sync*). Jest ona bardzo użyteczna, ponieważ wiele innych funkcji wymaga przynajmniej częściowej synchronizacji konfiguracji dwóch przełączników Nexus. Dwa z omawianych przeze mnie przykładów dały mi ostro popalić pod tym względem: FEX-y podłączone w schemacie *crossover* (patrz „Konfiguracja urządzeń FEX" na stronie 287) oraz kanały vPC (patrz „Virtual Port Channel" na stronie 291). W szczególności kanały vPC mają silną tendencję do zrywania połączenia za każdym razem, kiedy zachodzi jakakolwiek niespójność konfiguracji dwóch przełączników Nexus, do których vPC są podłączone.

Wyobraź sobie dwa przełączniki Nexus 5010 połączone z serwerami kasetowymi (ang. *blade servers*) poprzez trunki 10 Gb/s. Serwery wymagają redundancji, stworzyłem więc kanał vPC pomiędzy dwoma Nexusami 5K. Serwery widzą EtherChannel i wszystko działa świetnie. Konfiguracja dla trunku między przełącznikami może wyglądać następująco (zastosowana do obu Nexusów):

```
interface port-channel 10
  description BladeServer
  switchport mode trunk
  switchport trunk allowed vlan 10,20,30,40
  vpc 10
```

Wszystko będzie działać dobrze, dopóki pewnego dnia ktoś nie zdecyduje, że należy zmienić typ portów Spanning Tree. Nieważne dlaczego. Ważne, że jeśli nie zrobisz tego w ciągu 10 minut,

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

zostaniesz zwolniony. Znasz te sytuacje — typowe, dobrze zaplanowane zmiany. Komenda `spanning-tree port type edge trunk` zostaje zastosowana na NX-5K-1, co przynosi dość niefortunne rezultaty. Zaraz po dokonaniu zmiany na pierwszym 5K kanał vPC zostaje zerwany:

```
NX-5K-1(config-if)# sho vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id                     : 1
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status: success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 1
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ -------------------------------------------------
1    Po1    up     1
vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
10     Po10        up     failed      vPC type-1 configuration   -
                                      incompatible - STP
                                      interface port type
                                      inconsistent //Brak kompatybilności konfiguracji vPC typu 1 —
                                                   // niezgodny typ portu interfejsu STP
```

Jeśli wygląda to źle, to dlatego, że tak właśnie jest. Konfiguracja vPC pomiędzy dwoma przełącznikami przestała być zgodna, więc Nexus zasadniczo wyłączył kanał. Teoretycznie vPC powinien zostać zerwany, a jeden z linków powinien pozostać aktywny, umożliwiając przepływ pakietów. Nie powinno dochodzić do przestoju w działaniu, ale ludzie od serwerów narzekali za każdym razem, gdy to robiliśmy. Przed wprowadzeniem funkcji synchronizacji konfiguracji oznaczało to zakaz wprowadzania zmian w konfiguracji poza zaplanowanymi okresami przestojów. Jednak w tego typu środowisku nie planuje się okresów przestoju.

Zaraz po wykonaniu komendy na drugim przełączniku vPC jest z powrotem online, ponieważ konfiguracja jest znów zgodna. Bez względu na to, jak szybko potrafiłbyś skonfigurować oba przełączniki — nawet wciskając jednocześnie *Enter* na dwóch komputerach — nigdy nie będzie to wystarczająco szybko. W każdym przypadku vPC musi przejść przez ten irytujący proces.

Funkcja synchronizacji konfiguracji została stworzona w celu łagodzenia skutków występowania tego typu problemów. W skrócie, dwa przełączniki Nexus są ze sobą połączone poprzez interfejs mgmt0. Polecenia konfiguracyjne ładowane są do bufora jednego z przełączników, a następnie po wykonaniu komendy `commit` konfiguracja jest jednocześnie wykonywana na obu przełącznikach. Brzmi nieźle, ma to jednak również pewne wady.

Pierwszą rzeczą, o której należy pamiętać, jest to, że ta funkcja obsługuje obecnie jedynie komendy dotyczące konfiguracji vPC oraz urządzeń FEX. Przykładowo nie możesz zastosować synchronizacji konfiguracji do interfejsów Ethernet.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Drugą istotną rzeczą jest to, że o ile zsynchronizowane komendy będą wyświetlane na listingu polecenia `show run`, to tak naprawdę ich tam nie ma. Aby skorzystać z tej funkcji, należy najpierw wejść w tryb konfiguracyjny synchronizacji, zwany, o dziwo, `configure sync`:

```
NX-5K-1# configure sync
Enter configuration commands, one per line.  End with CNTL/Z.
NX-5K-1(config-sync)#
```

Mógłbyś pomyśleć, że wszystkie polecenia, które wpiszesz w tym trybie, zostaną automatycznie zreplikowane na drugim przełączniku. Byłbyś jednak w błędzie. Po pierwsze, nie zdefiniowaliśmy jeszcze drugiego przełącznika do pary, a po drugie, byłoby to zbyt proste i mało zabawne.

Komendy przełącznika w trybie synchronizacji są zapisywane w **profilu przełącznika** (ang. *switch profile*). Musisz skonfigurować ten sam profil na obu przełącznikach, które mają być synchronizowane. Poniżej stworzyłem profil przełącznika o wdzięcznej nazwie GAD:

```
NX-5K-1(config-sync)# switch-profile GAD
Switch-Profile started, Profile ID is 1
NX-5K-1(config-sync-sp)#
```

Możesz posiadać tylko jeden profil i musi on być zgodny z profilem na drugim przełączniku z synchronizowanej pary. Skonfigurujmy więc profil na drugim przełączniku:

```
NX-5K-2# configure sync
Enter configuration commands, one per line.  End with CNTL/Z.
NX-5K-2(config-sync)# switch-profile GAD
Switch-Profile started, Profile ID is 1
NX-5K-2(config-sync-sp)#
```

Nadal nie poinstruowaliśmy żadnego z przełączników, w jaki sposób odnaleźć partnera. Zróbmy to najpierw na przełączniku NX-5K-1:

```
NX-5K-1(config-sync-sp)# sync-peers destination 10.0.0.2
```

A teraz na przełączniku NX-5K-2:

```
NX-5K-2(config-sync-sp)# sync-peers destination 10.0.0.1
```

Są to adresy IP interfejsów zarządzających w przełącznikach Nexus 5000.

Według mnie w tym momencie zaczyna się robić nieco dziwnie. Aby komendy się synchronizowały, nie mogą znajdować się w działającej konfiguracji. Jeśli chcesz edytować aktualną konfigurację przełącznika, możesz zaimportować działającą konfigurację lub jej część do profilu przełącznika. Zalecam, abyś robił to jedynie dla interfejsów lub FEX-ów, dla których ma być wykonana synchronizacja. Inaczej możesz spędzić wiele godzin, starając się zrozumieć, czemu to nie działa. Istnieje wiele komend, które nie są obsługiwane przez funkcję synchronizacji konfiguracji, i jak dotąd lista ta nie jest dobrze udokumentowana. Staraj się zachować prostotę i trzymaj się kanałów port channel oraz FEX-ów.

Zdecydowałem więc, że chcę zsynchronizować port channel 10, który jest już skonfigurowany. Świetnie — muszę jedynie zaimportować tę konfigurację i będę mógł użyć funkcji synchronizacji. Tutaj napotykamy pewne ograniczenie, gdyż nie mogę zaimportować konfiguracji po tym, jak skonfigurowałem tę parę przełączników do synchronizacji. Dobra wiadomość jest taka, że będę musiał wykonać import tylko jeden raz. Od tego momentu to funkcja synchronizacji będzie zarządzać odpowiednimi interfejsami i FEX-ami, zakładając, że byłem na tyle przewidujący, by zaimportować wszystko, co będzie mi potrzebne w przyszłości. Specjalnie zwracam Ci na to uwagę, gdyż nie jest to dobrze pomyślane, a ja spędziłem wiele dni w laboratorium oraz wykonałem mnóstwo telefonów do pomocy technicznej Cisco, żeby to rozgryźć.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Idźmy więc dalej i zaimportujmy konfigurację dla port channel 10. Wykonam to teraz na przełączniku NX-5K-2, ale i tak muszę to zrobić dla obu przełączników:

```
NX-5K-2(config-sync-sp)# import interface po10
NX-5K-2(config-sync-sp-import)#
```

Wydaje się proste. Za każdym razem przy dokonywaniu zmian w buforze synchronizacji powinniśmy je zweryfikować. Jeśli komendy nie zostaną zweryfikowane poprawnie, synchronizacja się nie powiedzie:

```
NX-5K-1(config-sync-sp-import)# verify
Failed: Verify Failed //Weryfikacja nieudana
```

Nienawidzę, gdy tak się dzieje. Nie jest to również zbyt pomocna informacja. Jak mam się zorientować, w czym tkwi problem? Pierwszym krokiem będzie sprawdzenie, co znajduje się w buforze. Użyj komendy `show switch-profile` *`nazwa-profilu`* `buffer`:

```
NX-5K-1(config-sync-sp-import)# sho switch-profile GAD buffer
switch-profile  : GAD
----------------------------------------------------------
Seq-no  Command
----------------------------------------------------------
1       interface port-channel10
1.1       switchport mode trunk
1.2       vpc 10
1.3       switchport trunk allowed vlan 10, 20, 30, 40
1.4       speed 1000
```

Wygląda całkiem dobrze, ale wciąż nie wiemy, co stanowi problem. Oto listing z użycia tej samej komendy dla przełącznika NX-5K-2:

```
NX-5K-2(config-sync-sp-import)# sho switch-profile GAD buffer
switch-profile  : GAD
----------------------------------------------------------
Seq-no  Command
----------------------------------------------------------
1       interface port-channel10
1.1       switchport mode trunk
1.2       vpc 10
1.3       switchport trunk allowed vlan 10, 20, 30, 40
1.4       speed 1000
```

Oba listingi wyglądają identycznie. Usuńmy dla zabawy jedną z linii konfiguracji. Może komenda `speed` nie jest obsługiwana w trybie synchronizacji (ta komenda akurat jest obsługiwana, ale zabawmy się). Usunę ją poleceniem `buffer-delete`, po którym podam numer linii do usunięcia. Mógłbym również przesunąć linię poleceniem `buffer-move`. W tym przypadku chcę usunąć linię 1.4:

```
NX-5K-1(config-sync-sp-import)# buffer-delete 1.4
NX-5K-1(config-sync-sp-import)# verify
Failed: Verify Failed
```

Szaleństwo — nadal brak weryfikacji. Aby lepiej zorientować się w sytuacji, zastosuj polecenie `show switch-profile` *`nazwa-profilu`* `status`:

```
NX-5K-1(config-sync-sp-import)# sho switch-profile GAD status
switch-profile  : GAD
----------------------------------------------------------
Start-time:   5257 usecs after Mon Jan 17 20:53:21 2011
End-time: 980506 usecs after Mon Jan 17 20:53:21 2011
Profile-Revision: 1
Session-type: Import-Verify
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Session-subtype: -
Peer-triggered: No
Profile-status: Verify Failed
Local information:
----------------
Status: Verify Failed
Error(s):
Peer information: //Informacja o parze przełączników
----------------
IP-address: 10.0.0.2
Sync-status: In sync //Status synchronizacji: zsynchronizowane
Status: Verify Failure //Status: weryfikacja nieudana
Error(s): //Błąd:
Following commands failed mutual-exclusion checks: //Poniższe komendy nie przeszły testu wzajemnego wykluczenia
interface port-channel10
        switchport mode trunk
interface port-channel10
        vpc 10
interface port-channel10
        switchport trunk allowed vlan 10, 20, 30, 40
```

Listing informuje nas, że powiodła się weryfikacja lokalna, ale nie powiodła się weryfikacja zdalna. Reasumując — ta prosta konfiguracja nie może zostać zweryfikowana poprawnie, ponieważ została zaimportowana, kiedy mieliśmy już skonfigurowaną parę przełączników do synchronizacji.

Kiedy korzystasz z funkcji synchronizacji konfiguracji i chcesz zaimportować aktualną konfigurację z działającej konfiguracji, musisz wyłączyć synchronizację pary przełączników. Jeśli tego nie zrobisz, konfiguracja nie zostanie zweryfikowana poprawnie i nie będziesz w stanie jej zatwierdzić.

Usunę komendę `sync-peers`:

```
NX-5K-1(config-sync-sp-import)# no sync-peers dest 10.0.0.2
```

Popatrz, cóż magicznego się teraz stanie:

```
NX-5K-1(config-sync-sp-import)# verify
Verification Successful //Weryfikacja udana
```

Zdecydowałem się na razie nie komentować tej kwestii, ponieważ nie mam nic miłego do powiedzenia, a matka ostrzegała mnie, bym w takich sytuacjach trzymał język za zębami.

Istnieją dwa etapy weryfikacji, które mogą się nie powieść. Wzajemne wykluczenie, czyli **mutual-exclusion** (mutex), sprawdza, czy konfiguracja, którą wprowadziłem w trybie synchronizacji, nie znajduje się w konfiguracji działającej. Sprawdzane są oba przełączniki, chyba że synchronizacja jest wyłączona lub nie została skonfigurowana. W takim przypadku mutex jest wykonywany lokalnie. Kontrole scalenia, czyli **merge checks**, polegają na porównywaniu z profilem drugiego przełącznika pod kątem ewentualnych konfliktów. Jeśli któryś z etapów weryfikacji się nie powiedzie, problem musi zostać usunięty ręcznie.

Kiedy już konfiguracja interfejsu zostanie poprawnie zweryfikowana, możesz ją zatwierdzić do profilu przełącznika poleceniem `commit`:

```
NX-5K-1(config-sync-sp-import)# commit
Verification successful...
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Po wykonaniu tych samych czynności na przełączniku NX-5K-2 mam obydwa przełączniki prawie gotowe do synchronizacji. Pamiętaj, że muszę przywrócić połączenie między tą parą, zanim będę mógł zsynchronizować konfiguracje. Służy do tego polecenie `sync-peers destination`. Muszę to zrobić na obydwu przełącznikach. Najpierw przełącznik 1:

```
NX-5K-1(config-sync)# switch-profile GAD
Switch-Profile started, Profile ID is 1
NX-5K-1(config-sync-sp)# sync-peers destination 10.0.0.2
```

A teraz przełącznik 2:

```
NX-5K-2(config-sync)# switch-profile GAD
Switch-Profile started, Profile ID is 1
NX-5K-2(config-sync-sp)# sync-peers destination 10.0.0.1
```

Mogę w końcu wykonać to, co było moim celem od początku. Pamiętasz problem `spanning-tree port type`? Skonfiguruję to, ale tym razem zastosuję synchronizację konfiguracji. Muszę być w trybie *config-sync* i w profilu przełącznika:

```
NX-5K-2(config-sync)# switch-profile GAD
Switch-Profile started, Profile ID is 1
```

Kiedy już tu jestem, muszę wpisać komendy, które chcę zsynchronizować. Pamiętaj, że komendy muszą dotyczyć interfejsu dla vPC lub FEX-u i zostały wcześniej zaimportowane lub wprowadzone bezpośrednio w trybie synchronizacji. Wpisuję to tylko na jednym przełączniku:

```
NX-5K-1(config-sync-sp)# int po 10
NX-5K-1(config-sync-sp-if)# spanning-tree port type edge trunk
```

Teraz sprawdzę, czy mam pozytywną weryfikację:

```
NX-5K-1(config-sync-sp-if)# verify
Failed: Session Database already locked, Verify/Commit in Progress.
```

Jeśli bawiłeś się tym w laboratorium i włączałeś/wyłączałeś synchronizację pomiędzy przełącznikami tak jak ja, możesz otrzymać tego typu informację. Możesz ją również zobaczyć, jeśli jesteś podłączony do przełącznika przez telnet i próbujesz zweryfikować lub zatwierdzić konfigurację, w czasie gdy robi to również ktoś inny. Odczekaj chwilę, upewnij się, że tylko Ty konfigurujesz przełącznik, a następnie spróbuj ponownie:

```
NX-5K-1(config-sync-sp)# verify
Verification Successful
```

Skoro mamy już weryfikację, możemy zatwierdzić zmiany. Zmiany konfiguracji zostaną zatwierdzone lokalnie oraz na drugim przełączniku. Może to zająć kilka sekund, więc nie przejmuj się, jeśli będzie Ci się wydawać, że przełącznik chwilowo „wisi”:

```
NX-5K-1(config-sync-sp)# commit
Verification successful...
Proceeding to apply configuration. This might take a while depending on amount
of configuration in buffer.
Please avoid other configuration changes during this time.

Commit Successful
```

Wprowadziłem zmiany z poziomu konsoli przełącznika NX-5K-1, sprawdźmy więc interfejs port channel 10 na NX-5K-2 i zobaczmy, czy zmiany odniosły skutek. Mam nadzieję, że drugi przełącznik również będzie posiadał komendę `spanning-tree port type edge trunk`:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-5K-1(config-sync)# sho run int po 10

!Command: show running-config interface port-channel10
!Time: Mon Jan 17 22:57:20 2011

version 5.0(2)N2(1)

interface port-channel10
  description BladeServer
  switchport mode trunk
  vpc 10
  switchport trunk allowed vlan 10,20,30,40
  spanning-tree port type edge trunk
  speed 1000
```

Jasna cholera, zadziałało! Tak właśnie zareagowałem, gdy po raz pierwszy doświadczyłem poprawnego działania funkcji synchronizacji konfiguracji. W końcu można nieco ogłupieć, siedząc przez trzy dni w laboratorium bez okien.

Synchronizacja konfiguracji jest fajna, ale przyjrzyjmy się zagrożeniom. Konfiguracja, którą wprowadzasz, jest widoczna w konfiguracji działającej, ale jak wynika z powyższego przykładu, w rzeczywistości jej tam nie ma. Konfiguracja ta znajduje się w bazie danych synchronizacji (ang. *config-sync database*). Możesz sprawdzić, co jest w tej bazie, korzystając z komendy `show running-config switch-profile`:

```
NX-5K-1# sho running-config switch-profile

switch-profile GAD
  sync-peers destination 10.0.0.2

  interface port-channel10
    spanning-tree port type edge trunk
```

W międzyczasie działająca konfiguracja wygląda normalnie. Jedynie ta linijka odsłania zagrożenie czające się pod spodem:

```
NX-5K-1# sho run | begin switch-profile
switch-profile GAD
  sync-peers destination 10.0.0.2
```

Nie wygląda bardzo niebezpiecznie, prawda? Teraz wyobraź sobie, że właśnie zatrudniłeś nowego młodszego inżyniera. A nawet lepiej, wyobraź sobie, że to Ty **jesteś** nowym młodszym inżynierem. Będąc młodym, upartym sieciowym półbogiem, czujesz, że komenda `switch-profile` najwyraźniej jest bezużyteczna, więc decydujesz się ją usunąć.

Myślisz, że taki scenariusz jest mało prawdopodobny? Mogę sobie przypomnieć nie mniej niż pięć, a może nawet siedem takich przypadków, kiedy to pojawiali się młodsi inżynierowie i postanawiali „zrobić porządki", co powodowało jedynie katastrofalne przerwy w pracy sieci. Nie bądź takim gościem:

```
NX-5K-1# conf sync
Enter configuration commands, one per line.  End with CNTL/Z.
NX-5K-1(config-sync)# no switch-profile GAD ?
  all-config    Deletion of profile, local and peer configurations
                                        //Usuwa profil, konfigurację lokalną oraz konfigurację zdalną
  local-config  Deletion of profile and local configuration
                                        //Usuwa profil oraz konfigurację lokalną
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Hmm. To dziwne. Wybiorę opcję lokalną, ponieważ nie chcę niczego zepsuć:

```
NX-5K-1(config-sync)# no switch-profile GAD local-config
Verification successful...
Proceeding to delete switch-profile. This might take a while depending on
amount of configuration under a switch-profile.
Please avoid other configuration changes during this time.
Delete Successful
```

Ojej. Przyjrzyjmy się teraz szkodom, dobrze? Spójrz najpierw na interfejs port channel 10, który synchronizowaliśmy wcześniej:

```
NX-5K-1(config-sync)# sho run int po10
interface port-channel10
  description BladeServer
  switchport mode trunk
  vpc 10
  switchport trunk allowed vlan 10,20,30,40
  speed 1000
```

Zauważyłeś, że czegoś brakuje? Tak jest — zniknęła komenda `spanning-tree port type edge trunk`. Ponieważ została ona usunięta tylko lokalnie, prawdopodobnie vPC jest teraz zerwany. Rzeczywiście:

```
NX-5K-1(config-sync)# sho vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id                     : 1
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status: success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 1
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ -------------------------------------------------
1    Po1    up     1

vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
10     Po10        up     failed      vPC type-1 configuration   -
                                       incompatible - STP
                                       interface port type
                                       inconsistent //Brak kompatybilności konfiguracji vPC typu 1 —
                                                    // niezgodny typ portu interfejsu STP
```

Tak, jest źle. Wykorzystajmy tę okazję do przeanalizowania jednej z korzyści z synchronizacji konfiguracji. Ponieważ zsynchronizowaliśmy konfigurację na drugim przełączniku, a wszystkie linie komend wykonanych w trakcie synchronizacji są przechowywane w bazie danych, możemy je łatwo odzyskać. Trzeba tylko skonfigurować odpowiedni profil przełącznika, wskazać partnera do synchronizacji i poczekać:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
NX-5K-1# config sync
NX-5K-1(config-sync)# switch-profile GAD
Switch-Profile started, Profile ID is 1
NX-5K-1(config-sync-sp)# sync-peers destination 10.0.0.2
```

Po kilku sekundach wszystko wraca do poprzedniego stanu. Możemy teraz zobaczyć te linie kodów w działającej konfiguracji (choć naprawdę ich tam nie ma):

```
NX-5K-1(config-sync-sp)# sho run int po 10

!Command: show running-config interface port-channel10
!Time: Mon Jan 17 23:05:25 2011

version 5.0(2)N2(1)

interface port-channel10
  description BladeServer
  switchport mode trunk
  vpc 10
  switchport trunk allowed vlan 10,20,30,40
  spanning-tree port type edge trunk
  speed 1000
```

A tutaj możemy zobaczyć linie na powrót w lokalnej bazie danych przełącznika:

```
NX-5K-1(config-sync-sp)# sho run switch-profile

switch-profile GAD
  sync-peers destination 10.0.0.2

  interface port-channel10
    spanning-tree port type edge trunk
```

Jest jeszcze jedna rzecz, na którą trzeba uważać przy synchronizacji konfiguracji. Jeśli skonfigurujesz interfejs lub FEX z wykorzystaniem tej funkcji, nie będziesz już mógł go konfigurować w trybie *config term*:

```
NX-5K-1# conf t
Enter configuration commands, one per line.  End with CNTL/Z.
NX-5K-1(config)# int po 10
NX-5K-1(config-if)# no spanning-tree port type edge trunk

Error: Command is not mutually exclusive //Błąd: komenda nie jest wzajemnie wykluczona
```

Synchronizacja konfiguracji jest bardzo użytecznym narzędziem. Może być jednak nieco kłopotliwa, szczególnie jeśli dopiero się z nią zapoznajesz. Gorąco zalecam, abyś najpierw pobawił się tą funkcją w warunkach laboratoryjnych lub choćby przetestował na osobnym sprzęcie, zanim zastosujesz ją do konfiguracji urządzeń sieci produkcyjnej.

Cofanie zmian w konfiguracji

Jedną z nowych, fajnych funkcji Nexusa jest możliwość łatwego cofania zmian dokonanych w konfiguracji (ang. *configuration rollback*). Jak prawdopodobnie wiesz, IOS posiada addytywne CLI. Oznacza to, że komendy są zawsze dodawane na górze wiersza poleceń. Nie ma łatwego sposobu na wymazanie całej konfiguracji i załadowanie innej bez konieczności restartowania urządzenia. Jest to szczególnie dotkliwe w środowisku laboratoryjnym, ponieważ jedynym sposobem wyczyszczenia konfiguracji jest wpisanie komendy `write-erase`, zrestartowanie systemu, a następnie wprowadzenie w jakiś sposób nowej konfiguracji. Pochłania to dużo czasu i jest dość nieporęczne, żeby nie użyć mocniejszych słów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Cofanie zmian w konfiguracji jest obsługiwane przez przełączniki Nexus 5000 z NX-OS w wersji 5.0(2)N1(1) lub nowszej.

Nexus umożliwia cofanie całego bloku konfiguracji — z kontekstami VDC i innymi rzeczami — do uprzednio zapisanego punktu kontrolnego. Zapisałem kilka punktów konfiguracji, ale zanim je zobaczymy, utwórzmy nowy punkt:

```
NX-7K-2# checkpoint 100526-1 description Before VDC Re-order
.............................Done
```

Możesz przejrzeć zapisane punkty kontrolne konfiguracji, stosując komendę wykonawczą `rollback running-config checkpoint ?`. Nie jest to komenda trybu konfiguracyjnego:

```
NX-7K-2# rollback running-config checkpoint ?
  100526-1  Checkpoint name
  GAD       Checkpoint name
  Clean     Checkpoint name
```

Komenda `show checkpoint summary` pokaże Ci również szczegóły dotyczące każdego punktu kontrolnego:

```
Annie# sho checkpoint summary
User Checkpoint Summary
--------------------------------------------------------------------------------
1) Clean:
Created by admin
Created at Thu, 01:29:40 20 May 2010
Size is 23,601 bytes
Description: None

2) GAD:
Created by gtac
Created at Thu, 08:27:05 20 May 2010
Size is 23,928 bytes
Description: None

3) 100526-1:
Created by gtac
Created at Wed, 23:55:30 26 May 2010
Size is 23,934 bytes
Description: Before VDC Re-order
```

Możesz przejrzeć zawartość punktu kontrolnego, wpisując polecenie `show checkpoint` *nazwa-*
↪*punktu*:

```
NX-7K-2# sho checkpoint GAD
--------------------------------------------------------------------------------
Name: GAD
!Command: Checkpoint cmd vdc 1

!Time: Thu May 20 01:29:40 2010

version 4.2(4)
poweroff module 1
poweroff module 2
poweroff module 9
power redundancy-mode combined force
feature interface-vlan
username admin password 5 $1$wlVidSSk$dMAUWLJLkkFPH3Dk7dEl7/  role network-admin
ip domain-lookup
ip host Annie 10.1.1.102
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
hostname Annie
ip access-list copp-system-acl-bgp
  10 permit tcp any gt 1024 any eq bgp
  20 permit tcp any eq bgp any gt 1024
ipv6 access-list copp-system-acl-bgp6
  10 permit tcp any gt 1024 any eq bgp
  20 permit tcp any eq bgp any gt 1024
ip access-list copp-system-acl-cts
[fragment usunięty]
```

Zachowaj ostrożność, kiedy będziesz się bawił się tą funkcją, ponieważ może ona spowodować spustoszenie w działającym systemie, jeśli nie będziesz uważał. Poniżej przywrócę konfigurację do punktu kontrolnego GAD:

```
NX-7K-2(config)# rollback running-config checkpoint GAD
Note: Applying config parallelly may fail Rollback verification
        //Uwaga: zastosowanie konfiguracji równolegle może uniemożliwić weryfikację przywracania punktu konfiguracji
Collecting Running-Config
Generating Rollback Patch
Executing Rollback Patch
Generating Running-config for verification
Generating Patch for verification
Annie(config)#
```

Wydaje się dość proste. Ostrzeżenie, które otrzymaliśmy, jest rezultatem tego, że cofnąłem zmiany w konfiguracji, będąc w trybie konfiguracyjnym przełącznika. System ostrzega mnie, że nie powinienem tego robić. Jestem jednak uparty (czytaj: głupi), więc brnę dalej. Ale zaraz — zmieniła się nazwa hosta. Co jeszcze się zmieniło?

```
Annie(config)# sho vdc
vdc_id  vdc_name                        state              mac
------  --------                        -----              ----------
1       Goober                          active             00:26:98:1b:f0:41
```

Goober? Co stało się z kontekstami VDC Annie oraz Guinness? Zostały wymazane, kiedy przywróciłem punkt kontrolny.

Jeśli chcesz pobawić się w cofanie zmian w konfiguracji i zamierzasz przywrócić poprzedni punkt kontrolny, zawsze utwórz najpierw nowy punkt kontrolny!

Nagły atak paniki — potrzebuję moje konteksty VDC z powrotem! Szczęśliwie utworzyłem punkt kontrolny, zanim zacząłem się w to bawić. Jak się nazywał ten punkt kontrolny?

```
NX-7K-2# rollback running-config checkpoint ?
  100526-1  Checkpoint name
  GAD       Checkpoint name
  Clean     Checkpoint name
```

Przywrócenie mojej konfiguracji będzie wymagało jedynie załadowania odpowiedniego punktu kontrolnego:

```
Annie(config)# rollback running-config checkpoint 100526-1
Note: Applying config parallelly may fail Rollback verification
Collecting Running-Config
Generating Rollback Patch
Executing Rollback Patch
Generating Running-config for verification
Generating Patch for verification
NX-7K-2(config)#
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Nazwa hosta jest teraz właściwa, a co z moimi kontekstami VDC?

```
NX-7K-2# sho vdc
vdc_id  vdc_name                        state               mac
------  --------                        -----               ----------
1       NX-7K-2                         active              00:26:98:1b:f0:41
2       Annie                           active              00:26:98:1b:f0:42
3       Guinness                        active              00:26:98:1b:f0:43
```

Wszystko jest na miejscu i jak trzeba.

Cofanie zmian w konfiguracji jest fantastycznym narzędziem. Zalecam jednak, byś zawsze tworzył nowy punkt kontrolny przed wprowadzeniem istotnych zmian. Może Ci to ocalić posadę.

Aktualizacja NX-OS

Przełącznik Nexus możesz zaktualizować podobnie, jak się to robi w przypadku IOS. Istnieje jednak znacznie lepszy sposób. W przypadku NX-OS pliki obrazów (ang. *image files*) są bardzo duże. Wersja 4.2.1 dla przełączników 5000 ma około 182 MB. Ponadto musisz zainstalować jeszcze odpowiedni obraz rozruchowy (ang. *kickstart image*). Pliki rozruchowe mają z reguły około 20 MB.

Poniżej opiszę etapy aktualizacji kodu dla przełącznika Nexus 5020 z wersji 4.1 do 5.0. Wykorzystam serwer TFTP zainstalowany na moim MacBook Pro, podłączonym do 5020 przez interfejs mgmt0 tego przełącznika. Zacznę od znajomej instrukcji `copy tftp: boot flash:`:

```
NX-5K-1# copy tftp: bootflash:
```

Przełącznik prosi mnie o nazwę nowego pliku, którą podaję, jak następuje:

```
Enter source filename: n5000-uk9-kickstart.5.0.2.N2.1.bin
```

Kolejne zapytanie może zmylić tych, którzy dopiero uczą się Nexusa. Przełącznik pyta, którego VRF chcesz użyć. Domyślnym VRF-em jest ten, który nazywa się VRF **domyślny** (ang. *default*), co jest niepoprawne. Musimy poinformować przełącznik, że wykorzystamy VRF zarządzający. Pamiętaj, że Nexus lubi separować magistralę sterowania od magistrali danych, a interfejs mgmt0 jest częścią VRF-u zarządzającego:

```
Enter vrf (If no input, current vrf 'default' is considered): management
```

Od tego momentu wszystko powinno wydawać się całkiem znajome:

```
Enter hostname for the tftp server: 10.1.1.1

Trying to connect to tftp server......
Connection to Server Established.
```

Teraz przełącznik dokona transferu pliku. Ten sam proces musi zostać wykonany również dla pliku głównego obrazu. Jeśli korzystasz z TFTP jak ja, przygotuj się, że trochę to potrwa. Oczywiście możesz skorzystać z FTP lub nawet z SCP bezpośrednio ze strony Cisco, jeśli przełącznik jest podłączony do internetu.

Po zakończeniu transferu pliku właściwym sposobem aktualizacji Nexusa jest komenda `install all`. Zainstaluje ona jednocześnie obraz rozruchowy i systemowy, a przy odpowiednich warunkach dla systemu może się to odbyć bez zakłócania przepływu pakietów przez przełącznik. Zostanie również zaktualizowany BIOS i obrazy dla modułów. Przejdźmy przez ten proces dla przełącznika Nexus 5020.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Na początek musimy się upewnić, że nasza konfiguracja nie jest zagrożona w procesie aktualizacji. Ta sprytna komenda przeanalizuje Twoją konfigurację i powiadomi Cię, czy jakieś jej elementy mogą zostać utracone podczas aktualizacji (lub degradacji wersji, bo i taki przypadek może zaistnieć):

```
NX-5K-1# sho incompatibility system n5000-uk9.5.0.2.N2.1.bin
No incompatible configurations //Brak niezgodności w konfiguracji
```

Jeśli o tym zapomnisz, system zrobi to automatycznie w trakcie procesu instalacji i ostrzeże Cię następującym komunikatem:

```
"Running-config contains configuration that is incompatible with the new image
(strict incompatibility).
        //Działająca konfiguracja zawiera elementy niekompatybilne z nowym obrazem (ścisła niezgodność)
 Please run 'show incompatibility system <image>' command to find out which feature
needs to be disabled."
        //Zastosuj polecenie 'show incompatibility system <obraz>', aby sprawdzić, które funkcje muszą zostać wyłączone
```

Teraz jestem już zadowolony, że nie spowoduję problemów z działaniem sieci, mogę więc przejść do instalowania obrazu komendą `install`:

```
NX-5K-1# install all kickstart n5000-uk9-kickstart.5.0.2.N2.1.bin
system n5000-uk9.5.0.2.N2.1.bin

Verifying image bootflash:/n5000-uk9-kickstart.5.0.2.N2.1.bin for boot
variable "kickstart".
[####################] 100% -- SUCCESS
Verifying image bootflash:/n5000-uk9.5.0.2.N2.1.bin for boot variable "system".
[####################] 100% -- SUCCESS

Verifying image type.
[####################] 100% -- SUCCESS

Extracting "system" version from image bootflash:/n5000-uk9.5.0.2.N2.1.bin.
[####################] 100% -- SUCCESS

Extracting "kickstart" version from image
bootflash:/n5000-uk9-kickstart.5.0.2.N2.1.bin.
[####################] 100% -- SUCCESS

Extracting "bios" version from image bootflash:/n5000-uk9.5.0.2.N2.1.bin.
[####################] 100% -- SUCCESS

Notifying services about system upgrade.
[####################] 100% -- SUCCESS
```

Jeśli aktualizacja nie przebiegnie bezproblemowo, otrzymasz poniższą informację (lub zbliżoną, w zależności od przyczyny):

```
Compatibility check is done: //Przeprowadzono sprawdzenie zgodności
Module  bootable          Impact  Install-type  Reason
------  --------  --------------  ------------  ------
     1       yes      disruptive         reset  Current running-config is not
supported by new image
                        //Obecnie działająca konfiguracja nie jest obsługiwana przez nowy obraz — wpływ destrukcyjny
```

Szczęśliwie, my nie mamy problemów, więc otrzymujemy taką informację:

```
Compatibility check is done:
Module  bootable          Impact  Install-type  Reason
------  --------  --------------  ------------  ------
     1       yes  non-disruptive         reset
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Destrukcyjny (ang. *disruptive*) wpływ aktualizacji oznacza, że przełącznik zostanie zrestartowany, co wstrzyma cały ruch przepływających przez niego pakietów. Niedestrukcyjna (ang. *nondisruptive*) czy też **bezproblemowa** (ang. *hitless*) aktualizacja spowoduje przeładowanie kernela, ale w trakcie tego procesu pakiety będą nadal płynąć. Jest to możliwe, gdyż Nexus oddziela przepływ pakietów od przetwarzania CPU lub, jak głoszą materiały marketingowe, magistrala sterowania jest tu niezależna od magistrali danych.

W końcu otrzymujemy poniższe zestawienie, które informuje nas, co się wydarzy:

```
Images will be upgraded according to following table:
Module       Image         Running-Version             New-Version  Upg-Required
------  ----------  ----------------------  ----------------------  ------------
     1      system             4.2(1)N2(1)             5.0(2)N2(1)           yes
     1   kickstart             4.2(1)N2(1)             5.0(2)N2(1)           yes
     1        bios        v1.2.0(06/19/08)        v1.3.0(09/08/09)           yes
     1   power-seq                    v1.2                    v1.2            no
```

W tym punkcie, jeśli Twój Nexus musi zostać zrestartowany (aktualizacja destrukcyjna), otrzymasz następujący komunikat:

```
Switch will be reloaded for disruptive upgrade.
                                        //Przełącznik zostanie zrestartowany z uwagi na aktualizację destrukcyjną.
Do you want to continue with the installation (y/n)?  [n] //Czy chcesz kontynuować instalację (tak/nie)?
```

W naszym przypadku nie ma takiej potrzeby, więc dostajemy taką wiadomość:

```
Do you want to continue with the installation (y/n)?  [n] y
```

Teraz zostaniemy potraktowani opisami poszczególnych etapów procesu aktualizacji:

```
Install is in progress, please wait.

Notifying services about the upgrade.
[####################] 100% -- SUCCESS

Setting boot variables.
[####################] 100% -- SUCCESS

Performing configuration copy.
[####################] 100% -- SUCCESS

Module 1: Refreshing compact flash and upgrading bios/loader/bootrom/power-seq.
Warning: please do not remove or power off the module at this time.
Note: Power-seq upgrade needs a power-cycle to take into effect.
On success of power-seq upgrade, SWITCH OFF THE POWER to the system and then,
power it up.
[####################] 100% -- SUCCESS

Upgrade can no longer be aborted, any failure will result in a disruptive upgrade.

Requesting Line Cards to stop communication.
[####################] 100% -- SUCCESS

Requesting Sup Apps to stop communication.
[####################] 100% -- SUCCESS

Freeing memory in the file system.
[####################] 100% -- SUCCESS

Loading images into memory.
[####################] 100% -- SUCCESS
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Saving supervisor runtime state.
[###################] 100% -- SUCCESS

Saving mts state.
 writing reset reason 88, <NULL>
[###################] 100% -- SUCCESS

Rebooting the switch to proceed with the upgrade. //Restart przełącznika w celu kontynuowania aktualizacji
```

Ta ostatnia linijka wygląda mocno niepokojąco. Jest to jednak tylko informacja o restarcie procesów związanych z przetwarzaniem przez CPU. Nasze połączenie z przełącznikiem zostanie teraz przerwane, ale przepływ pakietów nie zostanie wstrzymany w trakcie tego procesu:

```
All telnet and ssh connections will now be terminated.
Starting new kernel
Calling kexec callback
Moving to new kernel
Calling into reboot_code_buffer code
Starting kernel...
Usage: init 0123456SsQqAaBbCcUu
INIT: version 2.85 booting
platform_type cmdline parameter not found. Asssuming Oregon.
I2C - Mezz present
sprom_drv_init_platform: nuova_i2c_register_get_card_index
autoneg unmodified, ignoring
autoneg unmodified, ignoring
Checking all filesystems..... done.
.
Loading system software
Uncompressing system image: bootflash:/n5000-uk9.5.0.2.N2.1.bin

Loading plugin 0: core_plugin...
Loading plugin 1: eth_plugin...
ethernet switching mode
INIT: Entering runlevel: 3
Exporting directories for NFS kernel daemon...done.
Starting NFS kernel daemon:rpc.nfsd.
rpc.mountddone.

Setting envvar: SYSMGR_SERVICE_NAME to muxif_service
Set envvar SYSMGR_SERVICE_NAME to muxif_service
/isan/bin/muxif_config: argc:2
muxif_init....vacl: ret: 0
Set name-type for VLAN subsystem. Should be visible in /proc/net/vlan/config
Added VLAN with VID == 4042 to IF -:muxif:-
2011 Jan  6 17:49:39 NX-5K-1 %$ VDC-1 %$ %USER-2-SYSTEM_MSG: CLIS: loading cmd
files begin  - clis
2011 Jan  6 17:50:07 NX-5K-1 %$ VDC-1 %$ Jan  6 17:50:07 %KERN-2-SYSTEM_MSG:
Starting kernel... - kernel
2011 Jan  6 17:50:07 NX-5K-1 %$ VDC-1 %$ Jan  6 17:50:07 %KERN-0-SYSTEM_MSG:
platform_type cmdline parameter not found. Asssuming Oregon. - kernel
2011 Jan  6 17:50:07 NX-5K-1 %$ VDC-1 %$ Jan  6 17:50:07 %KERN-0-SYSTEM_MSG:
I2C - Mezz present  - kernel
2011 Jan  6 17:50:07 NX-5K-1 %$ VDC-1 %$ Jan  6 17:50:07 %KERN-0-SYSTEM_MSG:
sprom_drv_init_platform: nuova_i2c_register_get_card_index  - kernel

Continuing with installation process, please wait.
The login will be disabled until the installation is completed.

Performing supervisor state verification.
[###################] 100% -- SUCCESS
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Supervisor non-disruptive upgrade successful.
2011 Jan  6 17:50:40 NX-5K-1 %$ VDC-1 %$ %USER-2-SYSTEM_MSG: CLIS: loading cmd
files end  - clis
2011 Jan  6 17:50:40 NX-5K-1 %$ VDC-1 %$ %USER-2-SYSTEM_MSG: CLIS: init begin  - clis

Install has been successful.
NX-5K-1 login:
```

Z naszego punktu widzenia ponownie wygląda na to, że przełącznik został zrestartowany. Była to jednak aktualizacja bezproblemowa, więc przerwana została tylko komunikacja z samym przełącznikiem — taka jak SNMP, telnet, SSH itp.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 19.

Cechy przełącznika Catalyst 3750

Catalyst 3750 jest kolejnym krokiem w ewolucji bardzo popularnego przełącznika 3550 o stałej konfiguracji (ang. *fixed-configuration switch*). 3550 był pierwszym przełącznikiem wielowarstwowym oferującym bardzo szeroki wachlarz funkcji w stosunku do ceny. Został później zastąpiony modelem 3560. Przełącznik 3750 jest potężniejszym urządzeniem wprowadzającym m.in. funkcję prawdziwego łączenia w stos, której zabrakło w 3560.

Nie starczyłoby miejsca na opisanie wszystkich możliwości 3750 w jednym rozdziale, skupiłem się więc na tych funkcjach, które uznałem za najbardziej użyteczne w praktyce. Celowo pominąłem część szczegółów dotyczących opisanych funkcji. Zamiast tego przedstawiłem zagadnienia, które, jak wierzę, będą Ci potrzebne, aby w pełni wykorzystać zalety tych funkcji.

Nie wszystkie z omówionych funkcji są charakterystyczne jedynie dla przełącznika 3750. Polecenia mogą być identyczne również dla innych modeli, ale zamieściłem w tym rozdziale jedynie przykłady zaczerpnięte z 3750. Jak zawsze wszystkie szczegółowe informacje na temat funkcji znajdziesz w dokumentacji Cisco.

Łączenie w stos

Jedną z głównych wad przełączników 3550 i 3560 był sposób łączenia ich w stos. **Łączenie w stos** (ang. *stacking*) polega na zlinkowaniu ze sobą kilku przełączników (z reguły tego samego typu) w taki sposób, aby utworzyły jeden przełącznik logiczny z pojedynczym adresem IP do zarządzania. Kiedy połączysz się z tym adresem IP sesją telnet lub SSH, będziesz mógł zarządzać całym stosem, jakby było to jedno urządzenie.

W 3550 zastosowano tryb łączenia w stos, który wymaga zainstalowania w gigabitowych slotach GBIC modułów zwanych **stacking GBIC**. Wiąże się to z ograniczeniem prędkości płyty *backplane* do 1Gb/s oraz wykorzystaniem jedynych gigabitowych interfejsów dostępnych w tych przełącznikach. Tak więc, jeśli połączysz w stos swoje przełączniki, pozbawisz się gigabitowych uplinków. W modelu 3560 stosuje się specjalny kabel SFP do łączenia przełączników. Jest on dostępny również w wersji z portami RJ-45 10/100/1000, ale w takim przypadku nadal musisz wykorzystać jeden z dwóch uplinków światłowodowych.

W przełącznikach 3750 wprowadzono bardziej wydajne rozwiązanie, stosując specjalne kable do łączenia w stos, które podpinane są do złączy znajdujących się na tylnym panelu obudowy. Połączenie *backplane* ma wtedy prędkość 32 Gb/s (lub wyższą, w zależności od modelu) oraz

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

nie wymaga angażowania portów na przednim panelu przełącznika. Nie będę wdawał się w szczegóły fizycznego połączenia stosu przełączników, gdyż dokumentacja Cisco w tym temacie jest więcej niż wystarczająca.

Dla każdego 24-portowego przełącznika 3750 interfejsy ponumerowane są w zakresie Gi1/0/1 – – Gi1/0/23. Są one opisane następująco: **typ-interfejsu nr-członka-stosu/moduł-nr/port-nr**. **Typ interfejsu** przełącznika 3750 to zazwyczaj Gigabit Ethernet, choć niektóre modele obsługują porty 10 Gb/s. **Nr-członka-stosu** ma wartość 1 dla osobnego przełącznika. W przypadku stosu numer ten odzwierciedla pozycję, jaką przełącznik zajmuje w tym stosie. **Moduł-nr** dla 3750 zawsze wynosi 0. **Port-nr** określa fizyczny numer portu w danym przełączniku. W ten sposób port 14 na trzecim przełączniku w stosie będzie opisany Gi3/0/14.

Jeśli rozbijesz stos przełączników 3750, żeby użyć ich niezależnie, mogą nie powrócić do domyślnej numeracji. Za przykład niech posłuży przełącznik, który wyciągnąłem ze stosu. Był tam przełącznikiem 2, ale teraz jest osobnym urządzeniem. Problem w tym, że nadal „sądzi", że jest przełącznikiem 2:

```
IOS-3750#sho switch detail
Switch/Stack Mac Address : dc7b.94dd.2280
                                          H/W   Current
Switch#  Role   Mac Address     Priority Version  State
------------------------------------------------------------
*2       Master dc7b.94dd.2280     1      0       Ready
```

Może się to nie wydawać wielkim problemem. Jednak irytujące jest, że również wszystkie interfejsy odzwierciedlają nadal drugą pozycję tego przełącznika w stosie:

```
IOS-3750#sho ip int brie
Interface              IP-Address      OK? Method Status                Protocol
Vlan1                  unassigned      YES NVRAM  administratively down down
GigabitEthernet2/0/1   unassigned      YES unset  down                  down
GigabitEthernet2/0/2   unassigned      YES unset  down                  down
GigabitEthernet2/0/3   unassigned      YES unset  down                  down
[fragment usunięty]
```

Łatwo to poprawić, jednak robię to tak rzadko, że prawie zapomniałem jak. Musisz tylko wpisać komendę `switch switch-nr renumber nowy-switch-nr`, a następnie zrestartować przełącznik. Zobaczmy to w akcji. Choć jest to prosty proces, może nie działać tak, jak byś tego oczekiwał. Ponieważ chcemy, żeby przełącznik 2 miał teraz numer 1, wpiszę komendę `switch 2 renumber 1`:

```
IOS-3750(config)#switch 2 renumber 1
WARNING: Changing the switch number may result in a
configuration change for that switch.
    //Uwaga: zmiana numeru przełącznika może spowodować zmianę konfiguracji tego przełącznika
The interface configuration associated with the old switch
number will remain as a provisioned configuration.
    //Konfiguracja interfejsu odnosząca się do numeru starego przełącznika zostanie zachowana jako konfiguracja zabezpieczona
Do you want to continue?[confirm] //Czy chcesz kontynuować?
Changing Switch Number 2 to Switch Number 1 //Zmiana numeru przełącznika 2 na przełącznik 1
New Switch Number will be effective after next reboot //Nowy numer pojawi się po najbliższym restarcie
```

Tak jak zostałem poinstruowany, muszę teraz zrestartować przełącznik. Wpisuję więc polecenia `mem` oraz `reload`. Kiedy to zrobię, moje interfejsy będą miały właściwe nazwy. Hura!

```
IOS-3750#sho ip int brie
Interface              IP-Address      OK? Method Status                Protocol
Vlan1                  unassigned      YES NVRAM  administratively down down
GigabitEthernet1/0/1   unassigned      YES unset  down                  down
GigabitEthernet1/0/2   unassigned      YES unset  down                  down
GigabitEthernet1/0/3   unassigned      YES unset  down                  down
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
GigabitEthernet1/0/4   unassigned      YES unset  down                  down
[fragment usunięty]
```

Jest jednak pewien problem. Listing polecenia `show switch detail` pokazuje nam dwa przełączniki, choć oczywiście mamy tylko przełącznik 1:

```
IOS-3750#sho switch detail
Switch/Stack Mac Address : dc7b.94dd.2280
                                          H/W   Current
Switch#  Role   Mac Address     Priority Version  State
-----------------------------------------------------------
*1       Master dc7b.94dd.2280     1      0       Ready
 2       Member 0000.0000.0000     0      0       Provisioned
```

Nabierze to sensu, jeśli spojrzysz na konfigurację i uświadomisz sobie, że konfiguracja dla wszystkich Twoich interfejsów zniknęła. Z technicznego punktu widzenia nie jest to całkiem właściwe stwierdzenie. Konfiguracja nadal istnieje, jednak przełącznik 2 został uznany za nieobecny. Z tego powodu jego adres MAC składa się z samych zer, a jego status to „provisioned" (zabezpieczony). Możesz odnaleźć te interfejsy w konfiguracji, wpisując inne komendy `show`:

```
IOS-3750#sho ip int brie | beg 1/0/50
GigabitEthernet1/0/50  unassigned      YES unset  down                  down
GigabitEthernet1/0/51  unassigned      YES unset  down                  down
GigabitEthernet1/0/52  unassigned      YES unset  down                  down
GigabitEthernet2/0/1   unassigned      YES unset  down                  down
GigabitEthernet2/0/2   unassigned      YES unset  down                  down
[fragment usunięty]
```

Zapamiętaj tylko, że ta komenda nie przenumeruje Twoich interfejsów i nie przeniesie konfiguracji. Musisz zrobić to sam.

Zakresy interfejsów

Zakresy interfejsów (ang. *interfaces ranges*) są bardzo użytecznym dodatkiem do IOS. Zamiast wprowadzać te same komendy dla wielu interfejsów po kolei, możesz określić zakres interfejsów, których komendy mają dotyczyć. Kiedy zastosujesz komendy do zakresu interfejsów, parser zreplikuje je na każdy interfejs z podanego zakresu. W przypadku przełącznika z 96 interfejsami może to zaoszczędzić Ci godziny pracy — szczególnie przy pierwszej konfiguracji urządzenia.

Zakresy interfejsów składają się z list interfejsów. Możesz definiować interfejsy indywidualnie lub — jeśli jest to kilka interfejsów następujących bezpośrednio po sobie — grupować je w zakresy. Interfejsy podawane indywidualnie musisz rozdzielać przecinkami, natomiast zakresy interfejsów określa się, podając interfejs początkowy i końcowy, rozdzielone myślnikiem. Zdefiniuję teraz dla funkcji zakresu dwa indywidualne interfejsy `g1/0/10` i `g1/0/12`:

```
3750(config)#interface range g1/0/10 , g1/0/12
3750(config-if-range)#
```

Kiedy jesteś już w trybie konfiguracji dla zakresu interfejsów, czyli **config-if-range**, wszystkie komendy, które wpiszesz, będą zastosowane do każdego interfejsu zdefiniowanego w zakresie.

Poniżej przykład zdefiniowania zakresu portów od `g1/0/10` do `g1/0/12` (włącznie). Podając końcowy interfejs z zakresu, wystarczy wpisać jedynie jego numer portu bez podawania całej składni opisującej interfejs. Nie musisz więc definiować zakresu w taki sposób: `g1/0/10 – g1/0/12`. Wystarczy tak: `g1/0/10 – 12`:

```
3750(config)#interface range g1/0/10 - 12
3750(config-if-range)#
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Aby odnieść się do kilku zakresów interfejsów jednocześnie, wprowadzaj je rozdzielone przecinkami. Poniżej zdefiniowałem zakresy od `g1/0/10` do `g1/0/12` oraz od `g1/0/18` do `g1/0/20`:

```
3750(config)#interface range g1/0/10 - 12 , g1/0/18 - 20
3750(config-if-range)#
```

Oprócz definiowania list interfejsów do bieżącego użytku możesz również zapisywać te listy do zastosowania w przyszłości. Aby to zrobić, zdefiniuj zakres interfejsów komendą `define interface-range` *`nazwa-makro`*. Tutaj stworzyłem makro o nazwie `Servers`:

```
3750(config)#define interface-range Servers g1/0/10 , g1/0/12 - 14
3750(config)#
```

Nie pomyl tych makr z inną funkcją, zwaną **smartport macros** (opisana w następnym podrozdziale). Te dwie funkcje nie są ze sobą powiązane. Żeby jeszcze bardziej skomplikować temat, możesz stosować *smartport macro* do makra zakresu interfejsów. Jeśli chodzi o programistów, to funkcja *smartport macros* powinna wydać się Wam bardziej zbliżona do standardowych makr, do których jesteście przyzwyczajeni.

Kiedy masz już zdefiniowane makro zakresu interfejsów, możesz każdorazowo odwoływać się do niego, stosując polecenie `interface range macro` *`nazwa-makro`*:

```
3750(config)#interface range macro Servers
3750(config-if-range)#
```

Makra

Makra, zwane przez Cisco **smartport macros**, to grupy komend zapisane pod nazwą referencyjną. Makra mogą się przydać, jeśli okazuje się, że często musisz wpisywać określoną sekwencję poleceń. Przykładowo, jeśli dodajesz wiele serwerów. Za każdym razem, kiedy dodajesz serwer, wykonujesz tę samą konfigurację komend dla przeznaczonego dla tego serwera interfejsu przełącznika. Mógłbyś stworzyć makro, które będzie wykonywać wszystkie te komendy automatycznie. Wtedy po prostu uruchamiasz makro za każdym razem, kiedy musisz dodać kolejny serwer do przełącznika.

Makra tworzysz poleceniem `macro`. Istnieją dwa typy makr: **globalne** (ang. *global*) oraz **makra interfejsu** (ang. *interface*). Makro interfejsu (jest to domyślny typ makr) jest stosowane do jednego lub kilku interfejsów. Aby stworzyć makro globalne, dodaj argument `global` przy tworzeniu makra.

Sposób tworzenia makr jest nieco dziwny, ponieważ polecenia nie są analizowane w trakcie ich wpisywania. W rezultacie możesz podać błędną komendę, nie wywołując komunikatu błędu. Aby stworzyć makro, wpisz najpierw polecenie `macro name` *`nazwa-makro`*. Ponieważ nie dodajemy argumentu `global`, będzie to makro interfejsu. Następnym krokiem jest podanie po kolei wszystkich komend, które mają być zawarte w tym makrze. Komendy nie są sprawdzane pod kątem składni. Jeśli chcesz zakończyć wpisywanie komend, umieść osobno w ostatniej linii znak (`@`).

Utworzę makro o nazwie SetServerPorts. Składać się będzie z komend: `spanning-tree`, `hhhhhh` (komenda niepoprawna) oraz `description <[ Server ]>`:

```
3750(config)#macro name SetServerPorts
Enter macro commands one per line. End with the character '@'.
spanning-tree portfast
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
hhhhhh
description <[ Server ]>
@
3750(config)#
```

Umieszczenie komendy pomiędzy nawiasami <[oraz]> nie wywołuje żadnego konkretnego skutku w IOS. Jest to coś, co robię od lat, żeby opis wyróżniał się w działającej konfiguracji oraz w listingu komend `show`, takich jak `show interface`.

Jak widzisz, przełącznik zaakceptował wszystkie komendy bez żadnego problemu. Makro, wraz z niepoprawną komendą, pojawiło się teraz w działającej konfiguracji:

```
!
macro name SetServerPorts
spanning-tree portfast
hhhhhh
description <[ Server ]>
@
!
```

Kiedy zastosujesz utworzone makro, parser uruchomi analizę składni i sprawdzanie poprawności komend. Teraz zobaczysz błędy, jeśli wprowadziłeś niepoprawne polecenia. Makro nie zostanie zatrzymane w momencie wykrycia błędów, musisz ich więc wypatrywać:

```
3750(config-if)#macro apply SetServerPorts
%Warning: portfast should only be enabled on ports connected to a single
 host. Connecting hubs, concentrators, switches, bridges, etc... to this
 interface  when portfast is enabled, can cause temporary bridging loops.
 Use with CAUTION

%Portfast has been configured on GigabitEthernet1/0/20 but will only
 have effect when the interface is in a non-trunking mode.
hhhhhh
^
% Invalid input detected at '^' marker. // Wykryto niewłaściwy wpis od miejsca oznaczonego '^'
3750(config-if)#
```

Zauważ, że w przypadku poleceń, które nie generują listingów, nie dostajemy żadnej informacji o ich wykonaniu. To dlatego, że makro jest tylko grupą poleceń, które są uruchamiane, gdy makro jest wywoływane. Jeśli potrzebujesz więcej informacji na temat wykonywania makra, dodaj argument `trace` przy jego uruchamianiu. Zostaną dodane linie tekstu wskazujące na wykonywanie poszczególnych poleceń makra:

```
3750(config-if)#macro trace SetServerPorts
Applying command... 'spanning-tree portfast' //Wykonywanie komendy...
%Warning: portfast should only be enabled on ports connected to a single
 host. Connecting hubs, concentrators, switches, bridges, etc... to this
 interface  when portfast is enabled, can cause temporary bridging loops.
 Use with CAUTION

%Portfast has been configured on GigabitEthernet1/0/20 but will only
 have effect when the interface is in a non-trunking mode.
Applying command... 'hhhhhh' //Wykonywanie komendy...
hhhhhh
 ^
% Invalid input detected at '^' marker.
Applying command... 'description <[ Server ]>' //Wykonywanie komendy...
3750(config-if)#
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kiedy uruchamiasz makro, **opis makra** (ang. *macro description*) jest dodawany do interfejsu lub interfejsów, dla których zostało ono zastosowane. Konfiguracja interfejsu jest uzupełniana o komendę `macro description`, po której następuje nazwa makra:

```
interface GigabitEthernet1/0/20
 description <[ Server ]>
 switchport mode access
 macro description SetServerPorts
 storm-control broadcast level bps 1g 900m
 spanning-tree portfast
```

Możesz dodać własny opis makra komendą `macro description` umieszczoną w samym makrze lub wykonaną z poziomu wiersza poleceń:

```
3750(config-if)#macro description [- Macro Description -]

interface GigabitEthernet1/0/20
 description <[ Server ]>
 switchport mode access
 macro description SetServerPorts | [- Macro Description -]
 storm-control broadcast level bps 1g 900m
 spanning-tree portfast
```

Jak widzisz, za każdym razem, kiedy makro jest uruchamiane lub wykonywane jest polecenie `macro description`, zdefiniowany opis (albo nazwa makra) jest dodawany do komendy `macro description` w konfiguracji interfejsu, do którego makro zostało zastosowane. Iteracje rozdzielane są pionową kreską.

Prostszym sposobem sprawdzenia, gdzie zostały zastosowane makra, jest użycie komendy `show parser macro description`. Poniżej zobaczysz, że uruchomiłem to samo makro kilkakrotnie dla interfejsu `Gi1/0/20`:

```
SW2#sho parser macro description
Interface    Macro Description(s)
------------------------------------------------------------
Gi1/0/20     SetServerPorts | SetServerPorts | SetServerPorts | [- Macro
Description -]
------------------------------------------------------------
```

Aby zobaczyć wszystkie makra na przełączniku, użyj polecenia `show parser macro brief`:

```
3750#sho parser macro brief
    default global   : cisco-global
    default interface: cisco-desktop
    default interface: cisco-phone
    default interface: cisco-switch
    default interface: cisco-router
    default interface: cisco-wireless
    customizable     : SetServerPorts
```

Jak widzisz na listingu, w IOS znajduje się sześć makr domyślnych (na liście oznaczone `default`). Możesz wykorzystać komendę `show parser macro name` *`nazwa-makro`*, aby przejrzeć szczegóły składni wybranego makra:

```
SW2#sho parser macro name cisco-desktop
Macro name : cisco-desktop
Macro type : default interface
# macro keywords $access_vlan
# Basic interface - Enable data VLAN only
# Recommended value for access vlan should not be 1
switchport access vlan $access_vlan
switchport mode access
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
# Enable port security limiting port to a single
# MAC address -- that of desktop
switchport port-security
switchport port-security maximum 1

# Ensure port-security age is greater than one minute
# and use inactivity timer
switchport port-security violation restrict
switchport port-security aging time 2
switchport port-security aging type inactivity

# Configure port as an edge network port
spanning-tree portfast
spanning-tree bpduguard enable
```

To makro zawiera nieco zaawansowanych funkcji, takich jak zmienne czy komentarze. Więcej informacji na temat funkcji makra znajdziesz w dokumentacji Cisco.

Jeśli zastanawiałeś się, jak zastosować *smartport macro* do makra zakresu interfejsów, oto Twoja odpowiedź (makro zakresu interfejsów nazywa się `Workstation`, a *smartport macro* nazywa się `SetPortsPortfast`):

```
SW2(config)#interface range macro Workstations
SW2(config-if-range)# macro apply SetPortsPortfast
```

Linki flex

Linki flex, czyli **flex links**, to interfejsy warstwy drugiej skonfigurowane ręcznie w parach redundantnych *primary/failover*. Protokół Spanning Tree (STP został opisany w rozdziale 8.) standardowo zapewnia funkcjonalność *primary/failover*, ale jego głównym celem jest zapobieganie powstawaniu pętli. Linki flex zapewniają tworzenie linków zapasowych (ang. *backup links*) dla linków podstawowych (ang. *primary links*). W danym momencie tylko jeden z linków pary flex-link jest wykorzystywany do transmisji pakietów.

Linki flex są przeznaczone wyłącznie dla przełączników, które nie pracują w protokole Spanning Tree, i tylko tam powinny być stosowane. Jeśli zostaną skonfigurowane na przełączniku z działającym Spanning Tree, nie będą uczestniczyć w STP.

Linki flex konfiguruje się na interfejsie podstawowym poprzez zdefiniowanie interfejsu zapasowego poleceniem `switchport backup interface`:

```
interface GigabitEthernet1/0/20
 switchport access vlan 10
 switchport backup interface Gi1/0/21
!
interface GigabitEthernet1/0/21
 switchport access vlan 10
```

Dla interfejsu zapasowego nie jest wymagana żadna konfiguracja.

Żaden z linków flex nie może być interfejsem należącym do EtherChannel. Za to EtherChannel **może być** zapasowym linkiem flex dla innego kanału port channel. Pojedynczy interfejs również może być linkiem zapasowym dla EtherChannel. Link zapasowy nie musi być takim samym typem interfejsu jak link podstawowy. Przykładowo interfejs 100 Mb/s może być linkiem zapasowym dla interfejsu 1 Gb/s.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Możesz monitorować linki flex, stosując komendę `show interface switchport backup`:

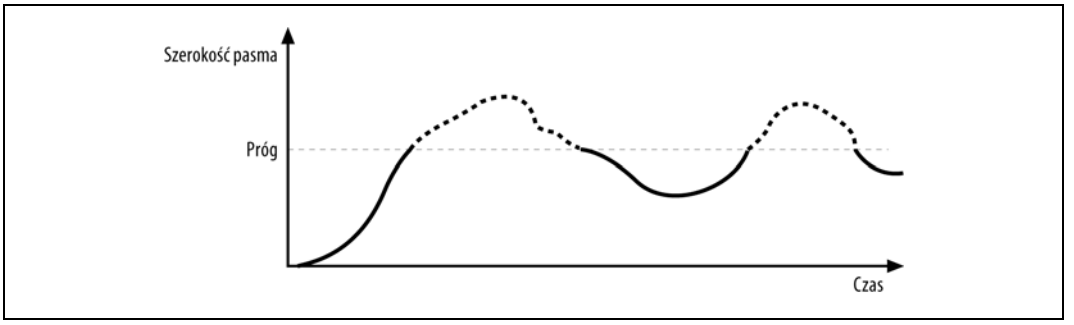
```
3750#sho int switchport backup
Switch Backup Interface Pairs:
Active Interface        Backup Interface        State
------------------------------------------------------------------------
GigabitEthernet1/0/20   GigabitEthernet1/0/21   Active Down/Backup Down
```

Kontrola burz

Kontrola burz (ang. *storm control*) zapobiega „zduszeniu" sieci przez burze rozgłoszeniowe, multicastowe oraz unicastowe. Burze tego typu mogą być powodowane przez wiele czynników, od pętli między mostkami po ataki wirusów. Dzięki kontroli burz możesz ograniczyć liczbę pakietów burzowych przychodzących na port przełącznika. Ruch wychodzący nie jest ograniczany.

Jeśli funkcja kontroli burz jest włączona, przełącznik monitoruje pakiety przychodzące na skonfigurowany interfejs. Co 200 milisekund odczytuje poziom emisji pakietów unicastowych, multicastowych i rozgłoszeniowych, a następnie porównuje go z ustawionym progiem. Pakiety, które przekraczają wyznaczony próg, są odrzucane.

Brzmi to dość prosto, ale w praktyce działa inaczej, niż można by oczekiwać. Kiedy pierwszy raz zetknąłem się z tą funkcją, potraktowałem powyższy opis dosłownie — czyli uznałem, że w dowolnym momencie typy pakietów, które skonfigurowałem do monitorowania, będą mogły napływać do przełącznika, dopóki nie zostanie osiągnięty pewien próg (podobnie jak na rysunku 19.1). Rzeczywistość jest jednak nieco bardziej skomplikowana.



Rysunek 19.1. Niepoprawny model kontroli burz

W rzeczywistości przełącznik monitoruje interfejs, gromadząc statystyki dla 200-milisekundowych przedziałów czasu. Jeśli pod koniec trwania jednego przedziału próg zostaje przekroczony, podejmowana jest zdefiniowana (lub domyślna) akcja dla kolejnego przedziału 200 ms.

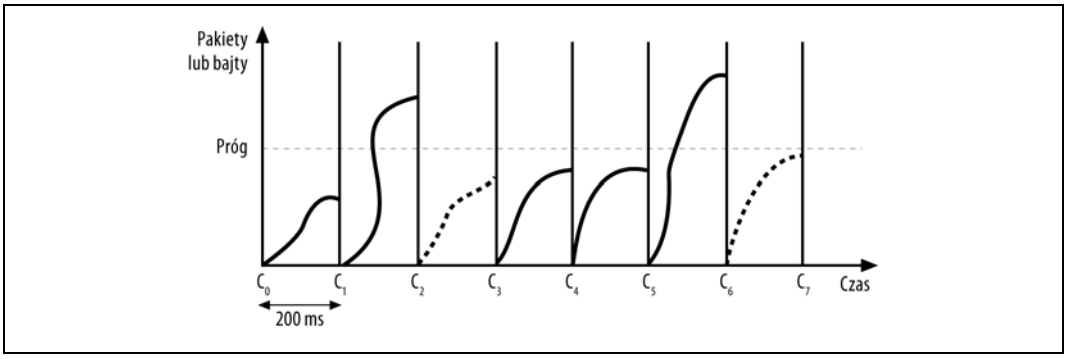
Rysunek 19.2 przedstawia właściwy sposób działania funkcji kontroli burz. Natężenie ruchu mierzone jest w przedziałach co 200 ms, przedstawionych na wykresie jako C0, C1 itd. Jeśli poziom natężenia dla monitorowanego rodzaju ruchu nie przekroczy progu w określonym

200-milisekundowym interwale, kolejny interwał pozostawiany jest bez zmian. W tym przykładzie przy osiągnięciu punktu C1 próg nie został przekroczony. Dla następnego przedziału (kończącego się w punkcie C2) nie jest więc podejmowana żadna akcja. Jednak w trakcie inter-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

wału C2 zostaje przekroczony próg, więc pakiety otrzymane w następnym interwale (C3) są odrzucane. Istotne jest to, że w trakcie trwania każdego interwału otrzymywane pakiety lub bajty są liczone jedynie od punktu początkowego dla tego interwału.

Dla interwału, w którym pakiety są odrzucane, ruch na interfejsie jest nadal monitorowany (pakiety są otrzymywane, ale nie są przekazywane do pozostałych interfejsów tego przełącznika). Jeśli liczba pakietów znów przekroczy zdefiniowany próg, pakiety dla następnego interwału będą ponownie odrzucane. Natomiast gdy liczba pakietów będzie poniżej progu, jak w przypadku interwału C3 na rysunku 19.2, pakiety w następnym interwale będą przyjmowane.



Rysunek 19.2. Właściwy model działania funkcji kontroli burz

Ponieważ pakiety nie są otrzymywane w sposób ciągły, zrozumienie, jak działa kontrola burz, pozwoli Ci pojąć, dlaczego może ona powodować sporadyczne przerwy w komunikacji dla standardowych aplikacji w zdrowej sieci. Przykładowo, jeśli w sieci znalazłby się wirus, który wysyła wystarczająco dużo pakietów rozgłoszeniowych, aby uruchomić kontrolę burz, istnieje możliwość, że port przestanie wysyłać jakiekolwiek pakiety rozgłoszeniowe, bo próg będzie stale przekroczony. Z drugiej strony, jeśli Twoja sieć działa poprawnie, ale typowe natężenie ruchu oscyluje wokół progu, prawdopodobnie będziesz tracił niektóre pakiety rozgłoszeniowe. Tylko uważne monitorowanie przełączników może dać Ci pewność, że nie utrudniasz normalnego ruchu w sieci. Jeśli kontrola burz powoduje problemy w sieci, która poza tym działa poprawnie, musisz prawdopodobnie poprawić ustawienia parametrów dla tej funkcji.

Konfiguracja kontroli burz odbywa się za pomocą polecenia interfejsu `storm-control`:

```
3750(config-if)#storm-control ?
  action     Action to take for storm-control //Rodzaj akcji, która ma być podjęta przez funkcję kontroli burz
  broadcast  Broadcast address storm control //Kontrola burz dla pakietów rozgłoszeniowych
  multicast  Multicast address storm control //Kontrola burz dla pakietów multicastowych
  unicast    Unicast address storm control //Kontrola burz dla pakietów unicastowych
```

Funkcja kontroli burz jest dostępna tylko dla interfejsów fizycznych. Mimo że można wpisać komendy tej funkcji dla interfejsów EtherChannel (port channel), zostaną one zignorowane.

Funkcja kontroli burz zmieniła się nieco wraz z kilkoma ostatnimi wersjami IOS. Pierwotnie komendami uruchamiającymi tę funkcję były polecenia: `switchport broadcast`, `switchport multi ↪cast` oraz `switchport unicast`. Możesz nadal znaleźć te polecenia w starszych wersjach IOS oraz w przełączniku 3550.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Dodatkowo ostatnie wersje IOS oferują kilka nowych możliwości. Jedną z nich jest możliwość wysyłania pułapki SNMP zamiast zamykania portu. Możesz to skonfigurować poleceniem interfejsu `storm-control action`:

```
3750(config-if)#storm-control action ?
  shutdown  Shutdown this interface if a storm occurs //Zamknij interfejs, jeśli wystąpi burza
  trap      Send SNMP trap if a storm occurs //Wyślij pułapkę SNMP, jeśli wystąpi burza
```

Uprzednio mogłeś definiować próg jako funkcję procentowego udziału w dostępnej szerokości pasma. Teraz dostępne są opcje konfiguracji w procentach, rzeczywistych bitach na sekundę oraz pakietach na sekundę. Dla każdego rodzaju kontroli burz (rozgłoszeniowa, multicastowa i unicastowa) można dowolnie skonfigurować jeden z dostępnych typów progu:

```
3750(config-if)#storm-control broadcast level ?
  <0.00 - 100.00>  Enter rising threshold //Zdefiniuj próg wznoszenia
  bps              Enter suppression level in bits per second //Podaj poziom tłumienia w bitach na sekundę
  pps              Enter suppression level in packets per second
                                                    //Podaj poziom tłumienia w pakietach na sekundę
```

Próg (dowolnego typu) o wartości 0 wskazuje, że nie będą przepuszczane żadne pakiety wybranego rodzaju. Procentowy próg o wartości 100 nigdy nie będzie blokował określonych pakietów.

Konfigurując progi typu bity-na-sekundę lub pakiety-na-sekundę, możesz podać wartość w standardowym zapisie dziesiętnym albo z przyrostkiem metrycznym. Dopuszczalne przyrostki to: `k` (**kilo**), `m` (**mega**) oraz `g` (**giga**):

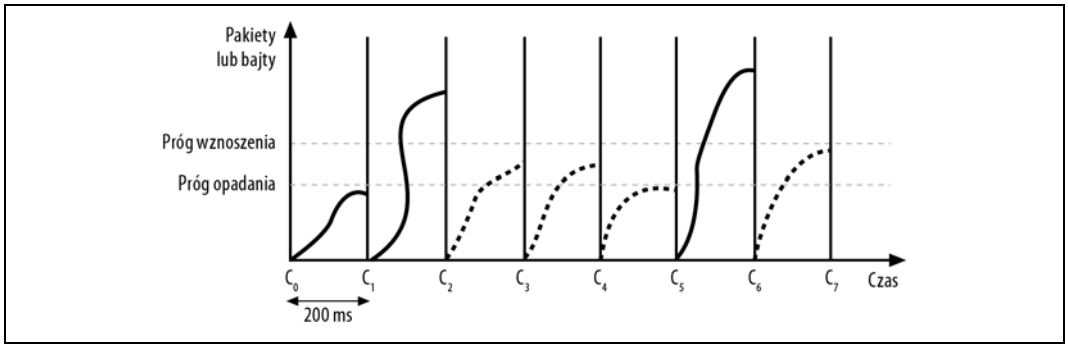
```
3750(config-if)#storm-control broadcast level bps ?
  <0.0 - 10000000000.0>[k|m|g]  Enter rising threshold
```

Kolejną nową funkcją jest możliwość definiowania progu wznoszenia (ang. *rising threshold*) oraz progu opadania (ang. *falling threshold*). Kiedy próg wznoszenia zostaje przekroczony, skonfigurowany typ pakietów jest odrzucany w następnym interwale. Jeśli liczba pakietów spadnie poniżej progu opadania, w następnym interwale ponownie przepuszczane będą pakiety określonego typu.

Rysunek 19.3 przedstawia efekt skonfigurowania progów wznoszenia i opadania. Próg wznoszenia jest ustawiony wyżej niż próg opadania. Ma to znaczący wpływ na liczbę interwałów odrzucających pakiety. Kiedy C2 przekracza próg wznoszenia, C3 odrzuca pakiety. Podobnie by się stało, jeśli byłby to jedyny skonfigurowany próg. C3 nie przekracza progu wznoszenia, ale nie spada także poniżej progu opadania — w C4 pakiety są nadal odrzucane. Od chwili przekroczenia progu wznoszenia pakiety określonego typu będą dotąd odrzucane, aż ich liczba nie spadnie poniżej progu opadania. Dzieje się tak do końca interwału C5, kiedy to poziom pakietów spada w końcu poniżej progu opadania. W następnym interwale ten typ pakietów może być ponownie przekazywany.

Próg opadania definiuje się po skonfigurowaniu progu wznoszenia. Jeśli nie podasz żadnej wartości, próg opadania będzie taki sam jak próg wznoszenia:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 19.3. Progi wznoszenia i opadania

```
3750(config-if)#storm-control broadcast level bps 100 ?
  <0.0 - 10000000000.0>[k|m|g]  Enter falling threshold
  <cr>
```

Poniżej skonfigurowałem takie same progi, stosując różne formy zapisu wartości. Oba sposoby są poprawne:

```
3750(config-if)#storm-control broadcast level bps 100000 90000
3750(config-if)# storm-control broadcast level bps 100k 90k
```

Myślę, że najprostszym sposobem konfiguracji kontroli burz są progi procentowe. Jest to również jedyna metoda obsługiwana przez starsze wersje IOS:

```
3750(config-if)#storm-control multicast level 40.5 30
```

Bądź ostrożny, konfigurując kontrolę burz multicastowych. Kiedy pakiety multicastowe są tłumione, wpływa to na protokoły routingu wykorzystujące multicasty. Natomiast nie oddziałuje to na pakiety protokołu CDP (ang. *Cisco Discovery Protocol*). Problem w tym, że jeśli działa protokół CDP, a nie działają protokoły routingu, w sieci może dojść do pewnego zamieszania w trakcie tych przerw.

Do monitorowania funkcji kontroli burz możesz wykorzystać polecenie `show storm-control`. Listing nie jest obszerny, ale prawdopodobnie nie będziesz potrzebował niczego więcej:

```
3750#sho storm-control
Interface  Filter State   Upper        Lower        Current
---------  -------------  -----------  -----------  ----------
Gi1/0/20   Link Down          1g bps     900m bps        0 bps
Gi1/0/21   Link Down          50.00%       40.00%        0.00%
Gi1/0/22   Forwarding         1m pps     500k pps        0 pps
```

Kolumna „current" zawiera bieżącą wartość dla interfejsu. Właśnie to powinieneś najpierw sprawdzić, jeśli podejrzewasz, że pakiety są zrzucane w wyniku działania kontroli burz. Pamiętaj, iż wartości są mierzone co 200 ms, więc będziesz musiał wykonać tę komendę wiele razy, żeby zobaczyć, czy pakiety nie przekraczają progu.

Możesz również użyć tej komendy z argumentem określającym rodzaj pakietów dla kontroli burz (`broadcast`, `multicast` lub `unicast`). Listing jest taki sam, ale uwzględnia tylko ten typ pakietów, który wybrałeś:

```
3750#sho storm-control unicast
Interface  Filter State   Upper        Lower        Current
---------  -------------  -----------  -----------  ----------
Gi1/0/19   Link Down       50.00%       40.00%        0.00%
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kolejną opcją jest określenie interfejsu, na temat którego chcesz uzyskać informacje. Jako argument podajesz wtedy nazwę interfejsu:

```
3750#sho storm-control g1/0/20
Interface  Filter State   Upper        Lower        Current
---------  -------------  -----------  -----------  ----------
Gi1/0/20   Link Down          1g bps     900m bps        0 bps
```

Zabezpieczenie portów

Zabezpieczenie portów (ang. *port security*) to funkcja umożliwiająca blokowanie dostępu do portu przełącznika urządzeniom sieciowym. Na poziomie portu możesz zdefiniować wybrany adres MAC, któremu możesz zabronić (ang. *deny*) dostępu do portu lub go umożliwić (ang. *allow*). Możesz to zrobić statycznie lub dynamicznie. Przykładowo możesz skonfigurować przełącznik tak, aby zezwolił tylko trzem pierwszym podłączonym stacjom na korzystanie z portu. Inną opcją jest umożliwienie korzystania z portu jedynie urządzeniu o podanym adresie MAC lub każdemu urządzeniu **z wyjątkiem** tego o podanym adresie MAC.

Adresy MAC są wprowadzane ręcznie lub pozyskiwane dynamicznie. Dynamicznie pozyskane adresy mogą być zapisywane. Adresy skonfigurowane ręcznie nazywają się **static secure MAC addresses**, podczas gdy te pozyskane dynamicznie określa się terminem **dynamic secure MAC addresses**. Zapisane adresy dynamiczne to **sticky secure MAC addresses**.

Zabezpieczenie portów włącza się poleceniem interfejsu `switchport port-security`. Może ono być zastosowane tylko na interfejsie, który został ustawiony jako switchport. Trunki oraz interfejsy, które są dynamiczne (domyślne), nie mogą być skonfigurowane z funkcją zabezpieczenia portów:

```
3750(config-if)#switchport port-security
Command rejected: GigabitEthernet1/0/20 is a dynamic port.
                                    //Komenda odrzucona: GigabitEthernet1/0/20 jest portem dynamicznym.
```

Jeśli otrzymasz ten błąd, musisz skonfigurować port w trybie dostępu switchport, zanim przejdziesz do dalszych ustawień:

```
3750(config-if)#switchport mode access
3750(config-if)# switchport port-security
```

Nie możesz ustawić zabezpieczenia portów dla interfejsu, który został skonfigurowany jako port docelowy SPAN (ang. *SPAN destination*):

```
3750(config-if)#switchport port-security
Command rejected: GigabitEthernet1/0/20 is a SPAN destination.
```

Kiedy już włączysz zabezpieczenie portów, możesz skonfigurować opcje dla tej funkcji:

```
3750(config-if)#switchport port-security ?
  aging        Port-security aging commands
  mac-address  Secure mac address
  maximum      Max secure addresses
  violation    Security violation mode
  <cr>
```

Poniżej ustawię, aby interfejs akceptował pakiety jedynie z urządzenia o adresie MAC 1234.5678.9012:

```
3750(config-if)#switchport port-security mac-address 1234.5678.9012
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Mógłbyś pomyśleć, że można zastosować tę samą komendę, aby dodać kolejny adres MAC do listy dozwolonych urządzeń. Jeśli tego spróbujesz, otrzymasz informację o błędzie:

```
3750(config-if)#switchport port-security mac-address 1234.5678.1111
Total secure mac-addresses on interface GigabitEthernet1/0/20 has reached maximum
limit. //Osiągnięto maksymalny limit adresów mac dla interfejsu GigabitEthernet1/0/20
```

Domyślnie można wprowadzić tylko jeden adres MAC. Żeby zwiększyć ten limit, zastosuj polecenie `switchport port-security maximum`. Po zwiększeniu limitu możesz zdefiniować kolejny adres MAC:

```
3750(config-if)#switchport port-security maximum 2
3750(config-if)# switchport port-security mac-address 1234.5678.1111
```

Jeśli spróbujesz ustawić maksymalną liczbę zabezpieczonych adresów MAC, która jest mniejsza od liczby już zdefiniowanych adresów, otrzymasz informację o błędzie, a komenda zostanie zignorowana.

Możesz zastosować komendę `switchport port-security maximum`, nie podając żadnych adresów MAC. W ten sposób ustawisz, aby z portu mogła korzystać ograniczona liczba adresów MAC. Przykładowo przy ustawionej liczbie maksymalnie trzech adresów tylko pierwsze trzy pozyskane dynamicznie adresy MAC będą mogły korzystać z portu, a pozostałe będą odrzucane.

Jeśli chcesz, aby przełącznik wykrywał i zapisywał adresy MAC, dodaj do komendy argument `sticky`. Takie adresy są dodawane do działającej konfiguracji:

```
3750(config-if)#switchport port-security mac-address sticky
```

Żeby pozyskane adresy zostały zachowane, musisz skopiować działającą konfigurację do konfiguracji startowej (skorzystaj z komendy `write memory`) przed zrestartowaniem przełącznika.

Jeśli dla portu skonfigurowana jest funkcja zabezpieczenia, a przychodzące pakiety nie spełniają zdefiniowanych warunków, uznane jest to za naruszenie. Przy wykryciu naruszenia bezpieczeństwa portów przełącznik może podjąć jedną z trzech możliwych akcji:

Ochrona (ang. *protect*)
: Jeśli pojawi się naruszenie, przełącznik będzie odrzucał wszystkie pakiety od adresów MAC, które nie spełniają skonfigurowanych wymagań. Takiej sytuacji nie towarzyszą żadne powiadomienia.

Ograniczenie (ang. *restrict*)
: Jeśli pojawi się naruszenie, przełącznik będzie odrzucał wszystkie pakiety od adresów MAC, które nie spełniają skonfigurowanych wymagań. Generowana jest pułapka SNMP, dodawany jest wpis w dzienniku logów oraz wzrasta wartość licznika naruszeń.

Zamknięcie (ang. *shutdown*)
: Jeśli pojawi się naruszenie, przełącznik wyłączy port, ustawiając go w stan **error-disabled**. Zostanie zatrzymany cały ruch przychodzący i wychodzący z portu. Jest to domyślne zachowanie dla portów z włączoną funkcją zabezpieczenia portu. Aby wyjść z tego stanu, musisz zresetować interfejs komendami `shutdown` lub `no shutdown` albo komendą `errdisable reco ↪very cause psecure-violation`.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Do zmiany typu akcji podejmowanej przez przełącznik w przypadku naruszenia posłuży Ci polecenie `switchport port-security violation`:

```
3750(config-if)#switchport port-security violation ?
  protect   Security violation protect mode
  restrict  Security violation restrict mode
  shutdown  Security violation shutdown mode
```

Bezpieczne adresy MAC mogą wygasać po ogólnie zdefiniowanym czasie lub po określonym czasie nieaktywności. Ta druga opcja może być użyteczna w środowiskach dynamicznych, w których urządzenia są nieustannie podłączane i odłączane. Załóżmy, że masz salę, w której pracuje wielu konsultantów. Pierwsi trzej, którzy podłączą się rano, będą mogli korzystać z sieci. Reszta ma mniej szczęścia. Jeśli jeden z tych trzech opuści biuro wcześniej, możesz chcieć zwolnić jego miejsce, aby umożliwić komuś innemu dostęp do sieci. Alternatywnie, możesz zdecydować się zapewnić pierwszym trzem konsultantom dostęp do sieci przez cały dzień, bez względu na to, o której opuszczą biuro. Dla pozostałych konsultantów będzie to kara za spóźnienie. Pracowałem z przełożonymi, którzy byli szczęśliwi, mogąc zastosować takie rozwiązania do przymuszenia konsultantów, by przychodzili do pracy wcześniej! Aby skonfigurować sposób „starzenia się" (ang. *aging*) adresów MAC, użyj polecenia `switchport port-security aging type`:

```
3750(config-if)#switchport port-security aging type ?
  absolute    Absolute aging (default)
                                //Adresy MAC będą wygasać po ogólnie zdefiniowanym czasie (ustawienie domyślne)
  inactivity  Aging based on inactivity time period
                                //Adresy MAC będą wygasać po określonym czasie nieaktywności
```

Czas starzenia się podawany jest w minutach wraz z argumentem `time`:

```
3750(config-if)#switchport port-security aging time ?
  <1-1440>  Aging time in minutes. Enter a value between 1 and 1440
                                //Podaj czas starzenia się w minutach. Wprowadź wartość z przedziału 1 – 1440
3750(config-if)# switchport port-security aging time 30
```

Do sprawdzenia statusu funkcji zabezpieczenia portów służy polecenie `show port-security`. Pokazuje ono miłe podsumowanie dla wszystkich portów z włączoną funkcją zabezpieczenia, liczbę skonfigurowanych adresów dla poszczególnych portów, liczbę adresów wykrytych dynamicznie oraz informuje, ile wystąpiło naruszeń:

```
3750#sho port-security
Secure Port  MaxSecureAddr  CurrentAddr  SecurityViolation  Security Action
                (Count)       (Count)          (Count)
--------------------------------------------------------------------------
   Gi1/0/20              2            2                  0         Shutdown
--------------------------------------------------------------------------
Total Addresses in System (excluding one mac per port)     : 1
Max Addresses limit in System (excluding one mac per port) : 6272
```

Aby uzyskać więcej szczegółów, zastosuj komendę `show port-security interface` dla określonego interfejsu:

```
3750#sho port-security interface g1/0/20
Port Security              : Enabled
Port Status                : Secure-down
Violation Mode             : Shutdown
Aging Time                 : 0 mins
Aging Type                 : Absolute
SecureStatic Address Aging : Disabled
Maximum MAC Addresses      : 2
Total MAC Addresses        : 2
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Configured MAC Addresses   : 2
Sticky MAC Addresses       : 0
Last Source Address:Vlan   : 0000.0000.0000:0
Security Violation Count   : 0
```

SPAN

SPAN (ang. *Switched Port Analyzer*), czyli analizator przełączanego portu, to funkcja umożliwiająca zreplikowanie ruchu sieciowego z określonego źródła na wybrany port. Replikować można ruch z portu fizycznego, portu wirtualnego lub VLAN-ów. Nie można jednak mieszać typów źródeł dla pojedynczej sesji SPAN. Najbardziej powszechnym zastosowaniem SPAN jest przechwytywanie pakietów. Przykładowo, jeśli chcesz przechwycić pakiety przepływające przez VLAN 10, nie możesz zastosować sniffera na jednym z portów tego VLAN-u, ponieważ przełącznik będzie przekazywał jedynie pakiety przeznaczone dla sniffera. Jednak uruchomienie SPAN z VLAN-em jako źródłem oraz portem sniffera jako miejscem docelowym spowoduje przekazywanie do sniffera całego ruchu na VLAN-ie. Funkcja SPAN jest również często stosowana, jeśli w sieci zostały zaimplementowane systemy detekcji włamań (ang. *Intrusion Detection Systems* — IDS). Urządzenia IDS muszą czytać wszystkie pakiety z jednego lub większej liczby VLAN-ów, a SPAN umożliwia dostarczenie pakietów do urządzeń IDS.

Wykorzystując **Remote Switched Port Analyzer** (RSPAN), czyli zdalny SPAN, możesz nawet wysyłać pakiety do innego przełącznika. RSPAN może być użyteczny w centrach danych, gdzie urządzenie do przechwytywania pakietów jest na stałe zainstalowane na jednym z wielu połączonych ze sobą przełączników. Dzięki RSPAN możesz przechwytywać pakiety na przełącznikach innych niż ten z podłączonym snifferem (szczegóły konfiguracji RSPAN znajdziesz w dalszej części tego podrozdziału).

SPAN konfigurujesz poleceniem `monitor`. Możesz posiadać więcej niż jedną sesję SPAN. Każdą z nich identyfikuje numer sesji:

```
3750(config)#monitor session 1 ?
  destination  SPAN destination interface or VLAN //Docelowy interfejs lub VLAN dla funkcji SPAN
  filter       SPAN filter //Filtr SPAN
  source       SPAN source interface, VLAN //Źródłowy interfejs lub VLAN dla funkcji SPAN
```

Uruchomienie kilku sesji SPAN jest użyteczne, kiedy posiadasz w sieci urządzenie IDS oraz musisz przechwytywać pakiety. Urządzenie IDS będzie wymagało jednej sesji SPAN, a przechwytywanie pakietów drugiej.

Aby sesja `monitor` była aktywna, musisz skonfigurować źródłowy port lub VLAN oraz port docelowy. Osobiście konfiguruję zazwyczaj najpierw port docelowy, ponieważ urządzenie do przechwytywania pakietów jest już podłączone. Jeśli dany port ma włączoną funkcję zabezpieczenia portu, musisz ją wyłączyć, zanim będziesz mógł użyć go jako portu docelowego SPAN:

```
3750(config)#monitor session 1 destination interface g1/0/20
%Secure port can not be dst span port //Port zabezpieczony nie może być portem docelowym SPAN
```

Sesje mogą być ponumerowane od 1 do 66. Możesz posiadać jednak tylko dwie sesje skonfigurowane w danym momencie na przełączniku 3750. Poniżej skonfigurowałem dwie sesje (sesję 1 oraz sesję 10):

```
monitor session 1 source vlan 20 rx
monitor session 1 destination interface Gi1/0/10
!
monitor session 10 source vlan 10 rx
monitor session 10 destination interface Gi1/0/20
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli spróbujesz skonfigurować więcej niż dwie sesje SPAN na przełączniku 3750, wystąpi poniższy błąd:

```
3750(config)#monitor session 20 source int g1/0/10
% Platform can support a maximum of 2 source sessions //Platforma obsługuje maksymalnie dwie sesje źródłowe
```

W tym przykładzie skonfigurowałem dwa VLAN-y jako źródła. W obu przypadkach pakiety będą replikowane na interfejs `Gi1/0/20`:

```
monitor session 10 source vlan 20 rx
monitor session 10 source vlan 10
monitor session 10 destination interface Gi1/0/20
```

Monitorować możesz również jeden lub kilka interfejsów. W przypadku wprowadzania kilku interfejsów możesz je skonfigurować osobno lub w jednej linii poleceń:

```
3750(config)#monitor session 11 source interface g1/0/11
3750(config)# monitor session 11 source interface g1/0/12
```

Wprowadzenie powyższych komend spowoduje dodanie w konfiguracji następującej linii:

```
monitor session 11 source interface Gi1/0/11 - 12
```

Źródła w sesji monitorowania mogą być skonfigurowane jako: odbiorcze (`rx`), nadawcze (`tx`) lub odbiorczo-nadawcze (`both`). Domyślną konfiguracją jest `both`:

```
3750(config)#monitor session 1 source int g1/0/12 ?
  ,     Specify another range of interfaces //Oddziela definiowane zakresy interfejsów
  -     Specify a range of interfaces //Stosowany przy określaniu zakresu interfejsów
  both  Monitor received and transmitted traffic //Monitorowanie pakietów odbieranych i nadawanych
  rx    Monitor received traffic only //Monitorowanie tylko pakietów odbieranych
  tx    Monitor transmitted traffic only //Monitorowanie tylko pakietów nadawanych
  <cr>
```

Interfejsy powinny być z reguły monitorowane w obu kierunkach, podczas gdy VLAN-y tylko w jednym.

Kiedy przechwytujesz informacje z VLAN-ów, uważaj na zduplikowane pakiety. Pamiętaj, że w przypadku VLAN-u każdy pakiet przychodzi na jeden port i wychodzi na innym porcie. Przy domyślnym ustawieniu `both` dla monitorowania VLAN-u prawie każdy pakiet będzie duplikowany w procesie przechwytywania.

Nie wiem już, ile razy byłem przekonany, że odkryłem niepożądane urządzenie dupli-kujące pakiety w sieci, ale okazywało się, że znów sam się podłożyłem, monitorując VLAN w obu kierunkach. W przypadku VLAN-ów najbezpieczniej jest monitorować je tylko w kierunku `rx`. Ponieważ domyślnym ustawieniem jest `both`, mogę sobie tłumaczyć, że padłem ofiarą żartownisia z Cisco, zamiast uświadomić sobie, iż jestem kompletnym idiotą.

Żeby sprawdzić, jakie sesje są skonfigurowane lub aktywne, skorzystaj z polecenia `show monitor`:

```
3750#sho monitor
Session 1
---------
Type              : Local Session
Source VLANs      :
    RX Only       : 20
Destination Ports : Gi1/0/22
    Encapsulation : Native
          Ingress : Disabled
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Session 10
----------
Type              : Local Session
Source VLANs      :
    TX Only       : 20
    Both          : 10
Destination Ports : Gi1/0/20
    Encapsulation : Native
          Ingress : Disabled
```

Jeśli chcesz wyłączyć monitorowanie dla określonego SPAN-u, możesz usunąć całą sesję, wszystkie źródła lub miejsce docelowe. Dla wszystkich komend `monitor` możliwe jest stosowanie negacji:

```
3750(config)#no monitor session 11 source interface Gi1/0/11 - 12
```

Stosując argument `local`, `remote` lub `all`, możesz wykasować wszystkie sesje lokalne SPAN, wszystkie sesje RSPAN albo wszystkie sesje SPAN:

```
3750(config)#no monitor session ?
  <1-66>  SPAN session number
  all     Remove all SPAN sessions in the box
  local   Remove Local SPAN sessions in the box
  remote  Remove Remote SPAN sessions in the box
```

Zawsze powinieneś usuwać te sesje SPAN, których już nie potrzebujesz, gdyż SPAN wykorzystuje zasoby systemowe. Może też powstać zamieszanie, jeśli ktoś podepnie jakieś urządzenie do portu docelowego SPAN.

RSPAN działa w podobny sposób co SPAN, z tym wyjątkiem, że interfejs docelowy znajduje się na innym przełączniku. Oba przełączniki muszą być połączone **VLAN-em RSPAN**. Aby stworzyć VLAN RSPAN, skonfiguruj VLAN i dodaj komendę `remote-span`:

```
3750-1(config)#vlan 777
3750-1(config-vlan)# remote-span
```

Jeśli wykorzystujesz protokół VTP, nie zawsze będziesz musiał tworzyć nowy VLAN, ale wciąż będziesz musiał skonfigurować VLAN dla RSPAN. W obu przypadkach należy podjąć te same kroki. Najpierw zdefiniuj na przełączniku źródłowym miejsce docelowe jako RSPAN VLAN:

```
3750-1(config)#monitor session 11 destination remote vlan 777
```

Możesz podać docelowy VLAN, który nie został skonfigurowany jako RSPAN VLAN, ale niestety nie będzie to działać.

Teraz na przełączniku docelowym skonfiguruj ten sam VLAN jako RSPAN VLAN. Następnie zdefiniuj sesję monitorowania, aby otrzymywać RSPAN wysyłane z przełącznika źródłowego:

```
3750-2(config)#vlan 777
3750-2(config-vlan)# remote-span
3750-2(config)# monitor session 11 source remote vlan 777
```

Nie ma wymagania, aby numery sesji monitorowania były takie same. Jednak jak zawsze powtarzam, **prostota jest dobra**. Jeśli nie skonfigurowałeś źródłowego przełącznika tak, aby był źródłem RSPAN, wystąpi błąd:

```
3750-2(config)#monitor session 11 source remote vlan 777
% Cannot add RSPAN VLAN as source for SPAN session 11 as it is not a RSPAN
Destination session
```

//Nie można dodać RSPAN VLAN jako źródła dla sesji 11 SPAN, ponieważ nie jest to sesja docelowa RSPAN

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli stosujesz RSPAN, nie wykorzystuj istniejącego trunku dla VLAN-u RSPAN. SPAN może generować spory ruch sieciowy. Kiedy monitorujesz VLAN-y składające się z wielu interfejsów gigabitowych, pakiety SPAN mogą z łatwością „zdusić" pojedynczy link gigabitowy RSPAN. Jeśli to możliwe, skonfiguruj pomiędzy przełącznikami dedykowany link VLAN-u RSPAN.

VLAN głosowy

VLAN głosowy (ang. *voice VLAN*) umożliwia skonfigurowanie telefonu IP Cisco, który jest podłączony do przełącznika 3750. Przełącznik wykorzystuje protokół CDP do przekazywania do telefonu informacji dotyczących CoS (ang. *Class of Service*) oraz VLAN-ów, które mają być wykorzystane do transmisji głosu i danych. Domyślnie funkcja ta jest wyłączona, co powoduje, że telefon nie otrzymuje instrukcji konfiguracyjnych od przełącznika. W takim przypadku telefon transmituje głos i dane przez VLAN domyślny (VLAN 0 w telefonie).

Telefony IP Cisco, takie jak model 7960, posiadają wbudowany 3-portowy przełącznik. Port 1 służy do podłączenia telefonu do przełącznika w sieci (tutaj zajmiemy się konfiguracją przełącznika 3750). Port 2 jest wewnętrznym połączeniem do samego telefonu. Port 3 to zewnętrzny port, służący zazwyczaj do połączenia z komputerem użytkownika.

Stosując komendę interfejsu `switchport voice vlan`, możesz skonfigurować przełącznik tak, aby umożliwić konfigurację telefonu IP podłączonego do określonego interfejsu. Możesz sam zdefiniować, który VLAN ma być wykorzystywany dla połączeń głosowych z telefonu. Alternatywnie możesz pozwolić, by przełącznik automatycznie poinstruował telefon, żeby wykorzystywał do połączeń głosowych standardowy VLAN używany do transmisji danych (ze skonfigurowanymi, lub nie, wartościami CoS):

```
3750(config-if)#switchport voice vlan ?
  <1-4094>  Vlan for voice traffic
  dot1p     Priority tagged on PVID
  none      Don't tell telephone about voice vlan
  untagged  Untagged on PVID
```

Aby wybrać VLAN, określ jego numer. Opcja `dot1p` instruuje telefon, żeby ustawiał bity CoS w pakietach głosowych, jeśli wykorzystywany jest VLAN danych. Opcja `untagged` powoduje wykorzystanie VLAN-u danych bez ustawiania wartości CoS.

Jeśli chcesz wykorzystać zalety VLAN-ów głosowych, musisz skonfigurować przełącznik tak, aby uznawał za zaufane wartości CoS, które wysyła telefon. Służy do tego polecenie interfejsu `mls qos trust cos`.

Zastosowanie komendy interfejsu `mls qos trust cos` nie przyniesie żadnego efektu, dopóki nie uruchomisz globalnie funkcji QoS (ang. *quality of service*) poleceniem `mls qos`.

Poniżej przykład interfejsu skonfigurowanego tak, aby wykorzystywał VLAN 100 do transmisji danych oraz VLAN 10 do transmisji głosowych. Przełącznik będzie ufał wartościom CoS wysyłanym przez telefon, a także poinstruuje telefon, żeby korzystał z VLAN-u 100 dla transmisji głosowych:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
interface GigabitEthernet1/0/20
 switchport access vlan 100
 switchport voice vlan 10
 mls qos trust cos
```

Kolejnym przyjemnym aspektem funkcji VLAN-u głosowego, jest możliwość zapewnienia tego, by telefon ufał wartościom CoS, które zostały ustawione przez urządzenie podłączone do jego portu zewnętrznego (zazwyczaj będzie to komputer użytkownika), lub je zmieniał. Tę funkcję konfigurujesz komendą interfejsu `switchport priority extend`. Dostępnymi opcjami są `cos` oraz `trust`. Wybierając opcję `cos`, możesz zdefiniować w polu CoS dowolną wartość:

```
3750(config-if)#switchport priority extend ?
  cos    Override 802.1p priority of devices on appliance //Nadpisuje priorytet 802.1p urządzenia
  trust  Trust 802.1p priorities of devices on appliance //Zaufaj priorytetom 802.1p urządzenia
```

Preferuję ufać wartościom CoS podłączonego komputera, ponieważ mogą one być różne w zależności od zastosowanego oprogramowania. Przykładowo użytkownik może korzystać z aplikacji softphone na swoim komputerze. Nadpisywanie wartości CoS ustawionych przez to oprogramowanie może prowadzić do problemów z jakością transmisji głosowych dla softphone.

Skonfigurowałem interfejs wykorzystujący VLAN 10 dla transmisji głosowych oraz ufający wartościom CoS ustawionym przez komputer użytkownika i telefon:

```
interface GigabitEthernet1/0/20
 switchport access vlan 100
 switchport voice vlan 10
 switchport priority extend trust
 mls qos trust cos
```

Aby sprawdzić, który VLAN został skonfigurowany jako VLAN głosowy, wpisz polecenie `show interface` *`interfejs-nazwa`* `switchport`:

```
3750#sho int g1/0/20 switchport
Name: Gi1/0/20
Switchport: Enabled
Administrative Mode: static access
Operational Mode: down
Administrative Trunking Encapsulation: negotiate
Negotiation of Trunking: Off
Access Mode VLAN: 1 (default)
Trunking Native Mode VLAN: 1 (default)
Administrative Native VLAN tagging: enabled
Voice VLAN: 10 (Inactive)
Administrative private-vlan host-association: none
Administrative private-vlan mapping: none
Administrative private-vlan trunk native VLAN: none
Administrative private-vlan trunk Native VLAN tagging: enabled
Administrative private-vlan trunk encapsulation: dot1q
Administrative private-vlan trunk normal VLANs: none
Administrative private-vlan trunk private VLANs: none
Operational private-vlan: none
Trunking VLANs Enabled: ALL
Pruning VLANs Enabled: 2-1001
Capture Mode Disabled
Capture VLANs Allowed: ALL

Protected: false
Unknown unicast blocked: disabled
Unknown multicast blocked: disabled
Appliance trust: none
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

QoS

Funkcja QoS została opisana szczegółowo w rozdziale 31., a jest to temat, który z łatwością zapełniłby kartki całej książki. W tym podrozdziale skupię się na niektórych funkcjach QoS charakterystycznych dla przełącznika 3750.

Jedną z użytecznych cech 3750 jest obsługa **AutoQoS**. AutoQoS przyjmuje pewne założenia dotyczące sieci i konfiguruje odpowiednio przełącznik. Choć nie jestem fanem umożliwiania urządzeniom kierowania się „własnymi" założeniami, to tym razem są one właściwe w większości przypadków. Nigdy nie miałem żadnych skrupułów, włączając AutoQoS w przełącznikach 3750, które zainstalowałem w sieciach VoIP posiadających setki telefonów obsługiwanych przez Cisco Call Manager. Nie ma z tym problemu, ponieważ założenia Cisco są zbudowane wokół koncepcji korzystania z oprogramowania Call Manager, telefonów IP Cisco oraz stosowania dla sieci kolejkowania z niskim opóźnieniem. Są duże szanse, że potrzebujesz włączyć QoS na przełącznikach właśnie dlatego, że implementujesz VoIP.

Możesz włączyć AutoQos dla interfejsu, stosując polecenie `auto qos voip`:

```
3750(config-if)#auto qos voip ?
  cisco-phone      Trust the QoS marking of Cisco IP Phone
  cisco-softphone  Trust the QoS marking of Cisco IP SoftPhone
  trust            Trust the DSCP/CoS marking
```

Dostępne są trzy opcje: `cisco-phone`, `cisco-softphone` oraz `trust`. Pierwsze dwie stosowane są dla interfejsów podłączonych odpowiednio do telefonów sprzętowych lub programowych. Przy zastosowaniu tych opcji wartości QoS otrzymywane w pakietach będą zaufane tylko wtedy, kiedy ich źródłem będą telefony IP Cisco. Opcja `trust` powoduje zaufanie wartościom QoS wszystkich pakietów.

Jeśli chciałbyś zobaczyć, za co odpowiada AutoQos, włącz debugowanie AutoQoS poleceniem `debug auto qos`, zanim skonfigurujesz interfejs:

```
3750#debug auto qos
3750# conf t
Enter configuration commands, one per line. End with CNTL/Z.
3750(config)# int g1/0/20
3750(config-if)# auto qos voip cisco-phone
3750(config-if)#
3d04h: mls qos map cos-dscp 0 8 16 26 32 46 48 56
3d04h: mls qos
3d04h: no mls qos srr-queue input cos-map
3d04h: no mls qos srr-queue output cos-map
3d04h: mls qos srr-queue input cos-map queue 1 threshold 3  0
3d04h: mls qos srr-queue input cos-map queue 1 threshold 2  1
3d04h: mls qos srr-queue input cos-map queue 2 threshold 1  2
3d04h: mls qos srr-queue input cos-map queue 2 threshold 2  4 6 7
3d04h: mls qos srr-queue input cos-map queue 2 threshold 3  3 5
3d04h: mls qos srr-queue output cos-map queue 1 threshold 3  5
3d04h: mls qos srr-queue output cos-map queue 2 threshold 3  3 6 7
3d04h: mls qos srr-queue output cos-map queue 3 threshold 3  2 4
3d04h: mls qos srr-queue output cos-map queue 4 threshold 2  1
3d04h: mls qos srr-queue output cos-map queue 4 threshold 3  0
[duży fragment usunięty]
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Konfiguracja interfejsu będzie wyglądała następująco:

```
interface GigabitEthernet1/0/20
 srr-queue bandwidth share 10 10 60 20
 srr-queue bandwidth shape  10  0  0  0
 queue-set 2
 auto qos voip cisco-phone
```

Zmiany w konfiguracji globalnej przełącznika są nieco bardziej rozległe. Na szczęście AutoQoS wykona całą robotę za Ciebie:

```
mls qos map cos-dscp 0 8 16 26 32 46 48 56
mls qos srr-queue input bandwidth 90 10
mls qos srr-queue input threshold 1 8 16
mls qos srr-queue input threshold 2 34 66
mls qos srr-queue input buffers 67 33
mls qos srr-queue input cos-map queue 1 threshold 2  1
mls qos srr-queue input cos-map queue 1 threshold 3  0
mls qos srr-queue input cos-map queue 2 threshold 1  2
mls qos srr-queue input cos-map queue 2 threshold 2  4 6 7
mls qos srr-queue input cos-map queue 2 threshold 3  3 5
mls qos srr-queue input dscp-map queue 1 threshold 2  9 10 11 12 13 14 15
mls qos srr-queue input dscp-map queue 1 threshold 3  0 1 2 3 4 5 6 7
mls qos srr-queue input dscp-map queue 1 threshold 3  32
mls qos srr-queue input dscp-map queue 2 threshold 1  16 17 18 19 20 21 22 23
mls qos srr-queue input dscp-map queue 2 threshold 2  33 34 35 36 37 38 39 48
mls qos srr-queue input dscp-map queue 2 threshold 2  49 50 51 52 53 54 55 56
mls qos srr-queue input dscp-map queue 2 threshold 2  57 58 59 60 61 62 63
mls qos srr-queue input dscp-map queue 2 threshold 3  24 25 26 27 28 29 30 31
mls qos srr-queue input dscp-map queue 2 threshold 3  40 41 42 43 44 45 46 47
mls qos srr-queue output cos-map queue 1 threshold 3  5
mls qos srr-queue output cos-map queue 2 threshold 3  3 6 7
mls qos srr-queue output cos-map queue 3 threshold 3  2 4
mls qos srr-queue output cos-map queue 4 threshold 2  1
mls qos srr-queue output cos-map queue 4 threshold 3  0
mls qos srr-queue output dscp-map queue 1 threshold 3  40 41 42 43 44 45 46 47
mls qos srr-queue output dscp-map queue 2 threshold 3  24 25 26 27 28 29 30 31
mls qos srr-queue output dscp-map queue 2 threshold 3  48 49 50 51 52 53 54 55
mls qos srr-queue output dscp-map queue 2 threshold 3  56 57 58 59 60 61 62 63
mls qos srr-queue output dscp-map queue 3 threshold 3  16 17 18 19 20 21 22 23
mls qos srr-queue output dscp-map queue 3 threshold 3  32 33 34 35 36 37 38 39
mls qos srr-queue output dscp-map queue 4 threshold 1  8
mls qos srr-queue output dscp-map queue 4 threshold 2  9 10 11 12 13 14 15
mls qos srr-queue output dscp-map queue 4 threshold 3  0 1 2 3 4 5 6 7
mls qos queue-set output 1 threshold 1 138 138 92 138
mls qos queue-set output 1 threshold 2 138 138 92 400
mls qos queue-set output 1 threshold 3 36 77 100 318
mls qos queue-set output 1 threshold 4 20 50 67 400
mls qos queue-set output 2 threshold 1 149 149 100 149
mls qos queue-set output 2 threshold 2 118 118 100 235
mls qos queue-set output 2 threshold 3 41 68 100 272
mls qos queue-set output 2 threshold 4 42 72 100 242
mls qos queue-set output 1 buffers 10 10 26 54
mls qos queue-set output 2 buffers 16 6 17 61
mls qos
```

Jeśli zajrzysz do czyjegoś routera i zobaczysz te wszystkie elementy, to nie powinieneś raczej uznawać osoby, która go konfiguruje, za jakiegoś geniusza QoS. Najprawdopodobniej uruchomiła ona właśnie AutoQoS!

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Do sprawdzenia, na których interfejsach został włączony AutoQoS, służy komenda globalna `show auto qos`:

```
3750#show auto qos
GigabitEthernet1/0/20
auto qos voip cisco-phone
```

Aby wyłączyć AutoQos na interfejsie, użyj komendy interfejsu `no auto qos voip`. Jeśli chcesz wyłączyć AutoQoS globalnie, wpisz polecenie `no mls qos`. Bądź ostrożny, gdyż wyłącza ono **cały** QoS na przełączniku.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 20.

Nomenklatura telekomunikacyjna

Świat telekomunikacji różni się nieco od świata informatyki, o czym z pewnością poinformuje Cię rzesza inżynierów telekomunikacji. Przykładowo większa część obecnej infrastruktury telekomunikacyjnej opiera się na standardach ustanowionych ponad sto lat temu. Samuel Morse wynalazł telegraf w 1835 r. Alexander Graham Bell skonstruował telefon w 1876 r. W roku 1961 laboratoria Bell Labs opracowały standard T1[1] służący do agregacji łączy pomiędzy biurami centralnymi (ang. *central offices* — CO) kompanii telefonicznych. Potrzeba było prawie stu lat od powstania pierwszego telefonu do wynalezienia T1.

Nieco inaczej rzecz ma się ze światem danych: ARPANET wystartował w 1969 r., Robert Metcalfe i David Boggs zbudowali pierwszy Ethernet w roku 1973, a Vincent Cerf i Bob Kahn opublikowali oryginalny standard TCP/IP w 1974 r. Hayes wprowadził pierwszy modem w 1977 r. (rety, aż 300 b/s!), a 3Com dostarczył pierwszą kartę Ethernet 10 Mb/s w 1981 r. Pierwszy router do zastosowań komercyjnych został sprzedany w 1983 r.

Zastanówmy się nad tym przez chwilę — pierwszy router został sprzedany w roku 1983. Spróbuj zapytać znajomych, czy pamiętają jeszcze czasy, w których nie było telefonów.

Świat telekomunikacji został zbudowany na podstawie standardów, które są już w użyciu od bardzo dawna. Jak często Twój telefon przestaje działać? Infrastruktura telekomunikacyjna jest na tyle niezawodna, że przyzwyczailiśmy się do solidności usług telefonicznych nawet bardziej niż do solidności dostawców elektryczności (telefonia komórkowa to zupełnie inna para kaloszy i nie ma zastosowania w naszych rozważaniach).

Jak w przypadku każdej technologii inżynierowie w okopach (a ich szefowie zza biurek) lubią sobie „rzucić" slangiem. Jeśli przez całe życie zawodowe otaczałeś się sprzętem informatycznym, żargon telekomunikacyjny może brzmieć dla Ciebie całkiem obco. Aby zbliżyć nieco obie te dziedziny, przygotowałem listę pojęć, z którymi możesz się spotkać, mając do czynienia z technologiami telekomunikacyjnymi.

Znaczenie wielu szeroko stosowanych terminów z dziedziny telekomunikacji zmieniło się z biegiem lat. Kolejne regulacje prawne zmodyfikowały funkcje takich podmiotów jak IXC oraz LEC (oba terminy zdefiniowane są w poniższym słowniku). Opiszę tutaj pierwotne znaczenie wybranych terminów.

[1] Jest to standard stosowany w Ameryce Północnej; jego europejskim odpowiednikiem jest sieć E1 — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Większość stosowanych w telekomunikacji określeń i sformułowań ma standardowe znaczenie, zdefiniowane w publikacji Federal Standard 1037C, zatytułowanej „Telecommunications: Glossary of Telecommunication Terms". Definicje te są zazwyczaj dość proste i nie zawierają zbyt wielu szczegółów. Pojęcia, które omówię w tym rozdziale, należą do najczęściej spotykanych w pracy administratora sieci lub inżyniera. Jeśli chcesz wiedzieć, czym jest występujący w sieci ważonej CCIF poziom szumów napięcia w obwodach mierzony za pomocą psofometru, z pewnością powinieneś zajrzeć do publikacji 1037C. Innym doskonałym źródłem informacji, które powinno się znaleźć w bibliotece każdego, kto ma do czynienia z telekomunikacją, jest *Newton's Telecom Dictionary* Harry'ego Newtona (CMP Books).

Słownik terminów telekomunikacyjnych

ACD

ACD (ang. *Automatic Call Distributor*) to automatyczny dystrybutor połączeń. Jest to system spotykany zazwyczaj w centrach obsługi telefonicznej typu *call center*, gdzie przychodzące rozmowy muszą być kierowane do następnego wolnego operatora lub kolejkowane do czasu, aż zwolni się któryś z operatorów.

Add/drop

Termin *add/drop* jest stosowany w telekomunikacji do opisania funkcjonalności polegającej na „odrywaniu" kanałów z obwodu i rutowaniu ich gdzie indziej. *Add/drop* CSU/DSU umożliwia wyodrębnienie zakresu kanałów do wydzielonych portów danych, dzięki czemu T1 może być podzielona na dwie częściowe sieci T1. Jedna z nich może być wykorzystana dla transmisji głosowych, a druga dla transmisji danych. Obie z nich mogą być też wykorzystane dla jednej z tych funkcji, ale z zastosowaniem innego routingu. Możesz wyłączyć funkcję *add/drop* w danym urządzeniu, przydzielając wszystkie kanały do pojedynczego portu danych. Funkcjonalność *add/drop* zwiększa koszt urządzenia, powinna być więc stosowana tylko wtedy, gdy jest niezbędna.

Analogowy i cyfrowy

Chcesz się trochę rozerwać? Poproś kogoś, kto zajmuje się komputerami, by zdefiniował pojęcie „analogowy". Niektóre z odpowiedzi z pewnością Cię zadziwią.

Analogowy znaczy dosłownie: **taki sam**. Jeśli jeden przedmiot jest analogiczny do drugiego, to jest taki jak ten drugi przedmiot. W telekomunikacji i informatyce termin „analogowy" odnosi się do sygnału, który jest ciągły w amplitudzie i czasie.

Sygnał analogowy nie składa się z wartości dyskretnych: każda drobna fluktuacja sygnału jest istotna. Fale radiowe są analogowe podobnie jak fale sygnału zasilania. Dźwięk jest również analogowy. Kiedy mówisz, wytwarzają się fale powietrza, które uderzają w bębenki ludzkiego ucha. Fala dźwiękowa jest sygnałem analogowym.

Pojęcie „cyfrowy" odnosi się do sygnału, który posiada wartości dyskretne. Jeśli przeanalizujesz falę dźwiękową, a następnie przypiszesz każdej próbce fali pewne wartości w określonych interwałach czasowych, stworzysz cyfrową reprezentację fali analogowej.

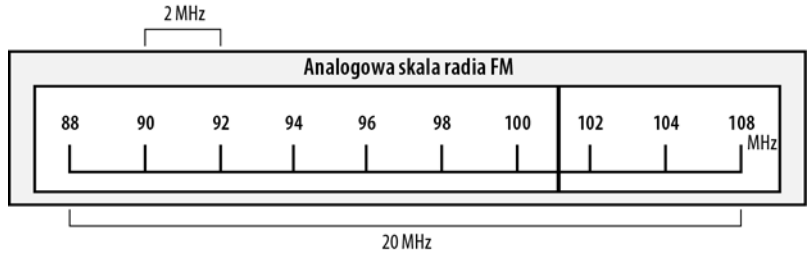
Ponieważ sygnał cyfrowy posiada wartości dyskretne, a sygnał analogowy nie, konwersja sygnału analogowego na cyfrowy zawsze skutkuje utratą informacji.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Podnosząc częstotliwość próbkowania sygnału (między innymi), można zwiększyć jakość końcowej reprodukcji. Jednak z technicznego punktu widzenia sygnał nie może być odtworzony dokładnie w ten sam sposób.

Szerokość pasma

Szerokość pasma (ang. *bandwidth*) jest jednym z tych terminów, które są powszechnie stosowane przez osoby nieznające ich właściwego znaczenia. Zakres częstotliwości zwany jest **pasmem** (ang. *band*). Szerokość, którą to pasmo obejmuje, nazywa się **szerokością pasma**. Dla tych, którzy nie pamiętają, jak wyglądało standardowe radio FM ze skalą analogową, zamieszczam rysunek 20.1.



Rysunek 20.1. Właściwy przykład szerokości pasma

Przedstawiona skala obejmuje zakres częstotliwości dopuszczonych przez rząd USA dla nadawców transmisji radiowych. Dokładny zakres częstotliwości to 87,9 MHz – – 107,9 MHz. Szerokość pasma dla tego zakresu wynosi 20 MHz. Dla zakresu 90 MHz – – 92 MHz szerokość pasma to 2 MHz.

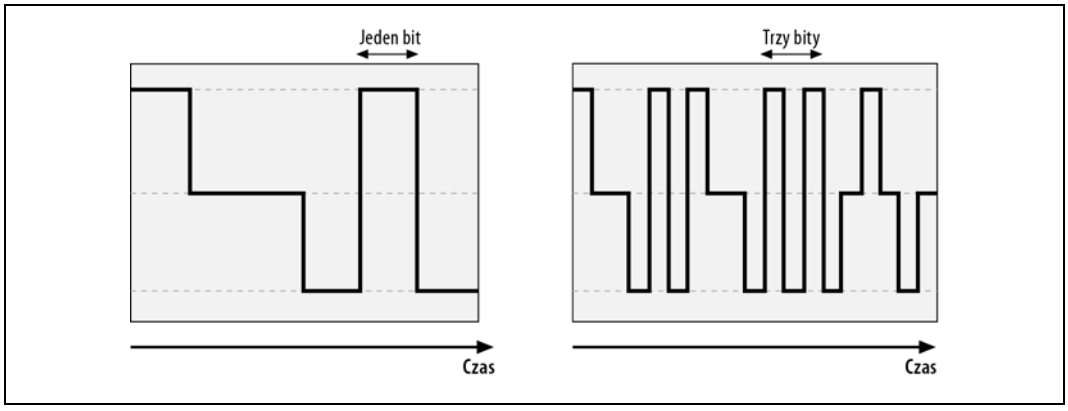
Kiedy mówimy o szerokości pasma dla łączy cyfrowych, odnosimy się tak naprawdę do **przepustowości** (ang. *throughput*). Przepustowość łączy cyfrowych mierzy się liczbą możliwych zmian stanu na sekundę.

Rysunek 20.2 ilustruje, jak prędkość łącza może wpływać na liczbę zmian stanu. Sygnał po lewej pokazuje sześć zmian stanu w określonym czasie. Dwie równorzędne wartości nie wymagały zmiany stanu, zaszło więc tylko pięć zmian stanu na sześć możliwych. Sygnał po prawej obrazuje 19 możliwych zmian stanu w tym samym czasie, z których 17 miało miejsce. Sygnał ten może być opisany jako sygnał o przepustowości charakteryzującej się większą liczbą bitów na sekundę (b/s).

Jeśli ktoś mówi, że DS3 posiada szersze pasmo niż T1, stwierdza właściwie, iż DS3[2] oferuje większą przepustowość. W tym przypadku 45 milionów bitów na sekundę (Mb/s) w porównaniu do zaledwie 1,54 Mb/s dla T1. W powszechnym użyciu terminy szerokość pasma i przepustowość są stosowane wymiennie. W stosunku do łączy bardziej odpowiednim terminem może być **częstotliwość przepływu danych** (ang. *data rate*). DS3 jest w stanie dostarczyć określoną liczbę bitów w krótszym czasie niż T1. Innymi słowy, sieć DS3 jest szybsza niż T1.

[2] Jest to standard stosowany w Ameryce Północnej; jego europejskim odpowiednikiem jest sieć E3 — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 20.2. Zmiany stanu w czasie w przypadku sygnału cyfrowego

BERT

BERT (ang. *Bit Error Rate Test*) to test poziomu błędnych bitów. Pewnie spotkasz się z tym, że ktoś dziwnie na Ciebie spojrzy, kiedy powiesz: „Musimy wykonać kilka BERT-ów dla T1".

Testy BERT służą do wykrywania zakłóceń i są przeprowadzane na łączu w celu sprawdzenia integralności danych dla określonego obwodu. Wykonuje się je zazwyczaj poprzez ustawienie pętli zwrotnej na odległym końcu łącza, a następnie przesyłanie specjalnych wzorów zer i jedynek. Sposób zwracania danych z odległego końca pętli zwrotnej pomaga zdiagnozować rodzaj problemów występujących na tej linii. Jedne z bardziej powszechnych typów testów BERT, z którymi możesz się spotkać, to: QRSS (ang. *Quasi-Random Signal Source*), 3 z 24, 1:7, 1 z 8, Min/Max, Same Jedynki (ang. *All Ones*) oraz Same Zera (ang. *All Zeros*). Każdy z tych wzorów powoduje „naprężenie" łącza w szczególny, przewidywany sposób. Do wykonywania testów BERT na T1 stosuje się urządzenie zwane **T-Berd**.

Biuro centralne

Biuro centralne (ang. *central office* — CO) to punkt, w którym fizycznie kończą się linie telefoniczne prowadzące z domów prywatnych lub przedsiębiorstw. Biuro centralne posiada odpowiedni sprzęt przełączający, który umożliwia trasowanie rozmów lokalnie lub na inną podnośną (w razie potrzeby). Jeśli wykonujesz rozmowę skierowaną poza obszar nieruchomości, w której się znajdujesz, pierwszym przystankiem jest CO.

W przypadku technologii typu T1 łącze miedziane z Twojej lokalizacji będzie prawdopodobnie kończyć się w biurze centralnym, gdzie może być agregowane jako część większego systemu SONET.

Bank kanałów

Bank kanałów (ang. *channel bank*) jest urządzeniem, które służy do rozdzielenia łącza T1 na 24 indywidualne linie telefoniczne lub łączenia ich na powrót w jedno łącze. Obecne centrale PBX (ang. *Private Branch Exchange*) same pełnią tę funkcję, nie ma więc potrzeby stosowania banku kanałów.

CSU/DSU

CSU (ang. *Channel Service Unit*) jest jednostką usług kanału, a DSU (ang. *Data Service Unit*) to jednostka usługi danych. CSU odpowiada za relacje z łączem WAN, podczas gdy DSU odpowiada za relacje ze sprzętem informatycznym, takim jak routery. CSU/DSU jest połączeniem obu funkcji w jednym urządzeniu. Zazwyczaj **demarc** łączy się z CSU/DSU kablem

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

z końcówką RJ-45, a CSU/DSU połączony jest z routerem za pomocą kabla V.35. CSU/DSU umożliwia z reguły konfigurację obsługi wszystkich popularnych opcji sygnalizacji i tworzenia ramek dla T1. Nowoczesne routery Cisco obsługują karty interfejsu WAN (ang. *WAN interface cards* — WIC), które posiadają zintegrowane funkcje CSU/DSU.

CPE

CPE (ang. *customer premises equipment*) oznacza sprzęt znajdujący się w siedzibie klienta. Przykładem takiego sprzętu są centrale PBX, telefony, routery, a nawet modemy kablowe. Pierwotnie termin ten dotyczył sprzętu, który znajduje się w siedzibie klienta, ale jest własnością dostawcy usług telekomunikacyjnych. W obecnym znaczeniu tego pojęcia nie jest istotny podmiot będący właścicielem sprzętu.

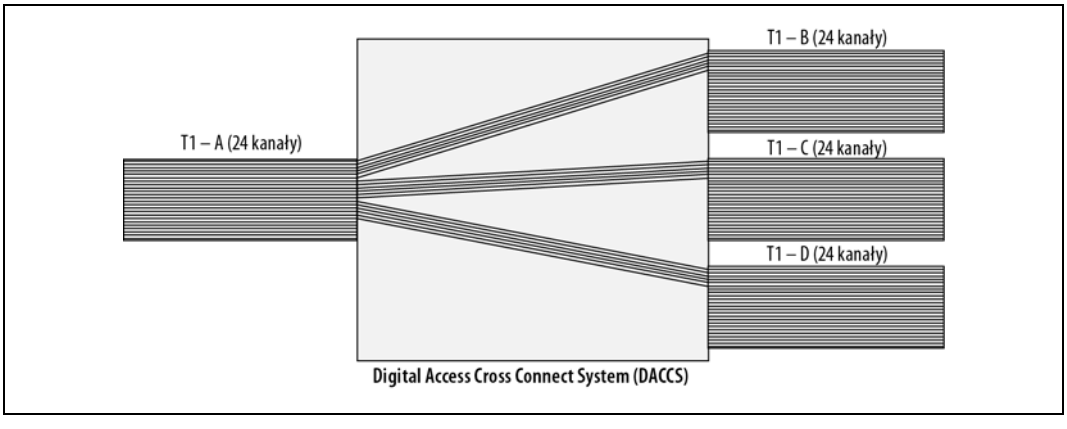
Inżynierowie telekomunikacji często skracają w mowie potocznej słowo **premises** do formy **prem**.

DACCS

DACCS (ang. *Digital Access Cross-Connect System*) jest mechanizmem przełączania kanałów w telekomunikacji cyfrowej. Termin ten występuje również pod nazwami DAX lub DACS®. Ta ostatnia jest zarejestrowanym znakiem towarowym AT&T.

DACCS jest urządzeniem, które umożliwia zmiany sposobu, w jaki kanały głosowe są połączone między trunkami. Odbywa się to za pomocą oprogramowania.

Rysunek 20.3 przedstawia logiczny schemat działania DACCS. T1-A łączy się z DACCS i wykorzystuje 18 ze swoich 24 kanałów. Kanały te muszą być skierowane w różne miejsca. Dzięki DACCS możemy połączyć pierwsze sześć kanałów (1 – 6) T1-A do kanałów 1 – 6 na T1-B, następne sześć kanałów (7 – 12) T1-A do kanałów 1 – 6 na T1-C i kolejne sześć kanałów (13 – 18) T1-A do kanałów 1 – 6 na T1-D. Kanały nie muszą być grupowane i mogą być mapowane pomiędzy łącza w dowolny sposób, zakładając, że na tych łączach są wolne kanały.



Rysunek 20.3. DACCS

Demarc

Demarc jest slangowym skrótem dla angielskiego określenia *demarcation point*, które oznacza punkt graniczny. Punkt graniczny to punkt, w którym kończy się odpowiedzialność

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

dostawcy usług telekomunikacyjnych, a zaczyna się odpowiedzialność klienta. Punkty graniczne przybierają najczęściej postać szaf telekomunikacyjnych lub podobnych miejsc, które mogą zabezpieczyć dostęp tylko dla inżynierów usługodawcy. Jeśli chcesz, aby łącze T1 kończyło się w innym miejscu budynku niż tam, gdzie znajduje się punkt graniczny, dostawca może **przesunąć punkt graniczny** (ang. *extended demark*), by dostosować się do Twoich potrzeb.

Hierarchia sygnału cyfrowego

Hierarchia sygnału cyfrowego (ang. *digital signal* — DS) opisuje poziomy sygnalizacji (ang. *signaling rate*) dla łączy trunkowych. Łącza te są fizycznymi linkami, na których logicznie opiera się system T-carrier, zwany podnośną T.

Numery podnośnej rosną wraz ze zwiększaniem się poziomu zwielokrotnienia DS0. DS0 jest podstawowym poziomem przenoszenia sygnału i jest to poziom wymagany dla pojedynczej linii telefonicznej. Ta hierarchia przedstawiona jest w tabeli 20.1.

Tabela 20.1. Hierarchia sygnału cyfrowego

Zwielokrotnienie	Podnośna	Przepływność	Liczba kanałów głosowych
DS0	Niedostępna	64 kb/s	1
DS1	T1	1,544 Mb/s	24
DS2	T2	6,312 Mb/s	96
DS3	T3	44,736 Mb/s	672
DS4	T4	274,176 Mb/s	4032

Podnośna E

Hierarchia podnośnej E (ang. *E-carrier*) jest podobna do podnośnej T (została ona opisana w dalszej części) stosowanej w USA. Nieco odmienne są jednak prędkości i sposób sygnalizacji. Podnośna E jest standardem europejskim, a jej hierarchię przedstawia tabela 20.2.

Tabela 20.2. Hierarchia podnośnej E stosowanej w Europie

Zwielokrotnienie	Przepływność	Liczba kanałów głosowych
E0	64 kb/s	1
E1	2,048 Mb/s	30
E2	8,448 Mb/s	120
E3	34,368 Mb/s	480
E4	139,268 Mb/s	1920
E5	565,148 Mb/s	7680

ISDN

ISDN (ang. *Integrated Services Digital Network*) jest technologią cyfrowej transmisji głosu oraz danych. W przeciwieństwie do tradycyjnych linii POTS lub usług kanału T1 — stosujących ścieżkę głosową dla sygnalizacji — ISDN wykorzystuje dla sygnalizacji wydzielony kanał zwany **data channel**. Pozostałe kanały określane mianem **bearer channels** mogą być wykorzystywane wyłącznie do transmisji głosowych. Ponieważ dla sygnalizacji przeznaczony jest oddzielny kanał, ISDN oferuje większą funkcjonalność.

Dla *bearer channels* stosuje się czasem określenie **B-channels**, a *data channel* nazywany jest **D-channel**.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jedną z zalet ISDN jest możliwość obsługi zarówno standardowych połączeń telefonicznych, jak i cyfrowych połączeń ISDN. We wczesnych latach 90. ISDN rozbudził wielkie oczekiwania: miał zrewolucjonizować usługi telefoniczne i usługi transmisji danych. Wyróżniamy dwa rodzaje łączy ISDN:

BRI

BRI (ang. *Basic Rate Interface*) jest łączem ISDN składającym się z dwóch kanałów *B-channel* 64 kb/s i jednego kanału *D-channel* 16 kb/s.

PRI

PRI (ang. *Primary Rate Interface*) jest łączem ISDN T1 składającym się z dwudziestu trzech kanałów *B-channel* 64 kb/s oraz jednego kanału *D-channel* 16 kb/s. Łącza PRI są często stosowane do łączenia systemów PBX oraz przez dostawców usług internetowych dla linii wdzwanianych (ang. *dial-up lines*).

Obwody PRI są dziś stosowane dla transmisji głosowych. Natomiast obwody BRI, niegdyś używane powszechnie do transmisji danych, są szeroko zastępowane tańszymi alternatywami, jak DSL.

IXC

IXC (ang. *interexchange carrier*) jest terminem stosowanym w USA na określenie przedsiębiorstwa telefonicznego, które zapewnia połączenia pomiędzy lokalnymi centralami. Połączenie między obszarami LATA może wymagać udziału IXC.

Podnośna J

Hierarchia podnośnej J (ang. *J-carrier*) jest znacznie bardziej zbliżona do stosowanej w USA podnośnej T (jeśli chodzi o prędkości) niż hierarchia europejska. Wartości te zmieniają się wraz ze wzrostem prędkości. Podnośna J jest nadal spotykana w praktyce, jednak większość obwodów, nad którymi pracowałem w Japonii, była w rzeczywistości łączami E1 lub T1.

Hierarchię podnośnej J, stosowanej w Japonii, przedstawia tabela 20.3.

Tabela 20.3. Hierarchia podnośnej J stosowanej w Japonii

Zwielokrotnienie	Przepływność	Liczba kanałów głosowych
J0	64 kb/s	1
J1	1,544 Mb/s	24
J2	6,312 Mb/s	96
J3	32,064 Mb/s	480
J4	397,200 Mb/s	5760

LATA

LATA (ang. *local access and transport area*) to lokalny obszar dostępu i transportu. Obszary LATA są geograficznie zdefiniowanymi strefami, w obrębie których dostawca telekomunikacyjny zapewnia lokalną usługę. Przykładowo Regionalni Operatorzy Telekomunikacyjni Bell (ang. *Regional Bell Operating Companies* — RBOC) nie byli zazwyczaj dopuszczani do świadczenia usług między obszarami LATA (interLATA), mogli jednak świadczyć usługi wewnątrz tych obszarów (intraLATA).

Obszary LATA mają zastosowanie w przypadku obwodów punkt-punkt, takich jak T1. Jeśli łącze T1 ma swój początek i koniec w obrębie tego samego obszaru LATA, koszty takiego

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

obwodu są zazwyczaj znacznie mniejsze niż w przypadku obwodu zaczynającego się w jednym obszarze, a kończącego w drugim. Wiąże się to z udziałem IXC w łączeniu obszarów LATA. Kolejną komplikacją są tutaj obszary geograficzne LATA, które nie zawsze pokrywają się z politycznymi granicami hrabstw lub stanów.

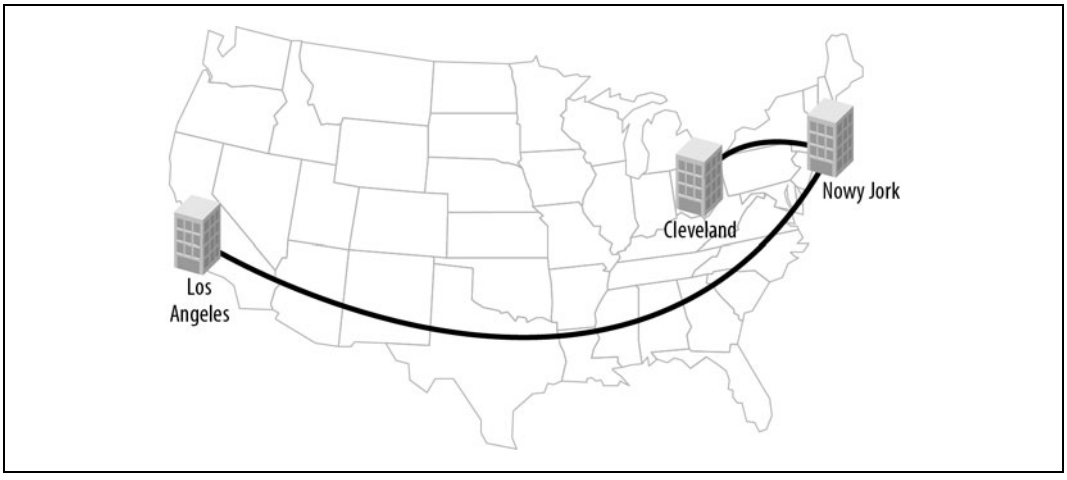
Latencja

Latencja (ang. *latency*) to termin odnoszący się do czasu, który jest potrzebny na przetworzenie danych i przekazanie ich dalej. Latencja nie ma nic wspólnego z przepustowością, szerokością pasma ani z prędkością łącza. Latencja zależna jest od dystansu, prędkości światła oraz szybkości sprzętowego przetwarzania danych.

Latencja na łączach sieciowych jest kombinacją opóźnienia propagacji (ang. *propagation delay*) i opóźnienia przetwarzania (ang. *processing delay*):

Opóźnienie propagacji

Rysunek 20.4 przedstawia trzy lokalizacje: Nowy Jork, Cleveland oraz Los Angeles. Lokalizacje połączone są dwoma linkami: jedno łącze T1 pomiędzy Nowym Jorkiem i Cleveland, a drugie łącze T1 pomiędzy Nowym Jorkiem i Los Angeles. Oba łącza mają tę samą prędkość (1,54 Mb/s), jednak więcej czasu zajmuje transmisja pakietu z Nowego Jorku do Los Angeles niż z Nowego Jorku do Cleveland.



Rysunek 20.4. Różne opóźnienia propagacji

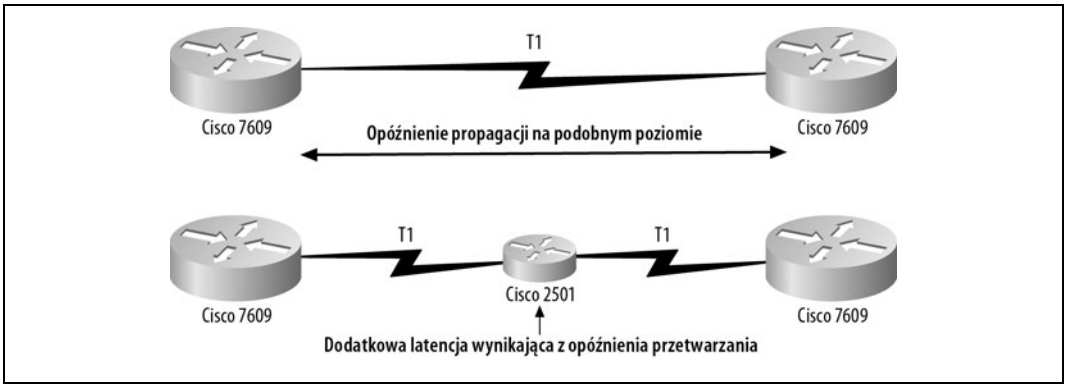
Rozbieżność ta wynika stąd, że Los Angeles jest dużo bardziej oddalone od Nowego Jorku niż Cleveland. Ten rodzaj latencji nazywany jest **opóźnieniem propagacji**. Opóźnienie propagacji zależy w dużym stopniu od praw fizyki. Nie może być więc usunięte, zmniejszone lub w jakikolwiek sposób zmienione (bez względu na to, czego życzy sobie Twój szef). W uproszczeniu, prędkość, z jaką elektrony mogą przenosić impulsy elektryczne, jest ograniczona. Podobnie ograniczona jest prędkość fotonów poruszających się w wiązce światłowodu.

Opóźnienie przetwarzania

Inna forma latencji, zwana opóźnieniem przetwarzania, wynika z czasu, jakiego urządzenie potrzebuje na przetworzenie otrzymanej informacji. W przeciwieństwie do opóźnienia propagacji, które nie może być zmienione, opóźnienie przetwarzania jest funkcją prędkości zastosowanego sprzętu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Rysunek 20.5 prezentuje dwa linki. Górny link jest bezpośrednim połączeniem pomiędzy dwoma nowoczesnymi routerami Cisco 7609 i charakteryzuje się zastosowaniem jedynie najnowszego sprzętu. Dolny link łączy te same dwa routery, jednak pomiędzy nimi znajduje się starszy router Cisco 2501.



Rysunek 20.5. Opóźnienie przetwarzania

Mimo że całkowity dystans między dwoma routerami Cisco 7609 jest taki sam od punktu do punktu w obydwu przypadkach, dodanie w drugim przypadku routera Cisco 2501 pośrodku dramatycznie zwiększa opóźnienie przetwarzania dla tego połączenia.

Innym przykładem wzrostu opóźnienia przetwarzania jest multilink PPP. Połączenie trzech linków T1 1,54 Mb/s w jeden link logiczny 4,5 Mb/s może wydawać się dobrym pomysłem (i tak często jest), jednak dodatkowe opóźnienie przetwarzania może być ogromne.

Zwróć uwagę na wartość opóźnienia dla interfejsu T1 pokazaną na listingu z użycia polecenia `show interface`:

```
Router#sho int s0/1 | inc DLY
  MTU 1500 bytes, BW 1544 Kbit, DLY 20000 usec,
```

Porównaj teraz tę wartość z opóźnieniem dla interfejsu multilink:

```
Router#sho int multilink 1 | inc DLY
  MTU 1500 bytes, BW 100000 Kbit, DLY 100000 usec,
```

Opóźnienie przetwarzania dla interfejsu multilink jest pięć razy większe niż dla interfejsu szeregowego T1.

Wartości przepustowości oraz opóźnienia pokazane dla danego interfejsu odzwierciedlają rzeczywistą przepustowość i opóźnienie, zakładając, że nie zostały one zmodyfikowane. Przepustowość i opóźnienie w IOS mogą być konfigurowane. Domyślne wartości oddają na tyle dobrze opóźnienie propagacji, aby zilustrować wpływ multilink PPP.

LEC

Mianem LEC (ang. *local exchange carrier*) określa się przedsiębiorstwo telekomunikacyjne, które zapewnia usługi lokalne, w przeciwieństwie do IXC, które łączy ze sobą przedsiębiorstwa typu LEC, zapewniając usługę o dalekim zasięgu. W większości przypadków funkcję LEC pełni RBOC.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Pętla lokalna

Pętla lokalna (ang. *local loop*) zwana jest też czasem **ostatnią milą** (ang. *last mile*). Jest to połączenie kablem miedzianym centrali firmy telekomunikacyjnej z obiektem takim jak biuro czy dom prywatny. O ile T1 może być konwertowane i zwielokrotniane w większe obwody, jak DS3 lub SONET, o tyle ostatnia mila opiera się zazwyczaj na miedzi.

Multipleksowanie

Multipleksowanie (ang. *multiplexing*), czyli zwielokrotnianie, jest to łączenie większej liczby sygnałów w jeden sygnał transmisyjny. Przykładem multipleksowania jest skonwertowanie dwudziestu czterech kanałów 64 kb/s na pojedynczy link T1.

PBX

PBX (ang. *private branch exchange*) jest zasadniczo systemem telefonicznym w takiej postaci, w jakiej zna go większość z nas. PBX umożliwia tworzenie sieci telefonicznej dla takich podmiotów jak przedsiębiorstwa. Jedną z podstawowych funkcji centrali PBX jest zdolność do współdzielenia ograniczonej liczby publicznych linii telefonicznych przez większą liczbę aparatów telefonicznych. Kolejną rzeczą jest możliwość przypisywania telefonom w obrębie centrali indywidualnych numerów wewnętrznych, składających się zazwyczaj z trzech lub czterech cyfr. Tradycyjne systemy PBX to duże urządzenia z zastrzeżonymi systemami kontroli, zawierające opatentowane rozwiązania konkretnych firm. Systemy VoIP są często kontrolowane przez programowe wersje PBX. Przykładem może być Cisco Unified Communications Manager (poprzednio Call Manager) albo środowisko Asterisk, które jest oprogramowaniem open source.

POTS

POTS (ang. *plain-old telephone service*) jest **najstarszą, podstawową usługą telefoniczną**. Linie POTS to takie linie telefoniczne, do których możesz podłączyć standardowy analogowy telefon lub faks. Większość domowych linii telefonicznych to linie POTS.

RBOC

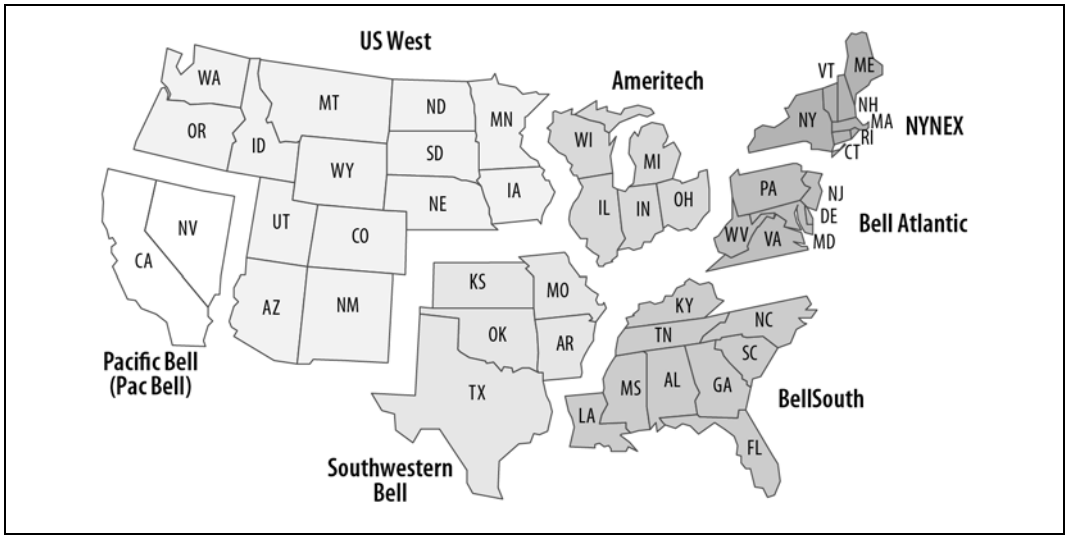
RBOC (ang. *Regional Bell Operating Company*) to Regionalni Operatorzy Telekomunikacyjni Bell.

W roku 1875 Alexander Graham Bell (przy udziale dwóch innych inwestorów) założył przedsiębiorstwo Bell Telephone Company. Firma ta, znana później jako Bell System, zaczęła skupować pakiety kontrolne innych spółek, takich jak Western Electric. W 1885 r. American Telephone and Telegraph (AT&T) przyłączyła się do budowy i eksploatacji pierwotnej sieci telefonicznej o dalekim zasięgu w USA. W 1889 r. AT&T wykupiło firmę Bella. Przez prawie 100 lat AT&T oraz należące do niego Bell System działały jako prawnie uregulowany monopol, tworząc najlepszą pod wieloma względami sieć telefoniczną na świecie.

Jednak w 1984 r. na podstawie przepisów antymonopolistycznych AT&T zostało rozdzielone z Bell System, znanym jako Ma Bell. W rezultacie wydzielono siedem regionalnych przedsiębiorstw typu RBOC powszechnie nazywanych **Baby Bells**. Po podziale AT&T pozostało dostawcą usług o dalekim zasięgu (lub IXC). W 1996 r. ustawa o deregulacji w usługach telekomunikacyjnych (ang. *Telecommunications Deregulation Act*) zezwoliła zarówno przedsiębiorstwom RBOC (zwanym też LEC), jak i kompaniom świadczącym usługi dalekiego zasięgu na sprzedawanie usług lokalnych, usług o dalekim zasięgu oraz usług międzynarodowych, czyniąc te rozgraniczenia bardzo mglistymi.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Każdemu ze wspomnianych siedmiu przedsiębiorstw przydzielony został określony region w kraju, w którym mogło ono prowadzić swoją działalność. Regiony te zostały przedstawione na rysunku 20.6.



Rysunek 20.6. Regiony RBOC

Większość z pierwotnych RBOC jest teraz częścią SBC (poprzednio Southwestern Bell), które nabyło również AT&T, przybliżając całą tę zawrotną aferę ponownie o krok bliżej w stronę monopolu. Oto jak przekształciły się pierwotne RBOC:

Bell Atlantic
W 1996 r. Bell Atlantic połączyło się z NYNEX, tworząc Verizon.

Southwestern Bell
Southwestern Bell zmieniło nazwę na SBC w 1995 r. SBC przejęło PacBell w 1996 r., a następnie zakupiło Ameritech, Southern New England Telecomunications oraz AT&T. Po przejęciu tej ostatniej spółki SBC przyjęło nazwę AT&T.

NYNEX
W 1996 r. NYNEX połączył się z Bell Atlantic, tworząc Verizon.

Pacific Bell (PacBell)
PacBell zostało przejęte przez SBC w 1996 r.

BellSouth
BellSouth zostało nabyte przez SBC/AT&T.

Ameritech
Ameritech zostało przejęte przez SBC w 1998 r.

US West
US West zostało przejęte przez Qwest Communications w roku 2000.

Smart jack
Smart jack jest urządzeniem, które znajduje się na końcu obwodu cyfrowego. Uważane jest za „mądre" (ang. *smart*), ponieważ może być zdalnie kontrolowane przez firmę telefoniczną. *Smart jack* służy również za punkt testowy dla urządzeń takich jak testery BERT. Łącza T1

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

są zazwyczaj zakończone interfejsami *smart jack*. W większych instalacjach montuje się zazwyczaj panele rackowe z tymi interfejsami (ang. *smart jack racks*).

SONET

SONET (ang. *synchronous optical network*) to synchroniczna sieć optyczna. SONET jest standardem ANSI dla systemów transmisji światłowodowej. Europejskim odpowiednikiem tego standardu jest synchroniczna hierarchia cyfrowa (ang. *synchronous digital hierarchy* — SDH).

Jak wskazuje nazwa, SONET jest siecią wyłącznie optyczną i przy tym bardzo szybką. Tabela 20.4 przedstawia poziomy **podnośnej optycznej** (ang. *optical carrier*) dla SONET.

Tabela 20.4. Poziomy podnośnej optycznej

Poziom podnośnej optycznej	Przepływność linii	Poziom obciążenia
OC1	51 Mb/s	50 Mb/s
OC3	155 Mb/s	150 Mb/s
OC12	622 Mb/s	601 Mb/s
OC48	2488 Mb/s	2405 Mb/s
OC192	9953 Mb/s	9621 Mb/s
OC768	39813 Mb/s	38486 Mb/s

Podnośna T

Podnośna T (ang. *T-carrier*) jest ogólną nazwą dla cyfrowych, zwielokrotnionych systemów przekazywania. Litera „T" oznacza **trunk**, ponieważ linki te zostały zaprojektowane do łączenia trunkiem wielu linii telefonicznych pomiędzy biurami centralnymi. Hierarchia podnośnej T jest stosowana w USA i Kanadzie. W Europie stosowana jest podobna skala, zwana hierarchią podnośnej E, a w Japonii stosuje się hierarchię podnośnej J. Hierarchię podnośnej T przedstawia tabela 20.5.

Tabela 20.5. Hierarchia podnośnej T stosowanej w Ameryce Północnej

Zwielokrotnienie	Przepływność	Liczba kanałów głosowych
T1	1,544 Mb/s	24
T1C	3,152 Mb/s	48
T2	6,312 Mb/s	96
T3	44,736 Mb/s	672
T4	274,176 Mb/s	4032

T-Berd

T-Berd (ang. *T1 Bit Error Rate Detector*), czyli **tester poziomu błędnych bitów dla T1**. Jest to nazwa stosowana dla każdego urządzenia umożliwiającego przeprowadzenie testów BERT dla łącza T1. Jeśli Twoje łącze T1 nie zachowuje się właściwie, usługodawca prawdopodobnie przyśle inżyniera z T-Berd w celu przeprowadzenia inwazyjnych testów.

TDM

TDM (ang. *time-division multiplexing*) to multipleksowanie z podziałem czasu. Link T1 jest linkiem TDM, ponieważ jego 24 kanały są podzielone na szczeliny czasowe (ang. *time slots*). Ponieważ T1 jest linkiem szeregowym, w danym momencie wysyłany jest tylko jeden bit. Kanały zmieniają się cyklicznie z dużą prędkością, a każdy kanał jest przypisany do odpowiedniego przedziału czasowego.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 21.

T1

W latach 50. ubiegłego stulecia jedyną metodą łączenia linii telefonicznych było zastosowanie jednoparowego przewodu miedzianego. Każda linia telefoniczna wchodząca do budynku wymagała jednej pary miedzianych kabli. Przeciążenia przewodów stanowiły w tym czasie ogromny problem w centralnych biurach (ang. *central offices* — CO) korporacji telefonicznych oraz w podziemnych tunelach kablowych na obszarach miejskich. Wyobraź sobie CO w sporym mieście, gdzie zbiegają się dziesiątki tysięcy linii telefonicznych, z których każda wymaga pary przewodów. Biura te musiały jeszcze komunikować się między sobą, co wymagało dodatkowego okablowania.

W roku 1961 laboratoria Bell Labs z New Jersey wynalazły łącza T1 jako narzędzie do cyfrowego trunkingu większej liczby kanałów głosowych pomiędzy wybranymi lokalizacjami. T1 oferowało zmniejszenie obciążenia w stosunku 12:1, ponieważ zastępowało 24 dwuprzewodowe linie telefoniczne zaledwie 4 przewodami. W tym okresie był to przełom w myśleniu. Pamiętaj, że wtedy technologia cyfrowa jeszcze praktycznie nie istniała. Pierwsze łącze T1 oddano do użytku w 1962 r., łącząc biuro firmy Bell na Dearborn w Chicago w stanie Illinois ze Skokie w stanie Illinois. Dzisiaj trudno byłoby znaleźć firmę, która nie wykorzystuje wielokrotnych łączy T1.

Rozdział ten poświęciłem szczegółom dotyczącym budowy, funkcjonowania oraz diagnozowania problemów z łączami T1. Chociaż zazwyczaj staram się upraszczać złożone tematy techniczne, w tym przypadku myślę, że ważne jest zrozumienie podstawowych zasad działania T1. Żyjemy w świecie, w którym wszystko jest ze sobą połączone, a znaczna część tego świata korzysta ze standardu T1. Znajomość działania łączy T1 może oszczędzić Ci wiele godzin spędzonych na usuwaniu problemów, które doprowadziły do przerwy w działaniu sieci[1].

Jak działa T1 w dupleksie

Link w pełnym dupleksie może wysyłać i odbierać dane w tym samym momencie. Każdy, kto kiedykolwiek kłócił się ze swoją drugą połówką przez telefon, może potwierdzić: obie strony mogą mówić (lub krzyczeć) w tym samym momencie. To jest właśnie rozmowa w pełnym

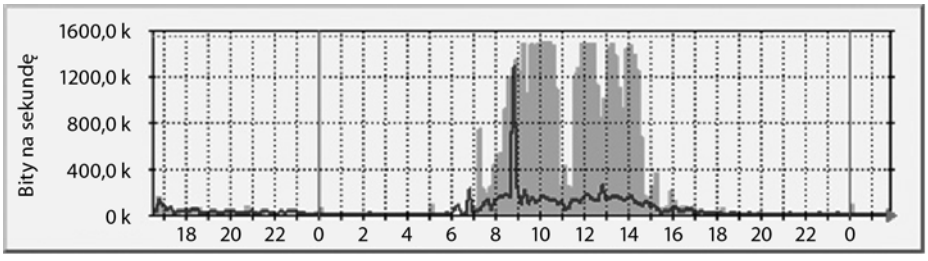
[1] Niniejszy rozdział poświęcony jest w całości standardowi T1, który jest podstawowym systemem zwielokrotnienia kanałów cyfrowych w telekomunikacji stosowanym w Ameryce Północnej. T1 jest standardem pierwszego zwielokrotnienia D1, o przepływności 1,544 Mb/s. Jego odpowiednikiem stosowanym w Europie jest specyfikacja E1 (2,048 Mb/s) — *przyp. tłum.*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

dupleksie. Gdyby tylko jedna osoba mogła mówić w danej chwili, mielibyśmy półdupleks. Przykładem półdupleksu jest rozmowa przez walkie-talkie. Aby mówić, musisz najpierw wcisnąć określony przycisk na słuchawce. Zazwyczaj przy wciśniętym przycisku nie możesz słyszeć osoby, do której mówisz (choć niektóre zaawansowane walkie-talkie stosują jedną częstotliwość do nadawania, a inną do odbierania, umożliwiając w ten sposób rozmawianie w pełnym dupleksie).

Łącza T1 działają w pełnym dupleksie. Kanały głosowe T1 transmitują i odbierają sygnał audio symultanicznie, a dane są przesyłane i odbierane jednocześnie na linkach WAN łącza T1. Wciąż jednak spotykam wiele osób parających się tą dziedziną, które nie rozumieją w pełni funkcjonowania dupleksu T1. Może to być związane ze sposobem, w jaki zazwyczaj opisywany jest przepływ danych na linkach WAN.

Rysunek 21.1 przedstawia wykorzystanie szerokości pasma dla łącza internetowego T1 zmierzone przez MRTG (ang. *Multi Router Traffic Grapher*). Liczby na osi poziomej odzwierciedlają godziny w cyklu dobowym (0 to północ). Wykres z szarym wypełnieniem, który przypomina krajobraz Arizony, to przepływ danych przychodzących. Czarna linia na wykresie symbolizuje przepływ danych wychodzących. Jest to typowy układ wykorzystania pasma dla bardzo obciążonego łącza internetowego T1. Większość danych pochodzi z internetu. Żądania otrzymywania danych z tego łącza są na bardzo niskim poziomie. Około 8:45 można zauważyć skok danych wychodzących — prawdopodobnie w tym czasie została wysłana duża liczba e-maili.



Rysunek 21.1. Wykorzystanie pasma według MRTG

Z wykresu nie wynika jasno, jaki typ dupleksu posiada to łącze. Można by łatwo założyć, że jest to T1 w półdupleksie, a przełączanie z nadawania na odbieranie odbywa się bardzo szybko. Ktoś, kto nie rozumie tej technologii, mógłby wywnioskować z tego wykresu, że T1 umożliwia przesyłanie danych tylko w jednym kierunku naraz. Zdziwiłbyś się, jak częsty jest to błąd.

Rodzaje T1

Terminologia jest bardzo ważna, kiedy rozmawia się o technologiach, a T1 nie jest tu wyjątkiem. Wiele terminów jest stosowanych niewłaściwie, nawet przez osoby od lat pracujące w branży. Pojęcia **T1** oraz **DS1** są często używane zamiennie, jednak możesz sobie napytać kłopotów, jeśli będziesz rozmawiał w ten sposób z osobami z długim stażem w telekomunikacji. Możesz również spotkać się z tym, że niektórzy nazywają PRI (ang. *Primary Rate Interface*) „cyfrowym T1", co nie jest do końca poprawne. Wszystkie łącza T1 są cyfrowe. Różnica z PRI polega na tym, że wykorzystuje ono cyfrowy system sygnalizacji dla kanału danych w przeci-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

wieństwie do analogowego systemu sygnalizacji dla kanałów głosowych. Nawet w przypadku „analogowego" T1 każdy kanał głosowy musi zostać przekonwertowany na postać cyfrową, zanim zostanie przesłany łączem T1.

Możesz natknąć się na wiele sprzecznych informacji, kiedy będziesz próbował dowiedzieć się czegoś o T1. Jest to szeroki zakres wiedzy, chociaż wyróżniamy tylko kilka rodzajów T1:

Kanał T1 (ang. *Channelized T1*)
: Kanał T1 jest obwodem głosowym składającym się z 24 kanałów głosowych. Każdy kanał posiada własną informację sygnalizacyjną, która jest umieszczana w strumieniu danych cyfrowej transmisji głosowej. Nazywa się to **sygnalizacją w paśmie** (ang. *in-band signaling*). Pod warunkiem prawidłowego przygotowania przewodu (patrz kolejne podrozdziały: „Kodowanie" na stronie 349 oraz „Ramkowanie" na stronie 351) z wykorzystaniem Add/Drop CSU/DSU, kanał T1 może zostać użyty do transmisji danych.

PRI
: PRI jest obwodem głosowym posiadającym 24 kanały, z których jeden jest przeznaczony na sygnalizację. Stąd też dostępna liczba kanałów głosowych wynosi 23. Kanały głosowe zwane są **bearer channels** (kanały B), a kanał sygnalizacyjny nazywa się **data channel** (kanał D). Ten rodzaj sygnalizacji określany jest mianem **sygnalizacji poza pasmem** (ang. *out-of-band signaling*).

Czysty kanał T1 (ang. *Clear-channel T1*)
: Czysty kanał T1 nie posiada struktur ramek. Nie ma też wydzielonych kanałów ani organizacji bitów przepływających przez łącze. Czyste kanały T1 są rzadkością, ponieważ większość linków stosuje ramkowanie ESF (ang. *Extended Super Frame*).

Sygnalizację w paśmie można porównać do wciskania przycisków na telefonie w trakcie rozmowy. Sygnały tonowe znajdują się w ścieżce głosowej połączenia telefonicznego. Są one wysyłane na początku rozmowy (czasem możesz je usłyszeć) z jednego przełącznika do drugiego (przy zastosowaniu sygnalizacji CAS), aby dostarczyć przełącznikowi sygnał do podjęcia decyzji o tym, w jaki sposób ma być rutowana rozmowa. Wewnątrz kanału znajdują się również sygnały, które nie są słyszalne. Są to bity osadzone w danych głosowych. Bity te zwane są bitami **ABCD** i służą do raportowania statusu telefonu (np. słuchawka podniesiona/odłożona).

Z kolei sygnalizacja poza pasmem działa podobnie do protokołu FTP: jeden kanał jest wykorzystywany do nawiązania połączenia, a następnie odrębny kanał jest wybierany do dostarczenia ładunku (w tym przypadku rozmowy głosowej).

Kodowanie

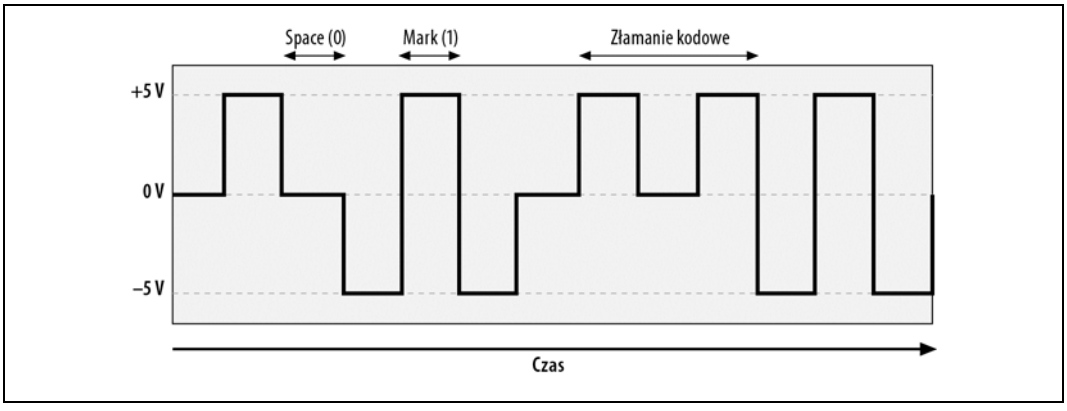
Kodowanie (ang. *encoding*) określa metodę, która jest stosowana do generowania i dekodowania sygnałów elektrycznych. Obecnie dla łączy T1 wykorzystuje się dwa rodzaje kodowania: **AMI** (ang. *Alternate Mark Inversion*) oraz **B8ZS** (ang. *Binary Eight Zero Substitution*). Zasadniczo kodowanie AMI stosowane jest dla obwodów głosowych, a B8ZS dla danych. B8ZS może również być wykorzystane do transmisji głosowych, jednak nie należy stosować AMI dla danych — o powodach napiszę poniżej.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

AMI

AMI to metoda kodowania wykorzystująca znaki o zmiennej polaryzacji. W sygnalizacji T1 występują dwa możliwe stany: *mark* oraz *space*. W uproszczeniu, **mark** reprezentowany jest przez cyfrę 1, a **space** przez cyfrę 0. Dla T1 *space* posiada napięcie 0 V, a *mark* przyjmuje wartości napięcia +5 V oraz –5 V. AMI koduje sygnał w taki sposób, że polaryzacja każdego znaku jest przeciwna do polaryzacji poprzedniego.

Takie rozwiązanie pozwala zastosować ciekawe techniki wykrywania błędów. Przykładowo rysunek 21.2 przedstawia dwie jedynki o tej samej polaryzacji występujące pod rząd. Jest to traktowane jako błąd zwany **złamaniem kodowym** (ang. *bipolar violation* — BPV). Gdyby wszystkie jedynki posiadały napięcie dodatnie, skok napięcia mógłby zostać błędnie potraktowany jako prawidłowa jedynka. Ponieważ naprzemienne znaki są odwrócone, uzyskujemy dodatkową korzyść w postaci średniego napięcia linii fizycznej wynoszącego zawsze 0 V. Pozwala to bezpiecznie chwytać fizyczne przewody T1.



Rysunek 21.2. Sygnalizacja T1 AMI

T1 są łączami asynchronicznymi, co oznacza, że tylko jedna strona łącza zapewnia taktowanie. Druga strona musi polegać na samym sygnale podczas określania początku i końca bitów. Ponieważ czas trwania znaku (*mark*) jest znany, synchronizację można osiągnąć w prosty sposób przez odbieranie znaków. Przy zastosowaniu kodowania AMI dłuższy okres trwania stanu *space* będzie skutkował utratą synchronizacji. Przy braku znaków w sygnale strona odbierająca łącza może w końcu stracić orientację, gdzie zaczynają się i kończą bity.

Istnieje ryzyko wystąpienia sygnału składającego się z **samych zer** (ang. *all zeros*). Dlatego też AMI ustawia wartość 1 dla każdego ósmego bitu bez względu na jego wartość oryginalną. Zapewnia to odpowiednią liczbę przejść sygnału w linii, co oznacza, że **zagęszczenie jedynek** (ang. *ones density*) w strumieniu sygnału jest wystarczające do zapewnienia synchronizacji. Zaledwie 16 zer pod rząd może doprowadzić do utraty synchronizacji na drugim końcu łącza.

Sygnały głosowe mogą być łatwo absorbowane, nawet jeśli posiadają wartość każdego ósmego bitu ustawioną na 1. Ludzkie ucho nie usłyszy różnicy w przypadku, kiedy pojedynczy bit w strumieniu został zmieniony. Sygnały danych nie tolerują jednak jakichkolwiek zmian bitów. Jeśli w pakiecie TCP zmienia się jeden bit, cykliczne sprawdzanie redundancji (ang. *Cyclic*

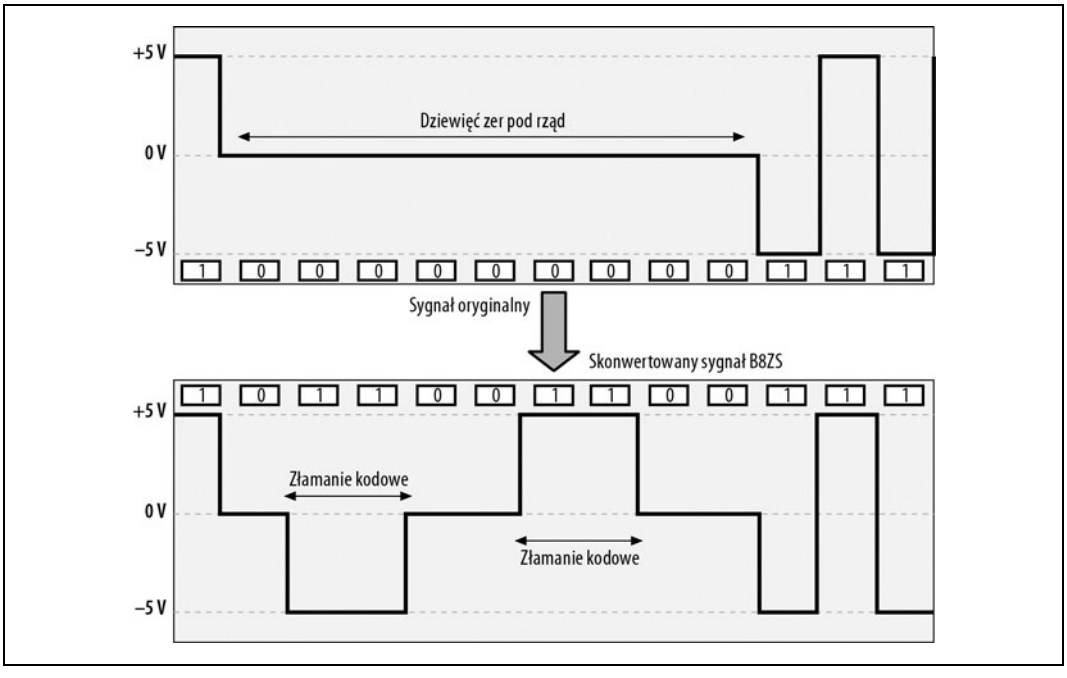
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Redundancy Check — CRC) nie powiedzie się, a sygnał zostanie wysłany ponownie (testy CRC nie są wykonywane dla pakietów UDP). Z uwagi na to ograniczenie AMI nie jest akceptowalnym rodzajem kodowania dla łączy T1 przesyłających dane.

B8ZS

Kodowanie B8ZS zostało wprowadzone, by wyeliminować ograniczenia AMI. Według koncepcji B8ZS, jeśli w sygnale zostanie wykrytych osiem zer pod rząd, są one konwertowane na wzorzec zawierający intencjonalne złamanie kodu (BPV). Kiedy na drugim końcu wzorzec zostanie rozpoznany, jest na powrót zamieniany na same zera.

Rysunek 21.3 pokazuje, jak długie ciągi zer są konwertowane w przewodzie. Górny sygnał składa się z jedynki, dziewięciu zer oraz kolejnych trzech jedynek. B8ZS zmienia pierwsze osiem zer na wzorzec zawierający dwa BPV. Nie spotkałbyś takiego wzorca w normalnym, zdrowym obwodzie. Kiedy drugi koniec otrzymuje wzorzec, zamienia go z powrotem na osiem zer. Dzięki tej technice strumień danych może posiadać tyle kolejnych zer, ile potrzeba, bez wpływu na zagęszczenie jedynek.



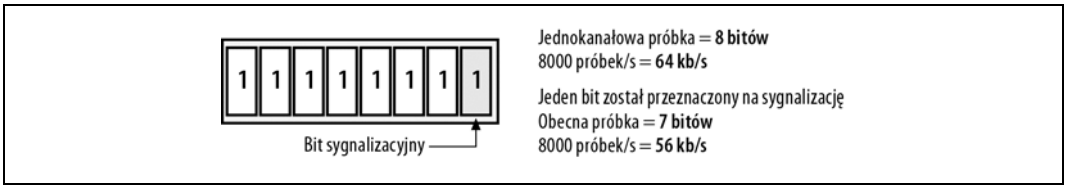
Rysunek 21.3. Kodowanie B8ZS

Ramkowanie

Telefoniczny sygnał audio jest próbkowany 8000 razy na sekundę, czyli z częstotliwością 8 kHz. Każda próbka jest konwertowana na postać ośmiobitowej wartości, w której jeden z bitów wykorzystywany jest do sygnalizacji.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Rysunek 21.4 przedstawia jednokanałową próbkę z jednym bitem dla sygnalizacji. Nazywa się to **sygnalizacją w paśmie** (ang. *in-band signaling*).



Rysunek 21.4. Jednokanałowa próbka

Kiedy łącze T1 jest skonfigurowane jako PRI, wszystkie osiem bitów w każdym kanale może być wykorzystanych dla danych. Jest to możliwe, ponieważ cały jeden kanał zarezerwowany jest dla sygnalizacji (w przeciwieństwie do wykorzystywania jednego bitu w każdym kanale). Zmniejsza to liczbę dostępnych kanałów z 24 na 23 i nazywane jest **sygnalizacją poza pasmem** (ang. *out-of-band signaling*).

T1 wykorzystuje multipleksowanie z podziałem czasu (ang. *time-division multiplexing*), co oznacza, że każdy kanał jest właściwie grupą szeregowych wartości binarnych. Kanały są przesyłane w kolejności, ale odbiornik musi wiedzieć, kiedy zaczyna się pierwszy kanał i kiedy kończy się ostatni. Odbywa się to za pomocą **ramkowania** (ang. *framing*).

D4/Superframe

W standardowym ramkowaniu głosowym, zwanym **D4** lub **superframe**, każda ośmiobitowa próbka jest transmitowana z każdego kanału w kolejności. Najpierw transmitowana jest jedna próbka z kanału pierwszego, następnie jedna próbka z kanału drugiego itd., aż do momentu, kiedy wszystkie 24 kanały prześlą po jednej próbce. Wtedy proces jest powtarzany.

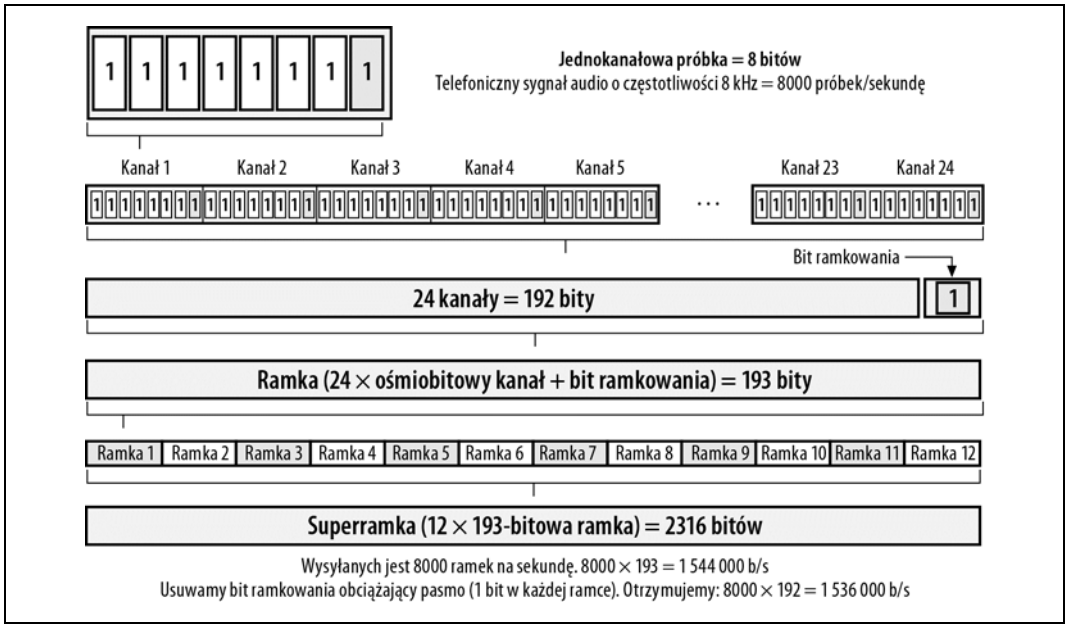
Aby na drugim końcu łącza była możliwa identyfikacja kanałów, bity ramkowania są dodawane po tym, jak wszystkie 24 kanały wyślą po jednej próbce. Ponieważ każda próbka to 8 bitów i istnieją 24 kanały, jedna iteracja dla wszystkich kanałów to 192 bity. Wliczając bit ramkowania, mamy w sumie 193 bity. Te 193-bitowe porcje zwane są **ramkami** (ang. *frames*). Każdy zestaw składający się z 12 ramek zwany jest *superframe* (superramką).

Schemat ramkowania został przedstawiony na rysunku 21.5. Urządzenia T1 mogą śledzić ramki dzięki umieszczaniu sekwencji 110111001000 w bitach ramkowania, które zawiera superramka.

Jeśli bity ramkowania nie odpowiadają oczekiwanej sekwencji, odbiornik zapisuje wystąpienie **błędu ramkowania** (ang. *framing error*). Kiedy urządzenie T1 osiąga określony próg błędów ramkowania, uruchamiany jest alarm.

Zapoznając się z różnymi informacjami na temat T1, mogłeś natrafić na pewne rozbieżności dotyczące prędkości tego łącza. Niektóre źródła podają, że T1 ma prędkość 1,544 Mb/s, podczas gdy inne stwierdzają, iż jest to 1,536 Mb/s. Ta różnica wynika z bitów ramkowania. Ponieważ bity ramkowania są wykorzystywane przez urządzenia T1 i nie są częścią danych, traktowane są jako obciążenie łącza. Tak więc użyteczna prędkość T1 wynosi 1,536 Mb/s, biorąc pod uwagę bity ramkowania.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 21.5. Ramkowanie DS1

Extended Super Frame

Standard D4/Superframe został opracowany dla transmisji głosowych i nie jest praktyczny dla przesyłu danych. Jednym z powodów, dla których D4 nie nadaje się do przesyłania danych, jest brak wykrywania błędów. Opracowano więc nowy standard oferujący detekcję błędów oraz lepsze wykorzystanie bitów ramkowania. Jest to **Extended Super Frame** (EFS), czyli rozszerzona superramka.

EFS działa na tych samych zasadach co D4/superframe, z tym wyjątkiem, że składa się z 24 ramek zamiast 12. Bit ramkowania w każdej ramce EFS jest wykorzystywany znacznie wydajniej niż w przypadku D4. Zamiast umieszczać określone sekwencje w bitach ramkowania dla całej superramki, EFS wykorzystuje te bity w następujący sposób:

Ramka 4, 8, 12, 16, 20 oraz 24 (co czwarta ramka)
W bitach ramkowania dla tych ramek umieszczana jest sekwencja 001011.

Ramka 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21 oraz 23 (wszystkie ramki nieparzyste)
Bity ramkowania tych ramek są wykorzystywane dla nowego wirtualnego kanału danych o prędkości 4000 b/s. Kanał ten jest stosowany do komunikacji poza pasmem (ang. *out-of-band*) pomiędzy urządzeniami sieciowymi w danym linku.

Ramka 2, 6, 10, 14, 18 oraz 22 (pozostałe ramki parzyste)
Bity ramkowania tych ramek są wykorzystywane do przechowywania sześciobitowej wartości CRC dla każdej superramki.

W każdym przypadku, kiedy obwód T1 jest wykorzystywany jako link WAN dla sieci danych, powinno być stosowane kodowanie B8ZS oraz ramkowanie EFS.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Monitorowanie wydajności

Urządzenia CSU/DSU przekazują status łączy T1, raportując występowanie zestawu standardowych zdarzeń. Niektóre ze zdarzeń oraz wywoływanych alarmów mogą być mylące dla profesjonalnych informatyków, którzy nie zetknęli się dotąd z tą technologią. Aby jeszcze pogorszyć sprawę, w większości raportów CSU/DSU do prezentacji błędów nie wykorzystuje się pełnych opisów, ale raczej dobrze znane (inżynierom telekomunikacji) skróty.

Różni producenci często stosują odmienne definicje zdarzeń, a znalezienie ich szczegółowych opisów może być wyzwaniem. Typy zdarzeń zostały wymienione w publikacji RFC 1232, zatytułowanej „Definitions of Managed Objects for the DS1 Interface Type" (Definicje zarządzalnych obiektów dla interfejsu DS1). Publikacja ta definiuje standardy stosowane w pułapkach SNMP i nie podaje elektrycznej specyfikacji tych alarmów. Zamieszczone opisy nie dotyczą konkretnych producentów, mogą więc nie być właściwe dla każdego urządzenia. Znajdziesz tam jednak sporo przejrzystych definicji zdarzeń.

Utrata sygnału

Utrata sygnału (ang. *loss of signal* — LOS) to stan, kiedy w określonym przedziale czasu nie zostały wykryte żadne impulsy elektryczne. Publikacja RFC 1232 opisuje LOS następująco:

> To zdarzenie jest zgłaszane na podstawie zaobserwowania 175 +/– 75 kolejnych pozycji impulsu, w których nie pojawiła się dodatnia lub ujemna polaryzacja (zwana również *keep alive*).

Znaczy to po prostu, że linia jest martwa. Nie ma żadnych alarmów, sygnałów itp. LOS jest równoważne z brakiem kabla w gnieździe T1.

Poza ramką

Stan poza ramką (ang. *out-of-frame* — OOF), zwany również utratą ramki (ang. *loss of frame* — LOF), określa sytuację, w której pewna liczba otrzymanych ramek posiada błędy bitów ramkowania. W takim przypadku dane nie mogą być uznane za zaufane, ponieważ synchronizacja na obu końcach łącza T1 jest niepoprawna. Nadmierna liczba błędów OOF uruchamia czerwony alarm. Publikacja RFC 1232 opisuje OOF następująco:

> Stan *Out of Frame* jest zgłaszany, kiedy urządzenie odbiorcze wykryje dwa lub więcej błędów bitów ramkowania w przedziale 3 milisekund lub co najmniej dwa błędy na pięć lub mniej kolejnych bitów ramkowania. W takim przypadku ustawiany jest stan OOF i rozpoczyna się wyszukiwanie właściwej sekwencji ramkowania. Stan OOF kończy się, kiedy zostanie wykryte poprawne ramkowanie.

Złamanie kodowe

Złamanie kodowe (ang. *bipolar violation* — BVP) występuje wtedy, kiedy dwa sygnały *mark* (jedynki) występują w sekwencji o tej samej polaryzacji. W sygnalizacji D1 każdy znak powinien mieć przeciwną polaryzację do znaku poprzedniego. Kiedy pojawiają się dwa znaki z tą samą polaryzacją (i nie są częścią ciągu podstawianego w kodowaniu B8ZS), traktowane jest to jako błąd. Nadmierna liczba BPV uruchamia alarm. Publikacja RFC 1232 opisuje złamania kodowe w następujący sposób:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Dla sygnałów kodowanych w B8ZS złamanie kodowe polega na odebraniu niezgodnych z systemem dwubiegunowości sygnałów, które nie należą do ciągu podstawianego zer w kodzie B8ZS. Złamanie kodowe może też zawierać inne błędy, takie jak osiem lub więcej kolejnych zer oraz niewłaściwa parzystość.

CRC6

CRC6 to sześciobitowy mechanizm cyklicznego sprawdzania redundancji (ang. *Cyclic Redundancy Check*) służący do wykrywania błędów w ESF. Błędy są wynikiem problemów z integralnością danych. Zdarzenia CRC6 nie są opisane w publikacji RFC 1232.

Errored Seconds

Errored Second to licznik wskazujący liczbę sekund w 15-minutowym interwale, w których wystąpiły błędy. Odczyt z tego licznika jest szybkim sposobem sprawdzenia, czy na łączu są jakieś problemy. Wskazuje on również poziom zakłóceń dla T1. Jeśli wskazanie licznika jest wysokie, mamy do czynienia z uporczywym problemem. Jeśli wskazanie jest niskie, jest to prawdopodobnie problem chwilowy (być może powtarzający się) lub przejściowy. Licznik *Errored Second* odnotowuje zazwyczaj sytuacje, kiedy w przeciągu jednej sekundy pojawia się jeden lub więcej problemów. Publikacja RFC 1232 opisuje błędy *Errored Second* w następujący sposób:

> *Errored Second* to sekunda, w której pojawia się jedno lub więcej zdarzeń typu *Code Violation Error Events* lub *Out of Frame*. W D4 oraz G.704 w sekcji 2.1.3.2 (np. G.704, które nie implementuje CRC) obecność BPV również uruchamia *Errored Second*.

Extreme Errored Seconds

Czasem określany również mianem *Severely Errored Seconds*. Jest to licznik, który odnotowuje przekroczenie określonego progu błędów w przeciągu sekundy. Konfiguracja progu oraz sposób naliczania błędów zależą od zastosowanego sprzętu. Nie powinieneś doświadczyć *Extreme Errored Seconds* na zdrowym łączu, ale pewne błędy *Errored Seconds* mogą się pojawić na normalnie działającym obwodzie. Publikacja RFC 1232 opisuje zdarzenia *Severely Errored Second* w następujący sposób:

> *Severely Errored Second* to sekunda, w której pojawia się 320 lub więcej zdarzeń typu *Code Violation Error Events* lub co najmniej jedno zdarzenie *Out of Frame*.

Alarmy

Alarmy to poważne sytuacje, wymagające uwagi. Mogą one być wywołane przez nadmierną liczbę błędów, problemy sprzętowe lub zakłócenia sygnału. Rodzaje alarmów są kodowane kolorami. Podobnie jak w przypadku zdarzeń dotyczących wydajności różni producenci definiują alarmy odmiennie, a znalezienie ich dokładnego opisu może stanowić wyzwanie. Większość alarmów opisana jest w publikacji RFC 1232, jednak ponownie uwzględnione jest głównie zastosowanie dla SNMP, a opisy nie stanowią standardu dla różnych rozwiązań sprzętowych.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Alarm czerwony

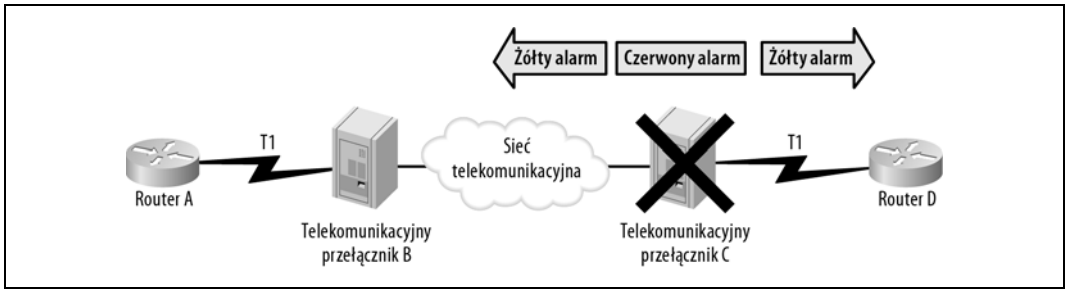
Publikacja RFC 1232 definiuje czerwony alarm (ang. *red alarm*) następująco:

> Czerwony alarm jest zgłaszany z powodu pojawiających się zdarzeń typu: *Loss of Signal*, *Loss of Framing* oraz *Alarm Indication Signal*. Po uruchomieniu czerwonego alarmu urządzenie wysyła żółty sygnał (ang. *yellow signal*) do drugiego końca łącza. Kiedy na drugim końcu łącza odebrany zostaje żółty sygnał, uruchamiany jest żółty alarm (ang. *yellow alarm*).

Ten rodzaj alarmu uruchamiany jest w przypadku wykrycia lokalnych awarii lub błędów OOF powtarzających się nieprzerwanie przez ponad **x** sekund (liczba sekund jest zależna od producenta sprzętu). Alarm zostaje wyłączony, kiedy upłynie określony czas, w którym nie wykryto błędów OOF (czas ten również zależy od zastosowanego sprzętu).

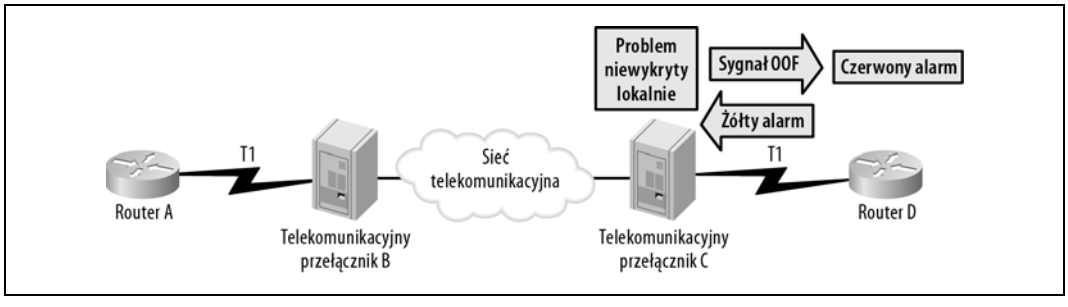
Kiedy urządzenie jest w stanie czerwonego alarmu, wysyła na zewnątrz sygnał alarmu żółtego.

Rysunek 21.6 przedstawia przykład wystąpienia czerwonego alarmu. Jakiś problem pojawił się na telekomunikacyjnym przełączniku C. Przełącznik uruchamia lokalnie czerwony alarm i wysyła sygnał żółtego alarmu, żeby ostrzec sąsiadów o problemie.



Rysunek 21.6. Czerwony alarm

Rysunek 21.7 pokazuje kolejny scenariusz czerwonego alarmu. W tym przykładzie telekomunikacyjny przełącznik C wysyła nieczytelne ramki do Routera D, który zauważa występujące pod rząd problemy OOF i wszczyna czerwony alarm. Po uruchomieniu czerwonego alarmu Router D wysyła z powrotem tym samym linkiem sygnał żółtego alarmu.



Rysunek 21.7. Żółty alarm

Sposób, w jaki zachowują się te alarmy, może być nieco mylący. Alarm czerwony zasadniczo wskazuje, że coś poważnego dzieje się z lokalnym sprzętem. Jednak w przykładzie z rysunku 21.7, Router D otrzymuje tak wiele błędów OOF, że sygnał staje się bezużyteczny. Ponieważ

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Router D nie potrafi znaleźć sposobu na odczytanie ramek przychodzących z drugiego końca łącza, uruchamia lokalnie czerwony alarm i wysyła na zewnątrz sygnał żółtego alarmu.

Alarm żółty

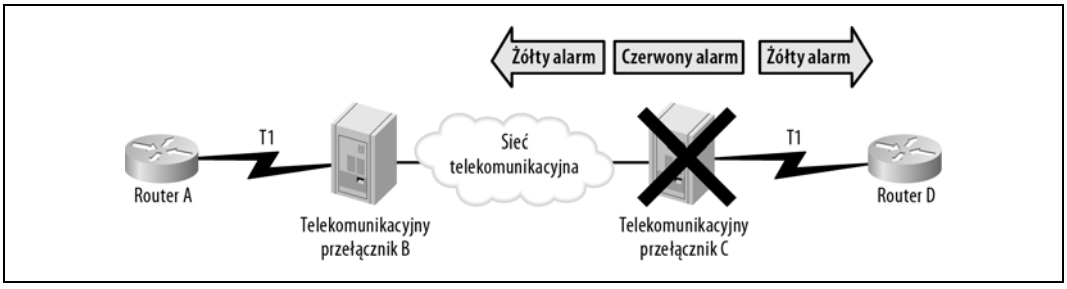
Żółty alarm, zwany również zdalnym wywołaniem alarmu (ang. *remote alarm indication* — RAI), informuje o istnieniu problemu na odległym końcu łącza. Publikacja RFC 1232 definiuje żółty alarm następująco:

> Żółty alarm jest zgłaszany w przypadku otrzymania żółtego sygnału z odległego końca łącza. W efekcie tego obwód staje się linkiem jednokierunkowym.

Jeden z producentów podaje, że żółty alarm jest uruchamiany dla linków *superframe*, kiedy szósty bit we wszystkich kanałach przyjmuje wartość zero przez co najmniej 335 ms, i zostanie wyłączony, gdy szósty bit w chociaż jednym kanale nie będzie miał wartości zero przez okres krótszy niż 1 – 5 sekund. Dla linków ESF żółty alarm jest uruchamiany, jeśli wzorzec sygnału pojawia się w co najmniej 7 – 10 kolejnych 16-bitowych interwałach, a wyłączany, kiedy wzorzec nie pojawia się w 10 kolejnych 16-bitowych interwałach.

Co za bełkot! W praktyce ta wiedza nie jest Ci potrzebna, jeśli nie musisz projektować podłączeń T1 CSU/DSU. Powinieneś wiedzieć tylko tyle: żółty alarm nie wskazuje koniecznie na problem z Twoim urządzeniem. Jest to raczej problem zgłaszany przez urządzenie, do którego jesteś podłączony.

Rysunek 21.8 przedstawia prostą sieć. Router A jest połączony linkiem T1 z routerem D. W rzeczywistości link T1 kończy się lokalnie w centrali usługodawcy T1, gdzie jest zazwyczaj agregowany w większy obwód łączący się z odległą centralą. Tam jest z powrotem konwertowany na link T1 dostarczany do zdalnej lokalizacji. Telekomunikacyjny przełącznik B działa poprawnie, ale przełącznik C ma awarię.



Rysunek 21.8. Kolejny przykład żółtego alarmu

Telekomunikacyjny przełącznik B otrzymuje sygnał żółtego alarmu od telekomunikacyjnego przełącznika C. Alarm ten może zostać przekazany do Routera A. Kiedy diagnozujesz przerwę w działaniu łącza, pamiętaj, że uruchomienie żółtego alarmu wskazuje zazwyczaj na problem z jakimś innym urządzeniem. W tym przypadku powodem przestoju jest telekomunikacyjny przełącznik C.

Uważaj, jakie czynisz założenia. Router D otrzyma żółty alarm. Żółty alarm może zostać wysłany również do Routera A. W takim przypadku administrator Routera A może szukać przyczyny problemu w awarii Routera D (i vice versa), ponieważ najprawdopodobniej nie będzie miał pojęcia o zaangażowaniu telekomunikacyjnych przełączników.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Alarm niebieski

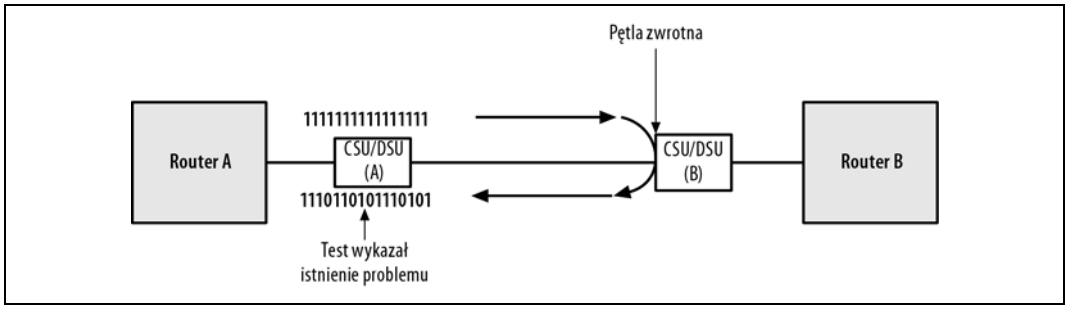
Alarm niebieski (ang. *blue alarm*) zwany jest również sygnałem wywołującym alarm (ang. *alarm indication signal* — AIS). Nie znajdziesz opisu tego alarmu w publikacji RFC. Alarm niebieski reprezentuje całkowity brak sygnału przychodzącego i jest wywoływany przez ciągły strumień jedynek bez struktury ramek. Możesz spotkać się z określeniem, że w przypadku niebieskiego alarmu interfejs otrzymuje „same jedynki". Jeśli otrzymasz niebieski alarm, jest duże prawdopodobieństwo, że rozłączony jest kabel lub zawiodło urządzenie.

Rozwiązywanie problemów z T1

Pierwszym krokiem w rozwiązywaniu kłopotów z linkami T1 jest określenie natury problemu. Zazwyczaj mamy do czynienia z kwestiami dotyczącymi okablowania, sprzętu oraz usług telekomunikacyjnych. Przeprowadzenie kilku prostych testów pozwoli Ci określić, jakie działania należy podjąć. Każdy z tych testów jest testem inwazyjnym, co oznacza, że musisz odłączyć usługę T1, aby je przeprowadzić.

Testy pętli zwrotnej

Test pętli zwrotnej (ang. *loopback test*) wymaga skonfigurowania jednego z urządzeń jako pętli zwrotnej, a następnie wysyłania danych tym łączem. Dane powinny wrócić w niezmienionym stanie. Jeśli dane nie powracają w takiej formie, w jakiej powinny, oznacza to, że coś po drodze je zmienia. Rysunek 21.9 przedstawia negatywny wynik testu pętli zwrotnej.



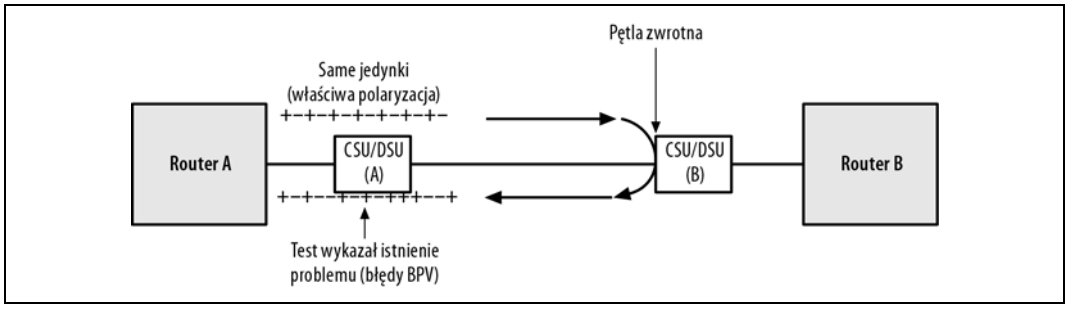
Rysunek 21.9. Koncepcyjne przedstawienie negatywnego wyniku testu pętli zwrotnej

Kiedy przeprowadzasz test pętli zwrotnej, jego negatywne wyniki nie będą się z reguły przedstawiały w tak prosty sposób, jak pokazana na rysunku zamiana jedynek w zera. Problem jest zazwyczaj raczej natury elektrycznej, tak jak w scenariuszu przedstawionym na rysunku 21.10. Ewentualnie błędy ramkowania mogą spowodować, że dane będą całkowicie nieczytelne.

Przy przeprowadzaniu testów pętli zwrotnej logicznym sposobem postępowania jest rozpoczęcie testu na jednym końcu łącza, a następnie poruszanie się wzdłuż łącza, aż pojawią się pierwsze symptomy.

Urządzenia CSU/DSU zasadniczo oferują opcje ustawienia kilku typów pętli zwrotnej, co jest bardzo przydatne w trakcie testów. Możesz skonfigurować pętlę zwrotną na interfejsie łączącym T1 z siecią telekomunikacyjną — **pętla zwrotna linii** (ang. *line loopback*), albo w miejscu, w którym sygnał minął już CSU/DSU — **pętla zwrotna bloku danych** (ang. *payload loopback*).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 21.10. Błędy BPV wykryte podczas testu pętli zwrotnej

Wiele urządzeń CSU/DSU umożliwia ustawianie pętli w obu kierunkach, co ułatwia wyizolowanie problemu.

Podczas diagnozowania problemów inżynierowie telekomunikacji zwykli mówić, że szukają **kłopotów** (ang. *trouble*) na linii. Może się to wydawać zbyt pedantyczne, ale kiedy posłuchasz ich przy pracy, usłyszysz z pewnością bardzo charakterystyczne określenia, takie jak np. **nie znaleziono kłopotów** (ang. *no trouble found* — NTF). Pamiętaj, że chodzi tu o standardy z ponad stuletnią historią.

Przeanalizujmy kolejny przykład. Na rysunku 21.11 Router A jest podłączony do Routera B linkiem T1 z wykorzystaniem dwóch urządzeń CSU/DSU. Możemy tutaj zobaczyć, jakie są możliwe punkty tworzenia pętli zwrotnej na urządzeniach CSU/DSU z perspektywy CSU/DSU (A).

Pamiętaj, że jeśli spojrzymy z perspektywy CSU/DSU (B), wszystkie punkty rozmieszczone będą odwrotnie.

Nie dla każdego modelu urządzenia CSU/DSU dostępne są wszystkie przedstawione tu punkty tworzenia pętli zwrotnej. Mogą się one również inaczej nazywać w zależności od modelu urządzenia.

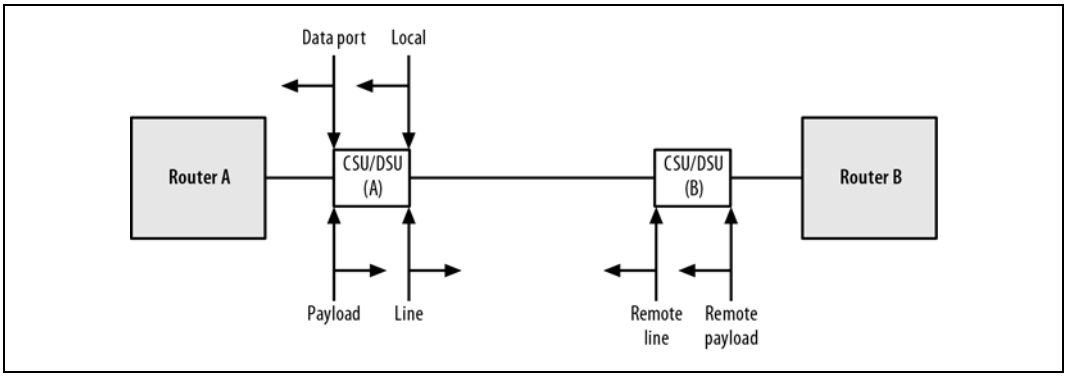
Poniższa lista zawiera terminy opisujące dostępne zazwyczaj punkty pętli zwrotnej. Pamiętaj, że te opisy bazują na strukturze sieci widzianej od strony Routera A na rysunku 21.11:

Data port/DTE
: W tym punkcie tworzona jest pętla dla sygnału pochodzącego z bezpośrednio podłączonego urządzenia służącego do przesyłania danych (w naszym przypadku jest to Router A). Sygnał kierowany jest z powrotem do tego urządzenia bez stosowania procesu ramkowania T1. W ten sposób testowany jest kabel V.35 oraz router.

Local (lokalna pętla zwrotna)
: W tym punkcie sygnał z bezpośrednio podłączonego urządzenia przesyłającego dane jest kierowany z powrotem do tego urządzenia z zastosowaniem ramkowania T1. Testowane jest w ten sposób fizyczne podłączenie do routera oraz CSU/DSU.

Payload (pętla zwrotna bloku danych)
: Ten punkt służy do zawracania sygnału z T1 z powrotem do T1 po zastosowaniu ramkowania T1. W naszym przykładzie test ten będzie administrowany po stronie urządzenia (B), jednak sama pętla będzie ustawiana lokalnie na CSU/DSU (A).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 21.11. Punkty pętli zwrotnych na urządzeniach CSU/DSU

Line (pętla zwrotna linii)
: W tym punkcie tworzona jest pętla kierująca sygnał z T1 z powrotem do T1 przed zastosowaniem ramkowania dla T1. Jest to skuteczny test linku T1 bez sprawdzania CSU/DSU. W naszym przykładzie pętla zwrotna będzie ustawiona lokalnie na CSU/DSU, ale test będzie przeprowadzany od strony urządzenia (B).

Remote line (zdalna pętla zwrotna linii)
: Pętla zwrotna *remote line* jest funkcją dostępną tylko na niektórych urządzeniach CSU/DSU. Umożliwia ona ustawienie przez lokalne CSU/DSU pętli zwrotnej linii na urządzeniu znajdującym się na odległym końcu. W naszym przykładzie komenda inicjująca pętlę zwrotną będzie wpisywana na CSU/DSU (A), mimo że sama pętla zostanie utworzona na CSU/DSU (B).

Remote payload (zdalna pętla zwrotna bloku danych)
: Pętla zwrotna *remote payload* jest funkcją dostępną tylko na niektórych urządzeniach CSU/DSU. Umożliwia ona ustawienie przez lokalne CSU/DSU pętli zwrotnej bloku danych na urządzeniu znajdującym się na odległym końcu. W naszym przykładzie komenda inicjująca pętlę zwrotną będzie wpisywana na CSU/DSU (A), mimo że sama pętla zostanie utworzona na CSU/DSU (B).

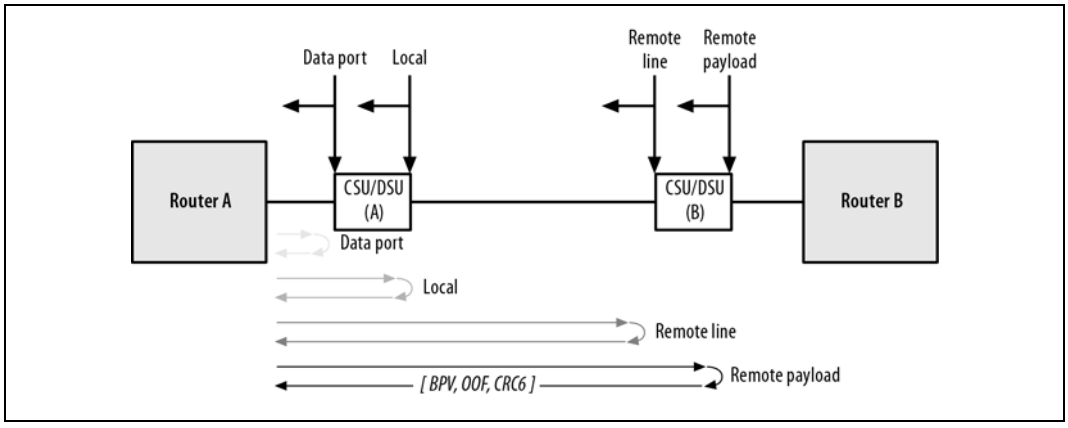
Załóżmy więc, że mamy problemy z linkiem T1 z rysunku 21.11 i musimy je zdiagnozować. Przeglądając statystyki błędów na CSU/DSU (A), zauważamy sporą liczbę błędów OOF, BPV oraz CRC6. Właściwe testy linii T1 będą zazwyczaj przebiegać tak, jak pokazano to na rysunku 21.12.

Zaczynamy od skonfigurowania pętli zwrotnej *data port* na CSU/DSU (A), a następnie uruchamiamy testy. Nie wykazują one żadnych błędów, co oznacza, że kabel V.35 jest w porządku.

Następnym krokiem jest usunięcie pętli *data port* na CSU/DSU (A) i ustawienie pętli zwrotnej *local*. Wykonujemy kolejne testy i wszystkie pakiety powracają bez błędów. Wyeliminowaliśmy już problemy z Routerem A, kablem V.35 oraz w 90 procentach problemy z CSU/DSU (A).

Teraz usuwamy pętlę zwrotną *local* na CSU/DSU i ustawiamy pętlę zwrotną *remote line* na CSU/DSU (B). Ponieważ jest to pętla zdalna, konfigurujemy ją z urządzenia CSU/DSU (A). Ewentualnie moglibyśmy poprosić drugą osobę, aby w lokalizacji (B) skonfigurowała ręcznie pętlę zwrotną linii na CSU/DSU (B). Uruchamiamy nasze testy i ponownie nie wykrywamy

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 21.12. Progresja testów pętli zwrotnej

żadnych błędów. W tym momencie mamy wyeliminowany Router A, kabel V.35 podłączony do Routera A, CSU/DSU (A) oraz samą linię T1 (łącznie z problemami leżącymi po stronie operatora sieci telekomunikacyjnej).

Kolejnym etapem jest usunięcie pętli *remote line* na CSU/DSU (B) i ustawienie pętli zwrotnej *remote payload* na CSU/DSU (B) — ponownie konfigurację przeprowadzamy z urządzenia CSU/DSU (A). Tym razem kiedy wykonujemy testy, CSU/DSU (A) raportuje znaczną liczbę błędów BPV, OOF oraz CRC6. Właśnie znaleźliśmy źródło problemów — CSU/DSU (B) nie działa prawidłowo. Systematycznie przesuwając punkty tworzenia pętli zwrotnej od początku łącza w stronę odległego końca, byliśmy w stanie określić punkt, w którym zaczęły pojawiać się problemy. Wymiana CSU/DSU (B) rozwiązuje sprawę i możemy przywrócić prawidłowe działanie łącza.

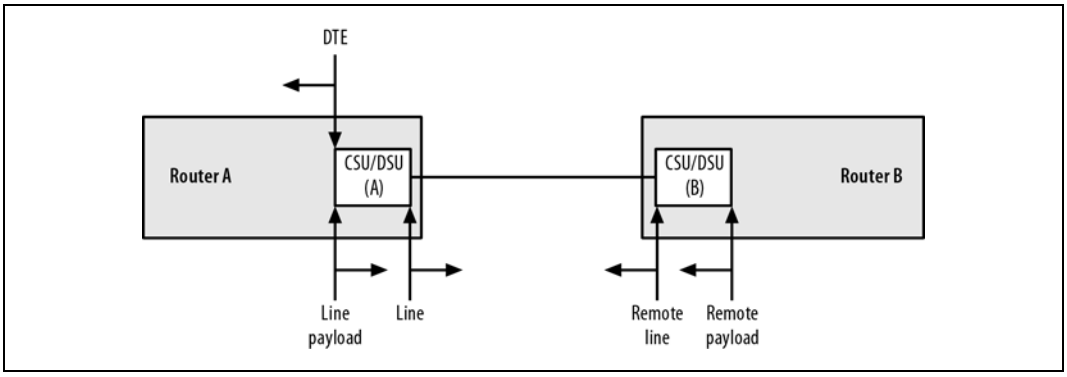
Zintegrowane CSU/DSU

Karty interfejsu T1 WAN (ang. *WAN interface cards* — WIC) ze zintegrowanymi CSU/DSU są najfajniejszym rozszerzeniem funkcji routerów wprowadzonym w ciągu ostatnich 15 lat. Możecie nazwać mnie świrem, ale koncepcja wyeliminowania kolejnego fizycznego urządzenia (łącznie z pozbyciem się tych okropnych kabli V.35 z szafy rackowej) niezbędnego dla każdej instalacji T1 jest czymś najbardziej według mnie zbliżonym do raju maniaków komputerowych.

CSU/DSU zintegrowane w kartach WIC służą tym samym celom co samodzielne jednostki tego typu. Ich dodatkową zaletą jest konfiguracja z poziomu IOS samego routera. Rysunek 21.13 przedstawia punkty tworzenia pętli zwrotnych dla T1 CSU/DSU WIC w taki sposób, w jaki zostały one opisane w IOS.

Niektóre CSU/DSU posiadają nawet funkcję umożliwiającą przeprowadzanie testów BERT. Może być to użyteczne, ale przeprowadzając te testy, powinieneś prawdopodobnie skontaktować się z dostawcą usług telekomunikacyjnych. Większość problemów wykrytych w testach BERT nie może być usunięta poprzez konfigurację routera.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 21.13. Punkty pętli zwrotnych dla zintegrowanych CSU/DSU

Konfiguracja T1

Konfiguracja T1 na routerze składa się z dwóch etapów. Pierwszym krokiem jest konfiguracja urządzeń CSU/DSU. Drugim etapem jest konfiguracja interfejsu routera. Jeśli posiadasz zintegrowane CSU/DSU, podział ten może się nieco zacierać, ale koncepcja pozostaje ta sama. Konfiguracja interfejsu routera przypomina konfigurację dowolnego interfejsu szeregowego.

Konfiguracja CSU/DSU

Aby uruchomić w pełni funkcjonalne łącze T1, musisz wykonać następujące czynności:

Skonfiguruj na obu stronach łącza takie samo kodowanie, odpowiadające kodowaniu dostarczonemu dla obwodu

Dostępne opcje kodowania to AMI oraz B8ZS. Dla T1 służącego do transmisji danych zawsze należy stosować kodowanie B8ZS. Konfiguracja kodowania na CSU/DSU WIC odbywa się za pomocą polecenia interfejsu `service-module t1 linecode`:

```
Router(config)#int s0/1
Router(config-if)# service-module t1 linecode b8zs
```

Skonfiguruj na obu stronach łącza takie samo ramkowanie, odpowiadające ramkowaniu dostarczonemu dla obwodu

Dostępne opcje ramkowania to D4/SF oraz ESF. Dla T1 służącego do transmisji danych zawsze należy stosować ramkowanie ESF. Konfiguracja ramkowania na CSU/DSU WIC odbywa się za pomocą polecenia interfejsu `service-module t1 framing`:

```
Router(config)#int s0/1
Router(config-if)# service-module t1 framing esf
```

Skonfiguruj liczbę kanałów, które zostaną wykorzystane dla linku, określ, które to będą kanały, oraz zdefiniuj ich prędkość

Jeśli link T1 został podzielony i dostarczono Ci mniej niż 24 kanały, musisz poinstruować CSU/DSU, ile kanałów jest wykorzystywanych. W tym celu dla CSU/DSU WIC stosujesz komendę interfejsu `service-module t1 timeslots`. Poniżej zdefiniowałem kanały 7 – 12 o prędkości 64 kb/s:

```
Router(config)#int s0/1
Router(config-if)# service-module t1 timeslots 7-12 speed 64
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Domyślnie wykorzystywane są wszystkie kanały i mają one prędkość 64 kb/s. Jeśli musisz przywrócić domyślną konfigurację wszystkich kanałów, skorzystaj z argumentu `all`:

```
Router(config)#int s0/1
Router(config-if)# service-module t1 timeslots all speed 64
```

Skonfiguruj jedną stronę łącza jako clock master, a drugą jako slave

Ponieważ łącza T1 są asynchroniczne, tylko jedna strona może posiadać aktywny zegar taktujący (ang. *active clock*). Druga strona będzie odczytywać taktowanie z samego strumienia danych, wykorzystując technologię **zamkniętej pętli fazowej** (ang. *phase-locked loop* — PLL). Aby skonfigurować taktowanie na CSU/DSU WIC, zastosuj polecenie interfejsu `service-module t1 clock source`. Dostępne argumenty to `internal` oraz `line`:

```
Router(config)#int s0/1
Router(config-if)# service-module t1 clock source internal
```

Argument `internal` oznacza, że to CSU/DSU będzie zapewniać taktowanie (master), a argument `line` definiuje odczyt taktowania ze strumienia danych na linii (slave). Domyślnym zachowaniem jest odczytywanie taktowania z linii.

W niektórych środowiskach może być wymagane ustawienie taktowania `line` na obu końcach. Jeśli nie masz pewności, jakie są wymagania taktowania dla Twojego łącza T1, skonsultuj się z usługodawcą.

Diagnozowanie problemów CSU/DSU

Zintegrowanie w routerze CSU/DSU jest doskonałym ulepszeniem w stosunku do samodzielnych urządzeń CSU/DSU. Możliwość wykorzystania sesji telnet dla połączenia się z CSU/DSU jest niezastąpionym udogodnieniem w momencie wystąpienia problemów z łączem. Samodzielne jednostki CSU/DSU często posiadają porty szeregowe, pozwalające na podłączenie serwerów konsoli, jednak w przeciętnych środowiskach korporacyjnych ta funkcja rzadko jest wykorzystywana. Ponadto wiele przedsiębiorstw stosuje różne marki i modele urządzeń CSU/DSU, a każde z nich posiada własne, charakterystyczne menu, polecenia, funkcje oraz dziwactwa.

Cisco T1 CSU/DSU WIC umożliwia przeglądanie statystyk CSU/DSU dzięki wpisywaniu komend w sesji telnet z routerem. Podstawową komendą diagnozowania problemów T1 CSU/DSU jest `show service-module` *`interfejs`*. Komenda ta generuje mnóstwo informacji dotyczących statusu CSU/DSU oraz ogólnie obwodu T1.

Przyjrzyjmy się listingowi tej komendy, gdy druga strona T1 nie jest podłączona:

```
Router#sho service-module s0/1
Module type is T1/fractional
    Hardware revision is 0.112, Software revision is 0.2,
    Image checksum is 0x73D70058, Protocol revision is 0.1
Transmitter is sending remote alarm.
Receiver has loss of frame,
Framing is ESF, Line Code is B8ZS, Current clock source is line,
Fraction has 24 timeslots (64 Kbits/sec each), Net bandwidth is 1536 Kbits/sec.
Last user loopback performed:
    dte loopback
    duration 08:40:48
Last module self-test (done at startup): Passed
Last clearing of alarm counters 08:45:16
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
    loss of signal        :    1, last occurred 08:45:07
    loss of frame         :    2, current duration 00:01:38
    AIS alarm             :    0,
    Remote alarm          :    0,
    Module access errors  :    0,
Total Data (last 34 15 minute intervals):
    2 Line Code Violations, 0 Path Code Violations
    1 Slip Secs, 200 Fr Loss Secs, 2 Line Err Secs, 0 Degraded Mins
    0 Errored Secs, 0 Bursty Err Secs, 0 Severely Err Secs, 200 Unavail Secs
Data in current interval (896 seconds elapsed):
    255 Line Code Violations, 255 Path Code Violations
    32 Slip Secs, 109 Fr Loss Secs, 34 Line Err Secs, 0 Degraded Mins
    0 Errored Secs, 0 Bursty Err Secs, 0 Severely Err Secs, 116 Unavail Secs
```

Możemy tutaj wyczytać, że CSU/DSU wysyła do T1 sygnał zdalnego alarmu (żółty alarm), ponieważ otrzymuje błędy utraty ramki (LOF). Co ważniejsze, pojawiło się jedno zdarzenie utraty sygnału (LOS), które miało miejsce około 8 godzin i 45 minut temu. Ten obwód T1 nie działa od czasu uruchomienia routera. Statystyki alarmów były zerowane ostatnio 8 godzin i 45 minut temu.

Informacje te zasadniczo mówią nam, że po drugiej stronie naszego obwodu nic nie ma. Z całą pewnością router po odległej stronie łącza jest wyłączony. Posiada on również zintegrowane CSU/DSU. Po włączeniu tego routera wyzerujmy liczniki i sprawdźmy, jak wygląda moduł usług:

```
Router#clear counters s0/1
Clear "show interface" counters on this interface [confirm]
09:00:04: %CLEAR-5-COUNTERS: Clear counter on interface Serial0/1 by console
Router#
Router#sho service-module s0/1
Module type is T1/fractional
    Hardware revision is 0.112, Software revision is 0.2,
    Image checksum is 0x73D70058, Protocol revision is 0.1
Receiver has no alarms.
Framing is ESF, Line Code is B8ZS, Current clock source is line,
Fraction has 24 timeslots (64 Kbits/sec each), Net bandwidth is 1536 Kbits/sec.
Last user loopback performed:
    dte loopback
    duration 08:40:48
Last module self-test (done at startup): Passed
Last clearing of alarm counters 00:03:01
    loss of signal        :    0,
    loss of frame         :    0,
    AIS alarm             :    0,
    Remote alarm          :    0,
    Module access errors  :    0,
Total Data (last 96 15 minute intervals):
    258 Line Code Violations, 257 Path Code Violations
    33 Slip Secs, 309 Fr Loss Secs, 37 Line Err Secs, 1 Degraded Mins
    1 Errored Secs, 1 Bursty Err Secs, 0 Severely Err Secs, 320 Unavail Secs
Data in current interval (153 seconds elapsed):
    0 Line Code Violations, 0 Path Code Violations
    0 Slip Secs, 0 Fr Loss Secs, 0 Line Err Secs, 0 Degraded Mins
    0 Errored Secs, 0 Bursty Err Secs, 0 Severely Err Secs, 0 Unavail Secs
```

Ten listing wygląda znacznie lepiej. Zawiera on mile widzianą frazę „receiver has no alarms" (brak alarmów na odbiorniku). Listing wskazuje również, że przed chwilą wyzerowaliśmy liczniki, a od tego czasu nie pojawiły się żadne błędy ani alarmy.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ostatnie dwa punkty listingu są szczególnie ważne. CSU/DSU śledzi z reguły wszystkie zdarzenia, które wystąpiły w ciągu ostatnich 24 godzin. Zdarzenia te są zapisywane i raportowane w 15-minutowych interwałach. Pierwszy punkt (`Total Data`), choć alarmujący, nie jest tak istotny jak punkt kolejny (`Data in current interval`), który przedstawia zdarzenia dla bieżącego interwału. Może to być nieco mylące, przyjrzyjmy się więc temu nieco bliżej.

Stosując komendę `show service-module` *`interfejs`* `performance-statistics`, możemy sprawdzić, które zdarzenia miały miejsce w każdym z ostatnich 96 interwałów 15-minutowych:

```
Router#sho service-module s0/1 performance-statistics
Total Data (last 96 15 minute intervals):
    258 Line Code Violations, 257 Path Code Violations
    33 Slip Secs, 309 Fr Loss Secs, 37 Line Err Secs, 1 Degraded Mins
    1 Errored Secs, 1 Bursty Err Secs, 0 Severely Err Secs, 320 Unavail Secs
Data in current interval (380 seconds elapsed):
    0 Line Code Violations, 0 Path Code Violations
    0 Slip Secs, 0 Fr Loss Secs, 0 Line Err Secs, 0 Degraded Mins
    0 Errored Secs, 0 Bursty Err Secs, 0 Severely Err Secs, 0 Unavail Secs
Data in Interval 1:
    1 Line Code Violations, 2 Path Code Violations
    0 Slip Secs, 0 Fr Loss Secs, 1 Line Err Secs, 0 Degraded Mins
    1 Errored Secs, 1 Bursty Err Secs, 0 Severely Err Secs, 0 Unavail Secs
Data in Interval 2:
    255 Line Code Violations, 255 Path Code Violations
    32 Slip Secs, 109 Fr Loss Secs, 34 Line Err Secs, 0 Degraded Mins
    0 Errored Secs, 0 Bursty Err Secs, 0 Severely Err Secs, 116 Unavail Secs
Data in Interval 3:
    0 Line Code Violations, 0 Path Code Violations
    0 Slip Secs, 0 Fr Loss Secs, 0 Line Err Secs, 0 Degraded Mins
    0 Errored Secs, 0 Bursty Err Secs, 0 Severely Err Secs, 0 Unavail Secs
Data in Interval 4:
    0 Line Code Violations, 0 Path Code Violations
0 Slip Secs, 0 Fr Loss Secs, 0 Line Err Secs, 0 Degraded Mins
 0 Errored Secs, 0 Bursty Err Secs, 0 Severely Err Secs, 0 Unavail Secs
```

Pierwszy punkt przedstawia podsumowanie dla wszystkich interwałów znajdujących się w pamięci urządzenia. Maksymalny okres, dla którego CSU/DSU będzie przechowywał informacje o zdarzeniach, to 24 godziny (96 15-minutowych interwałów). Najstarszy interwał, niemieszczący się w tym przedziale, zostaje zawsze usuwany na bieżąco z pamięci. Listing skonstruowany jest na podstawie numerów interwałów, a nie czegoś bardziej oczekiwanego, jak właściwy czas. Jest to uwstecznienie w stosunku do sposobu raportowania informacji historycznych przez samodzielne urządzenia CSU/DSU. Pierwszy na liście jest trwający właśnie interwał (`current interval`). Następny interwał, oznaczony jako **Interval 1**, jest najbardziej aktualnym zakończonym interwałem. Numeracja interwałów rośnie do momentu osiągnięcia numeru 96.

Analizując te informacje, szukaj powtarzających się układów. Jeśli zauważysz występowanie pewnej liczby błędów `line code violations` lub dowolnych innych błędów, które powtarzają się w większości interwałów albo nawet we wszystkich interwałach — znalazłeś problem. Błędy wystąpią, jeśli obwód T1 jest niestabilny z jakiegoś powodu. Interwał 2 pokazuje pewne poważne błędy, które wynikają z wypięcia kabla T1 z gniazdka.

Jeśli natrafisz na rosnącą liczbę błędów, zacznij diagnozować problem. I pamiętaj: najpierw sprawdź struktury fizyczne! Większość problemów jest spowodowanych usterką kabla lub urządzenia.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 22.

DS3

Standard strumienia cyfrowego DS3 (*Digital Signal 3*) zamierzam przedstawić w nieco inny sposób niż informacje na temat łączy T1. O ile mechanizmy ramkowania i sygnalizacji w rozwiązaniach T1 wydają mi się bardzo istotne dla inżyniera sieciowego, nie sądzę, żeby miały szczególne znaczenie w przypadku systemów DS3. W przeciwieństwie do tego, co wiele osób twierdzi (często z uporem), strumień DS3 nie jest definiowany jako 28-krotność strumienia DS1. Przepływność DS3 jest w rzeczywistości wynikiem multipleksacji siedmiu sygnałów DS2. Jeśli pomyślałeś sobie właśnie „nie ma czegoś takiego jak DS2", nie jesteś sam. Strumień DS2 jest grupą czterech sygnałów DS1, ale nie jest wyprowadzany poza system zwielokrotniania.

Choć informacja o tym, że strumień DS3 składa się z kilku strumieni DS2, może się wydać interesująca, nie pomoże w projektowaniu lub uruchamianiu sieci komputerowej. Dlatego w dalszej części rozdziału omówione zostały jedynie te zagadnienia, które są naprawdę istotne: ogólne wprowadzenie teoretyczne, przypadki błędów, zasady konfiguracji oraz sposoby usuwania usterek.

Standard DS3 nie jest tożsamy ze specyfikacją T3. Standard DS3 definiuje logiczną nośną przesyłaną w fizycznym łączu T3. W praktyce określenia te stosuje się wymiennie — większość osób nie zauważy różnicy, jeśli posłużysz się jednym z nich. Niemniej od tego miejsca będę używał terminu DS3 do określania obwodu transmisji danych, gdyż tak naprawdę nie interesuje nas fizyczne medium transmisyjne.

Obwody DS3 występują w dwóch odmianach: podzielone na kanały (ang. *channelized*) oraz bezkanałowe (ang. *clear-channel*). **Podzielony na kanały sygnał DS3** składa się z 672 strumieni DS0, z których każdy może przenosić jedną rozmowę telefoniczną systemu POTS. W terminologii Cisco taki sygnał DS3 jest określany jako „kanałowe łącze T3" (ang. *channellized T3*). **Bezkanałowy sygnał DS3** jest wykorzystywany w całości do przenoszenia danych bez dodatkowego podziału na poszczególne kanały[1].

[1] Niniejszy rozdział poświęcony jest w całości standardowi DS3, który jest systemem zwielokrotniania kanałów cyfrowych stosowanym w Ameryce Północnej. Jego europejskim odpowiednikiem jest standard E3 o przepływności 34 368 Mb/s — *przyp. tłum.*

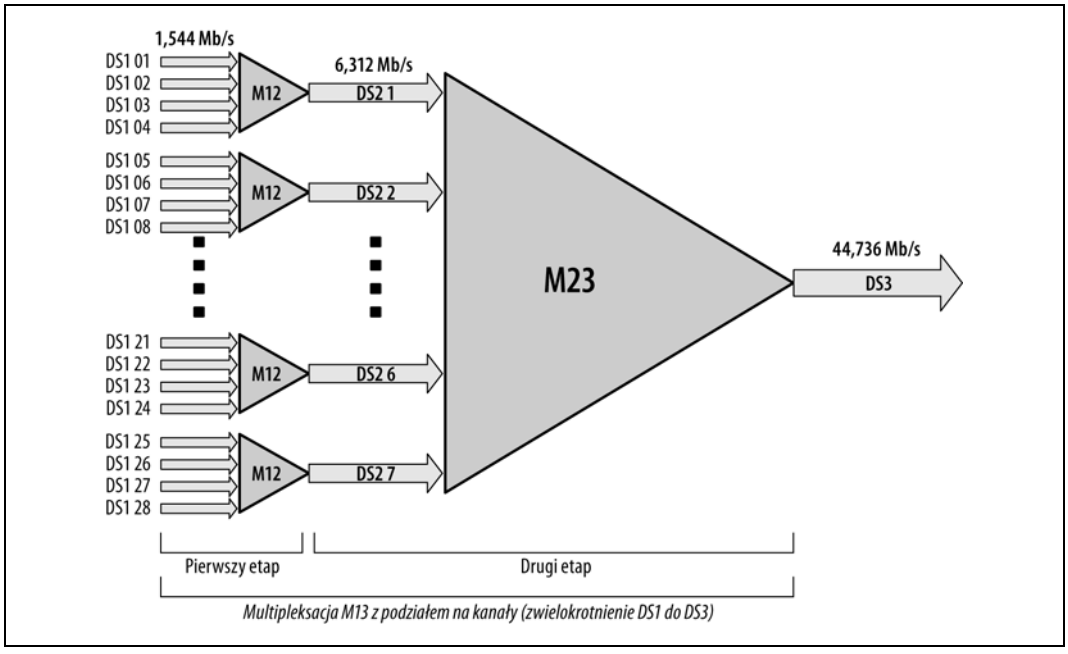
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ramkowanie

Zgodnie z wcześniejszymi informacjami strumień DS3 jest w rzeczywistości wynikiem multipleksacji siedmiu strumieni DS2. Wspominałem również o podziale sygnału DS3 na kanały. W latach 90. ubiegłego stulecia, gdy opracowywano standard, nie było takiego zapotrzebowania na obwody danych, jakie obserwujemy obecnie. Ramkowanie w strumieniu DS3 było więc niezbędne z tych samych powodów, z których stosuje się je w sygnałach DS1. Różnica między wymienionymi rozwiązaniami jest jedynie taka, że w pojedynczym strumieniu DS3 może być przekazywanych wiele sygnałów DS1. Każdy ze strumieni DS1 ma własny system zegarowy, ramkowania i kodowania, które trzeba zachować podczas transmisji w łączu DS3. Oczywiście system DS3 musi również dysponować własnymi mechanizmami generowania sygnału zegarowego, formowania ramek oraz kodowania danych. Co więcej, mechanizmy te nie mogą kolidować z rozwiązaniami stosowanymi w sygnałach składowych. Administrator sieci może wybrać jedną z wielu metod ramkowania sygnału zależnie od przeznaczenia obwodu DS3.

M13

Nazwa M13 jest skrótem od słów *Multiplexed DS1 to DS3*, które odnoszą się do multipleksacji (zwielokrotniania) sygnałów DS1 i generowania wynikowego sygnału DS3. Proces zwielokrotniania składa się z dwóch etapów — M12 i M23. Połączenie tych operacji jest ogólnie określane jako multipleksacja M13. Zasada generowania sygnału DS3 na podstawie 28 sygnałów DS1 została zilustrowana na rysunku 22.1.



Rysunek 22.1. Multipleksacja M13

Pierwotnie strumienie DS3 były wykorzystywane do agregowania sygnałów T1. Oznaczało to, że do centrali telefonicznej docierało 28 sygnałów T1, z których każdy był generowany przez

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

urządzenia pracujące w innych lokalizacjach o różnych odległościach od siedziby operatora. Ponieważ poszczególne strumienie T1 nie były ze sobą skorelowane, istniało duże prawdopodobieństwo, że nie będą zsynchronizowane. Projektanci systemu uwzględnili ten fakt i dlatego wydzielili pierwszą fazę przetwarzania, w której opisany problem jest rozwiązywany.

Przepustowość łącza T1 to 1,544 Mb/s. Jeśli pomnożymy 1,544 Mb/s przez 4, otrzymamy 6,176 Mb/s. Dlaczego więc przepływność strumienia DS2 (który odpowiada czterem strumieniom T1) wynosi 6,312 Mb/s? Aby zrównoważyć dostarczanie rozbieżności czasowej w dostarczaniu danych T1 (+/– 77 Hz), multiplekser M12 wstawia do strumienia dodatkowe bity, które pozwalają na ujednolicenie przepływności wszystkich sygnałów T1 w grupie.

Po **nadziewaniu bitami** (ang. *bit stuffing*) każdy sygnał T1 ma przepływność 1 545 796 b/s. W ciągu sekundy dodawanych jest 128 816 bitów ramkowania i narzutu, które zwiększają przepływność sumaryczną do 6,312 Mb/s (1 545 796·4+128 816). Po stronie odbiorczej nadmiarowe bity są usuwane przez demultiplekser.

Przetaktowywanie (ang. *overclocking*) strumienia T1 eliminuje ryzyko problemów z synchronizacją sygnałów DS3. Pamiętaj, że sygnał T1 ma własny zegar nadrzędny i podrzędny. Strumień DS3 zawiera natomiast własny mechanizm zegarowy, który nie wpływa w żaden sposób na synchronizację poszczególnych sygnałów T1 (w nowoczesnych sieciach bazujących na rozwiązaniach SONET ten problem w ogóle nie występuje, ale trzeba pamiętać, że opisywana technologia została opracowana wiele lat przed powstaniem sieci SONET).

Bity C

Ramkowanie M13 jest nieco przestarzałe. Zakłada się w nim, że łącza DS2 mogą być doprowadzane do zdalnych lokalizacji, tak jak DS1. W praktyce nie są one jednak wdrażane, a strumienie DS2 są przetwarzane jedynie w multiplekserach.

Oznacza to, że problem synchronizacji, który wymaga wstawiania dodatkowych bitów, występuje jedynie na etapie M12, a nigdy w fazie M23. Mechanizm M13 rezerwuje jednak pewne pozycje bitowe na operację nadziewania bitami w czasie multipleksacji M23.

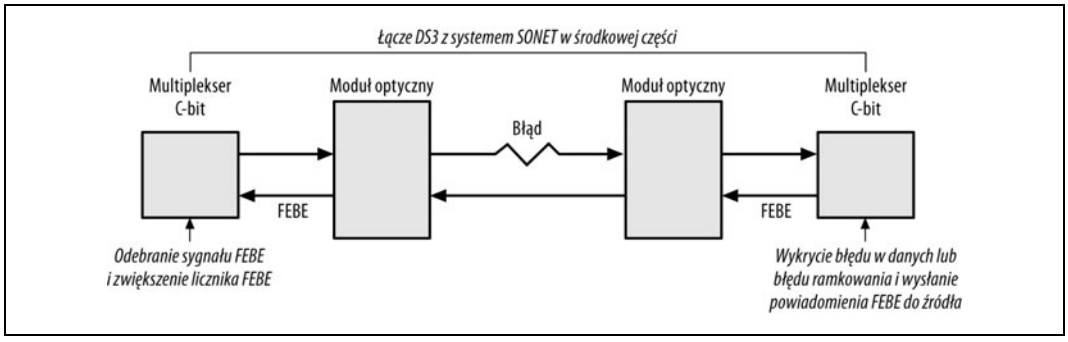
Aby wykorzystać nieużywane bity, opracowano jeszcze jedną technikę ramkowania. W pierwotnym rozwiązaniu pewne bity (nazywane **bitami C**) służyły do sygnalizowania obecności bitów dodawanych na etapie M23. Technika **ramkowania typu C-bit** (ang. *C-bit framing*) wykorzystuje je nieco inaczej.

Służą one do przekazywania informacji o błędach blokowych zarejestrowanych po drugiej stronie łącza (FEBE — ang. *Far-End Block Errors*). Pola FEBE są specjalnymi komunikatami alarmowymi strumienia DS3, które informują, że węzeł zdalny zarejestrował błąd parzystości w bicie C lub błąd ramkowania. Na rysunku 22.2 przedstawiono sposób przekazywania wartości FEBE w strumieniu DS3, który w środkowej części łącza jest przenoszony przez system SONET.

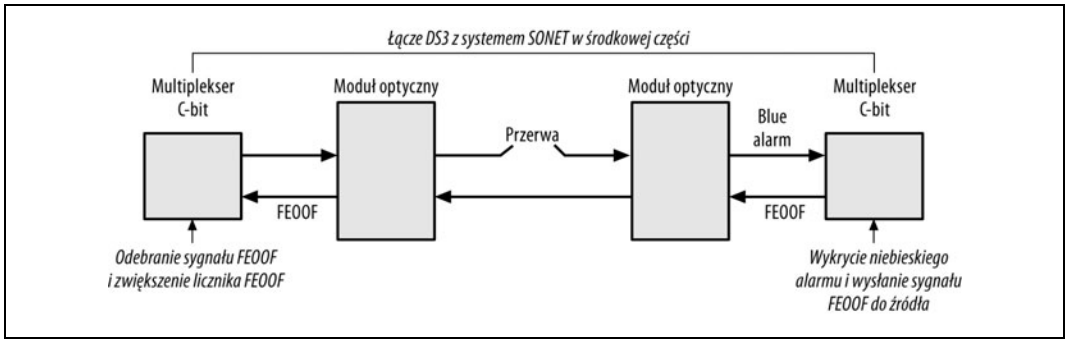
Ramkowanie z wykorzystaniem bitów C umożliwia również przekazywanie informacji o przerwaniu ramki po stronie zdalnej (FEOOF — ang. *Far-End Out-Of-Frame*). Gdy urządzenie zdalne wykryje przerwę w odbiorze danych, wysyła do źródła strumienia sygnał FEOOF. Działanie sygnalizacji FEOOF zostało przedstawione na rysunku 22.3.

Wykorzystywane są również inne kody alarmowe i sterujące generowane przez stację zdalną (FEAC — ang. *Far-End Alarm and Control*). Oto ich lista:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 22.2. Błędy blokowe rejestrowane w zdalnej stacji



Rysunek 22.3. Błędy końca ramki po stronie zdalnej

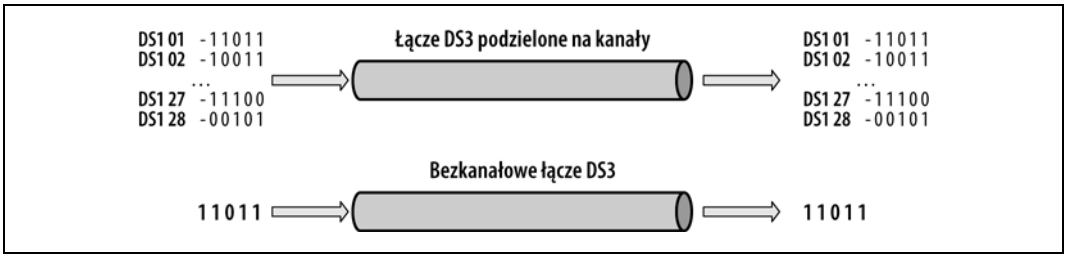
- usterka sprzętowa łącza DS3 mająca wpływ na usługę (SA — ang. *Service Affecting*);
- DS3 LOS/HBER;
- brak ramki DS3;
- odbiór sygnału AIS w łączu DS3;
- odbiór sygnału IDLE w łączu DS3;
- usterka sprzętowa łącza DS3 niemająca wpływu na usługę (NSA — ang. *Non Service Affecting*);
- typowa usterka sprzętowa (NSA);
- wielokrotne błędy DS1 LOS/HBER;
- usterka sprzętowa DS1 (SA);
- pojedynczy błąd DS1 LOS/HBER;
- usterka sprzętowa DS1 (NSA).

Kody FEAC są prezentowane w zestawieniu wynikowym polecenia `show controllers`.

Ramkowanie bezkanałowego strumienia DS3

W rozwiązaniu tym warstwa łącza danych wykorzystuje całą pojemność strumienia DS3 bez dzielenia jej na kanały (zgodnie z rysunkiem 22.4). Ramkowanie jest jednak nadal potrzebne. Do utrzymania synchronizacji między stacjami końcowymi można wykorzystać mechanizm C-bit

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 22.4. Łącze DS3 z podziałem na kanały i bezkanałowe

lub M13. Etapy M12 i M23 nie są potrzebne i nie wprowadzają dodatkowego narzutu bitowego. Zajmowane przez nie bity są więc wykorzystywane do przenoszenia danych użytkowych.

Dzięki wyeliminowaniu multipleksacji szerokość pasma bezkanałowego strumienia DS3 wynosi 44,2 Mb/s.

Używając standardu DS3 w łączach przeznaczonych do przesyłania danych, korzystaj z ramkowania C-bit, ponieważ obejmuje ono mechanizm zgłaszania błędów (opisany wcześniej).

Kodowanie liniowe

Łącza DS3 wykorzystują kodowanie liniowe AMI, B3ZS oraz HDB3. Kodowanie AMI przebiega dokładnie tak samo jak w łączu T1 (opisanym w rozdziale 21.). Mechanizm B3ZS jest zbliżony do B8ZS, który również został przedstawiony w rozdziale 21. Jedyna różnica polega na zastępowaniu trzech kolejnych zer (a nie ośmiu kolejnych zer) znanym stanem bipolarnym. System HDB3 jest stosowany przede wszystkim w Europie i Japonii. Domyślnym kodowaniem w interfejsach DS3 urządzeń Cisco jest B3ZS. W przypadku podziału na kanały wykorzystuje się często mechanizm AMI, ale jest to zależne od umowy z dostawcą usługi. Zazwyczaj jednak preferowanym rozwiązaniem jest B3ZS.

Konfiguracja łączy DS3

Zgodnie z informacjami zamieszczonymi na początku rozdziału łącza DS3 występują w dwóch odmianach: bezkanałowej i uwzględniającej podział na kanały. Wersja bezkanałowa jest przeznaczona do obsługi łączy danych, natomiast podział na kanały znajduje zastosowanie w połączeniach głosowych.

Prezentowana dalej konfiguracja odnosi się do urządzeń ze zintegrowanymi modułami CSU/DSU. Jeśli masz starszy router, który wymaga podłączenia zewnętrznej jednostki CSU/DSU, prawdopodobnie będzie on połączony z routerem za pomocą interfejsu HSSI, który dla systemu IOS jest po prostu interfejsem szeregowym. Nie można wówczas ustawić parametrów ramkowania ani sprawdzić błędów łącza DS3 (z uwagi na przyłączenie modułu CSU/DSU do portu HSSI jako zewnętrznego urządzenia).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Bezkanałowe łącze DS3

Konfiguracja bezkanałowego łącza DS3 jest wyjątkowo nudnym zadaniem. Do definiowania ramkowania służy polecenie `framing`. Sam interfejs parametryzuje się tak samo jak każdy inny interfejs szeregowy:

```
interface Serial3/1/0
 description DS3
 ip address 10.100.100.100 255.255.255.252
 framing c-bit
```

Sprawdzenie stanu interfejsu również nie odbiega w żaden sposób od weryfikacji innych interfejsów. Informacje o błędach oraz liczniki mają charakter ogólny, tzn. nie są szczególne dla standardu DS3 lub rozwiązań operatora telekomunikacyjnego.

```
7304#sho int s3/1/0
Serial3/1/0 is up, line protocol is up
  Hardware is SPA-4XT3/E3
  Description: DS3
  Internet address is 10.100.100.100/30
  MTU 4470 bytes, BW 44210 Kbit, DLY 200 usec,
  reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation HDLC, crc 16, loopback not set
  Keepalive set (10 sec)
  Last input 00:00:02, output 00:00:01, output hang never
  Last clearing of "show interface" counters 6d03h
  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  5 minute input rate 101000 bits/sec, 179 packets/sec
  5 minute output rate 98000 bits/sec, 170 packets/sec
    81589607 packets input, 2171970011 bytes, 0 no buffer
    Received 61914 broadcasts (229394 IP multicast)
    1072 runs, 0 giants, 0 throttles
             0 parity
    1136 input errors, 10 CRC, 0 frame, 0 overrun, 0 ignored, 54 abort
    80620669 packets output, 1165631313 bytes, 0 underruns
    0 output errors, 0 applique, 0 interface resets
    0 output buffer failures, 0 output buffers swapped out
    0 carrier transitions
```

Ciekawszych danych dostarcza instrukcja `show controllers`:

```
7304#show controllers s3/1/0
Interface Serial3/1/0 (DS3 port 0)
  Hardware is SPA-4XT3/E3
   Framing is c-bit, Clock Source is Line
   Bandwidth limit is 44210, DSU mode 0, Cable length is 10 feet
   rx FEBE since last clear counter 792, since reset 3693

   No alarms detected.

   No FEAC code is being received
   MDL transmission is disabled
   PXF interface number = 0x12

 SPA carrier card counters:
   Input: packets = 81583462, bytes = 6466441470, drops = 7
   Output: packets = 80614617, bytes = 5460208896, drops = 0
   Egress flow control status: XON
   Per bay counters:
   General errors: input = 0, output = 0
   SPI4 errors: ingress dip4 = 0, egress dip2 = 0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  SPA FPGA Packet Counters:
    Transmit : 80614638, Drops : 0
    Receive : 81583490, Drops : 0

  SPA FPGA Invalid Channel Packets:
    Transmit : 0, Receive : 0

  SPA FPGA IPC Counters:
    Transmit : 1057496, Drops : 0
    Receive : 1057496, Drops : 0

  SPA FPGA Packet Error Counters:
    202 Receive error packets
Framer(PM5383) Counters:
    Transmit : 80614555 packets, 1165231422 bytes
    Errors : 0 aborts, 0 underruns

    Receive : 81583399 packets, 2171463422 bytes
    Errors : 10 crc, 1072 runts, 0 giants, 54 aborts
```

Powyższy listing opisuje poprawnie działające bezkanałowe łącze DS3. Ramkowanie jest realizowane zgodnie ze standardem C-bit.

Zanotowano pewną liczbę sygnałów FEBE. Niewielki przyrost wartości licznika powiadomień FEBE w dłuższym czasie nie ma istotnego wpływu na transmisję danych. Z wcześniejszego zestawienia wynika, że liczniki zostały wyzerowane sześć dni przed wykonaniem instrukcji. Od tego czasu router zarejestrował 792 sygnały FEBE, oznacza to (przy założeniu równomiernego rozkładu) 5,5 wystąpienia w ciągu godziny. Należałoby więc wyzerować liczniki ponownie i poobserwować działanie interfejsu. Jeśli wartości licznika FEBE przyrastają regularnie, warto rozważyć dalsze testy i poszukiwanie przyczyn problemu.

Z zestawienia `show controllers` wynika również, że nie ma żadnych aktywnych błędów i nie odebrano też żadnych kodów FEAC. Na podstawie tych informacji można stwierdzić, że łącze jest w dobrym stanie.

W modułach PA-A3-T3 oraz NM-1A-T3 występuje znany problem zbyt dużej czułości odbiornika (zobacz wpis CSCds15318 w rejestrze błędów Cisco). Jeśli urządzenia są połączone za pomocą krótkiego kabla, mogą rejestrować dużą liczbę błędów. Oprócz wymiany interfejsu (na inny model) firma Cisco zaleca ograniczenie poziomu transmitowanego sygnału (w urządzeniu pracującym po drugiej stronie kabla) lub zainstalowanie tłumików. Wymienione karty są dość powszechnie wykorzystywane, więc warto zwrócić uwagę na opisany problem.

Łącze DS3 podzielone na kanały

Podzielone na kanały łącze DS3 można przeznaczyć do transmisji głosu, danych lub obydwu form informacji jednocześnie. Dzięki podziałowi na kanały pojedyncze strumienie T1 można sklasyfikować jako strumienie danych lub łącza głosowe. Odpowiadają za to instrukcje konfiguracji kontrolera.

W kolejnym przykładzie ustawiłem pierwsze 10 strumieni T1 łącza DS3 jako połączenia szeregowe przeznaczone do transmisji danych. W tym celu przypisałem odpowiednią liczbę kanałów DS0 do **grupy kanałów** (ang. *channel group*). Ponieważ konfigurowane łącza składają się z całych strumieni T1, do każdej grupy kanałów przypisałem wszystkie 24 kanały DS0 (opisywane jako **szczeliny czasowe** (ang. *timeslot*)). Ponieważ operacja odnosi się do łącza DS3 podzielonego na

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

kanały, poszczególne strumienie T1 można by poddać dalszemu podziałowi przez grupowanie odpowiednich strumieni DS0. Każda grupa otrzymałaby wówczas własny numer (numer grupy kanałów). Oto przykładowa konfiguracja kontrolera:

```
controller T3 2/0
 framing m23
 clock source line
 t1 1 channel-group 1 timeslots 1-24
 t1 2 channel-group 1 timeslots 1-24
 t1 3 channel-group 1 timeslots 1-24
 t1 4 channel-group 1 timeslots 1-24
 t1 5 channel-group 1 timeslots 1-24
 t1 6 channel-group 1 timeslots 1-24
 t1 7 channel-group 1 timeslots 1-24
 t1 8 channel-group 1 timeslots 1-24
 t1 9 channel-group 1 timeslots 1-24
 t1 10 channel-group 1 timeslots 1-24
 t1 1 clock source Line
 t1 2 clock source Line
 t1 4 clock source Line
 t1 6 clock source Line
 t1 7 clock source Line
 t1 8 clock source Line
 t1 9 clock source Line
 t1 10 clock source Line
 t1 11 clock source Line
 t1 12 clock source Line
 t1 13 clock source Line
 t1 14 clock source Line
 t1 15 clock source Line
 t1 16 clock source Line
 t1 17 clock source Line
 t1 18 clock source Line
 t1 19 clock source Line
 t1 20 clock source Line
 t1 21 clock source Line
 t1 22 clock source Line
 t1 23 clock source Line
 t1 24 clock source Line
 t1 25 clock source Line
 t1 26 clock source Line
 t1 27 clock source Line
 t1 28 clock source Line
```

Powyższy listing opisuje łącze DS3 przyłączone do interfejsu 2/0. Jest to łącze dzielone na kanały, więc ustawiono ramkowanie M23. Źródło zegara wynika z samego łącza (`Line`). W przeciwieństwie do łączy T1 parametr ten nie powinien być zmieniany, ponieważ sygnał zegarowy jest zazwyczaj narzucany przez sieć SONET. Zwróć uwagę na to, że są osobne polecenia `clock` dla łącza DS3 oraz dla poszczególnych strumieni T1 przenoszonych w ramach strumienia DS3.

W łączach DS3 podzielonych na kanały urządzenia Cisco obsługują jedynie ramkowanie M23 i C-bit. Jeśli więc w module telekomunikacyjnym zostało ustawione ramkowanie M13, po stronie urządzenia Cisco należy wybrać mechanizm M23.

Po sparametryzowaniu kontrolerów można skonfigurować interfejsy szeregowe T1 tak, jakby były klasycznymi łączami T1. Nazwy interfejsów są połączeniem nazwy fizycznych interfejsów oraz numerów strumieni T1 w łączu DS3, po których następuje znak dwukropka i numer grupy kanałów zdefiniowany w konfiguracji kontrolera:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
interface Serial2/0/1:1
 description T1 nr 1
 ip address 10.220.110.1 255.255.255.252
!
interface Serial2/0/2:1
 description T1 nr 2
 ip address 10.220.120.1 255.255.255.252
!
interface Serial2/0/3:1
 description T1 nr 3
 ip address 10.220.130.1 255.255.255.252
!
interface Serial2/0/4:1
 description T1 nr 4
 ip address 10.220.140.1 255.255.255.252
```

W ramach podzielonego na kanały łącza DS3 nie można utworzyć interfejsu szeregowego o większej przepustowości niż T1. Jeśli musisz uzyskać przepustowość na poziomie kilku megabitów, rozważ utworzenie kilku strumieni T1 i połączenie ich za pomocą mechanizmu Multilink-PPP lub rozłożenie obciążenia za pomocą mechanizmu CEF bądź odpowiedniego protokołu routingu.

W routerze, z którego pochodzą listingi (7304), zainstalowano karty interfejsów bezkanałowych oraz uwzględniających podział na kanały. Wykonanie polecenia `show version` może być więc nieco zaskakujące, ponieważ informacja o liczbie interfejsów obejmuje bezkanałowe połączenia DS3 oraz wszystkie łącza T1, które zostały wykorzystane w łączach DS3 dzielonych na kanały. Omawiany router dysponuje czterema bezkanałowymi łączami DS3 oraz czterema dzielonymi na kanały. Jeśli dzielone na kanały łącze DS3 nie zostało skonfigurowane na poziomie kontrolera, nie jest uwzględniane w wyniku zastosowania instrukcji `show version`. W prezentowanym rozwiązaniu wykorzystaliśmy cztery bezkanałowe łącza DS3 (które domyślnie są interfejsami szeregowymi). Skonfigurowaliśmy również 10 łączy T1 jako interfejsy szeregowe wchodzące w skład jednego łącza DS3 dzielonego na kanały. Z tego powodu router informuje o dostępności 14 interfejsów szeregowych:

```
7304#sho ver

[fragment usunięty]

1 FastEthernet interface
2 Gigabit Ethernet interfaces
14 Serial interfaces4 Channelized T3 ports
```

Poniżej znajduje się wykaz poszczególnych interfejsów szeregowych routera. Dziesięć z nich ma charakter logiczny (s2/0/1:1 – s2/0/10:1), a cztery są interfejsami fizycznymi (s3/1/0 – s3/1/3).

```
7304#sho ip int brie
Interface              IP-Address      OK? Method  Status                Protocol
FastEthernet0          unassigned      YES NVRAM   administratively down down
GigabitEthernet0/0     10.220.11.1      YES NVRAM    up                    up
GigabitEthernet0/1     10.220.12.1      YES NVRAM    up                    up
Serial2/0/1:1          10.220.110.1     YES NVRAM    up                    up
Serial2/0/2:1          10.220.120.1     YES NVRAM    up                    up
Serial2/0/3:1          10.220.130.1     YES NVRAM    up                    up
Serial2/0/4:1          10.220.140.1     YES NVRAM    up                    up
Serial2/0/5:1          unassigned      YES manual  administratively down down
Serial2/0/6:1          unassigned      YES manual  administratively down down
Serial2/0/7:1          unassigned      YES manual  administratively down down
Serial2/0/8:1          unassigned      YES manual  administratively down down
Serial2/0/9:1          unassigned      YES manual  administratively down down
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Serial2/0/10:1         unassigned      YES manual  administratively down down
Serial3/1/0            10.100.100.100  YES manual  up                     up
Serial3/1/1            unassigned      YES NVRAM   down                   down
Serial3/1/2            unassigned      YES NVRAM   down                   down
Serial3/1/3            unassigned      YES NVRAM   down                   down
```

Wynik wykonania polecenia `show controllers` w odniesieniu do łącza DS3 podzielonego na kanały odbiega istotnie od zestawienia dotyczącego bezkanałowego łącza DS3. Zawiera bowiem informacje o stanie linii z okresu ostatnich 24 godzin przedstawiane z uwzględnieniem 15-minutowych interwałów. Bieżące alarmy oraz dane na temat kodowania liniowego zostały wyróżnione.

```
7304#sho controllers t3 2/0
T3 2/0 is up. Hardware is 2CT3+ single wide port adapter
  CT3 H/W Version: 0.2.2, CT3 ROM Version: 1.0, CT3 F/W Version: 2.5.1
  FREEDM version: 1, reset 0 resurrect 0
  Applique type is Channelized T3
  Description:
  No alarms detected.
  Framing is M23, Line Code is B3ZS, Clock Source is Line
  Rx-error throttling on T1's ENABLED
  Rx throttle total 99, equipment customer loopback
  Data in current interval (29 seconds elapsed):
     0 Line Code Violations, 0 P-bit Coding Violation
     0 C-bit Coding Violation, 0 P-bit Err Secs
     0 P-bit Severely Err Secs, 0 Severely Err Framing Secs
     0 Unavailable Secs, 0 Line Errored Secs
     0 C-bit Errored Secs, 0 C-bit Severely Errored Secs
  Data in Interval 1:
     0 Line Code Violations, 0 P-bit Coding Violation
     0 C-bit Coding Violation, 0 P-bit Err Secs
     0 P-bit Severely Err Secs, 0 Severely Err Framing Secs
     0 Unavailable Secs, 0 Line Errored Secs
     0 C-bit Errored Secs, 0 C-bit Severely Errored Secs
  Data in Interval 2:
     0 Line Code Violations, 0 P-bit Coding Violation
     0 C-bit Coding Violation, 0 P-bit Err Secs
     0 P-bit Severely Err Secs, 0 Severely Err Framing Secs
     0 Unavailable Secs, 0 Line Errored Secs
     0 C-bit Errored Secs, 0 C-bit Severely Errored Secs
[fragment usunięty]
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 23.

Frame Relay

Protokół **Frame Relay** jest jednym z mechanizmów transportu informacji cyfrowych w sieci. Dane są dzielone na ramki, które następnie przesyła się między urządzeniami sieciowymi zarządzanymi najczęściej przez operatorów telekomunikacyjnych. Sama sieć Frame Relay jest często przedstawiana na schematach jako chmura, ponieważ użytkownik usługi zazwyczaj nie wie (i nie jest zainteresowany tym), w jaki sposób realizowane są połączenia między jej węzłami. Klient końcowy musi jedynie znać **obwód wirtualny** (VC — ang. *Virtual Circuit*), który zapewnia dostarczenie ramek do drugiego punktu końcowego chmury.

Wszystkie dane wprowadzone po jednej stronie obwodu VC powinny zostać wyemitowane po drugiej stronie. Punkt zdalny można więc postrzegać jako drugi koniec kabla, mimo że w rzeczywistości dostarczenie do niego informacji wymaga przekazania ich przez wiele węzłów w sieci.

Obwody VC mogą być **przełączane** lub **stałe**. Stałe obwody wirtualne (PVC — ang. *Permanent Virtual Circuit*) działają zawsze, nawet jeśli nie są wykorzystywane. Przełączane obwody wirtualne (SVC — ang. *Switched Virtual Circuit*) są ustanawiane jedynie wówczas, gdy są potrzebne. Większość sieci transmisji danych bazuje na rozwiązaniach PVC.

Na rysunku 23.1 przedstawiono typową (choć niezbyt skomplikowaną) sieć Frame Relay wykorzystującą obwody PVC. Obwody te zapewniają połączenie routera A z routerem B oraz routera A z routerem C. Połączenie takie nie występuje natomiast między routerami B i C. Obydwa obwody kończą się we wspólnym interfejsie routera A.

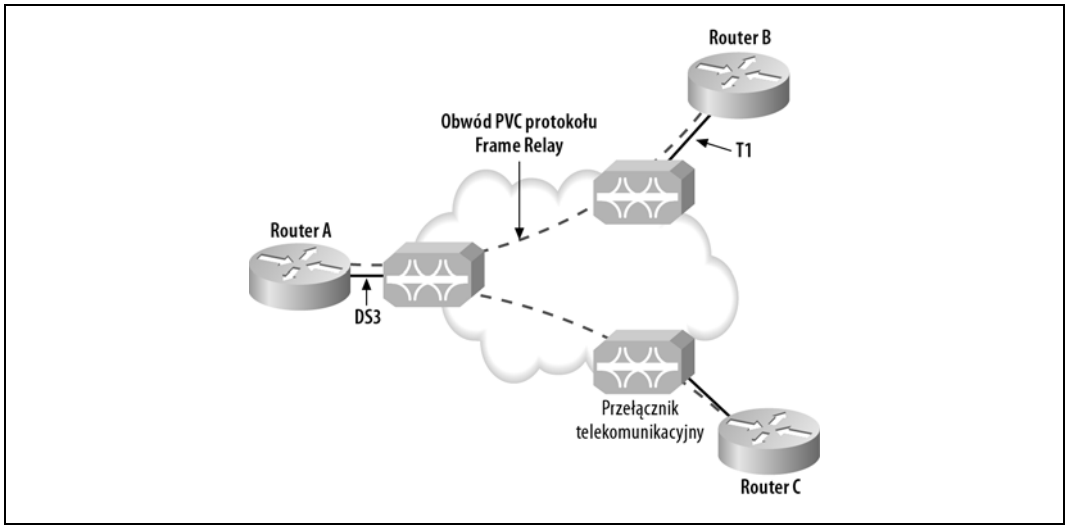
W warstwie fizycznej każdy router jest przyłączony do jednego przełącznika zarządzanego przez operatora telekomunikacyjnego. Wspomniane przełączniki stanowią punkty wejścia danych do chmury oraz ich wyjścia z chmury.

Router A jest przyłączony za pomocą łącza DS3, natomiast routery B i C pracują na końcu łączy T1. Jednak w warstwie logicznej mechanizm Frame Relay symuluje bezpośrednie połączenie dwóch odległych urządzeń za pomocą obwodu PVC.

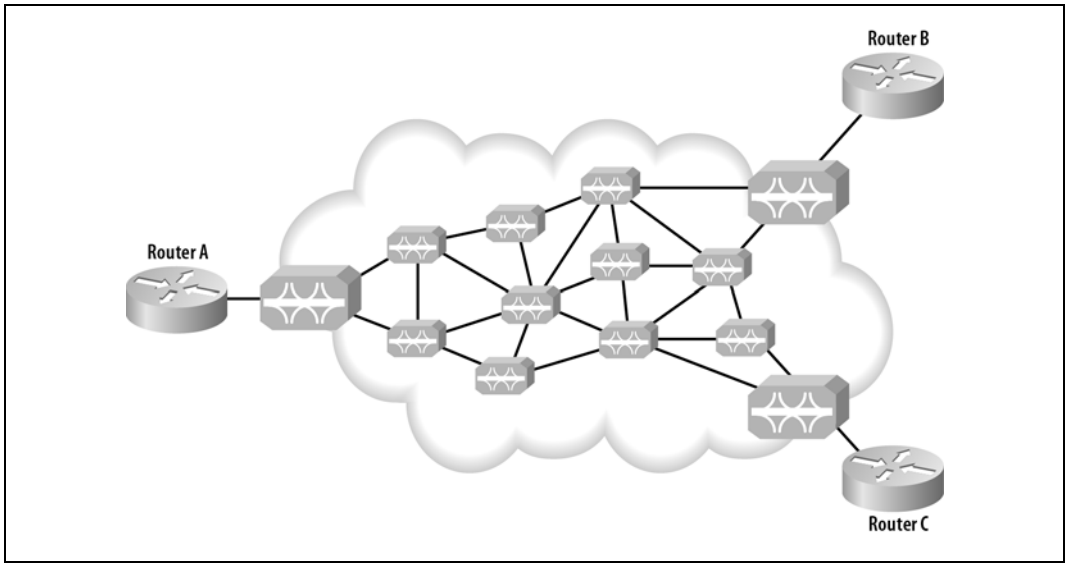
W rzeczywistości w chmurze może się znajdować wiele urządzeń, które pośredniczą w przekazywaniu ramek między punktami końcowymi. Przykład takiej konfiguracji został przedstawiony na rysunku 23.2.

Rozwiązania Frame Relay funkcjonują w sposób zbliżony do internetu. Gdy użytkownik wysyła żądanie do zdalnej witryny, nie zastanawia się, w jaki sposób pakiet ten zostanie dostarczony.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 23.1. Prosta sieć WAN w technologii Frame Relay



Rysunek 23.2. Budowa fizyczna sieci Frame Relay

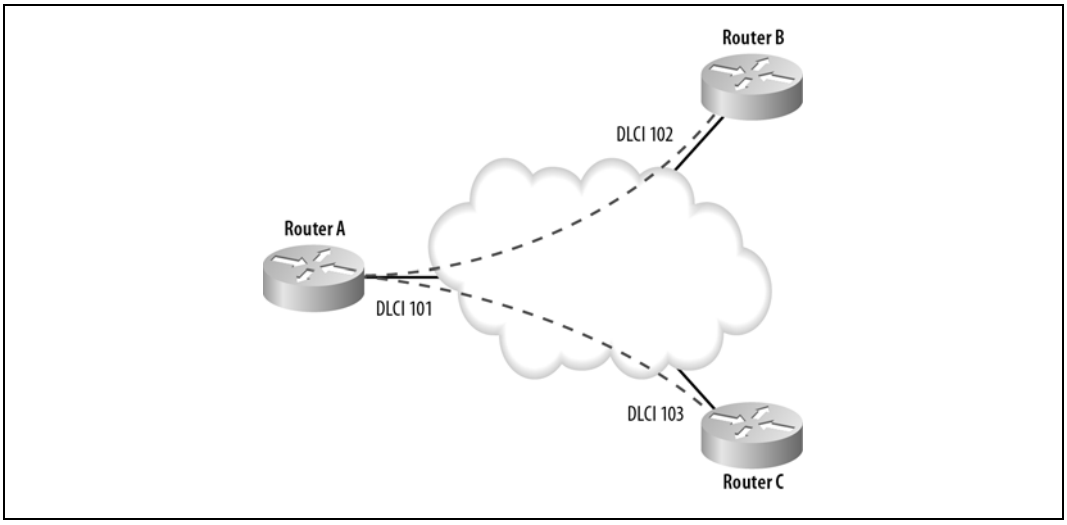
Nie interesuje go to, przez ile urządzeń pakiety są przekazywane. Musi jedynie wiedzieć, pod jakim adresem pracuje brama domyślna. O resztę martwi się dostawca usług internetowych. Mechanizm Frame Relay działa podobnie. Użytkownik sieci nie wie, przez jakie urządzenia pośredniczące dane są przesyłane lub w jaki sposób wyznaczany jest routing pakietów. Wiadomo jedynie, że liczba punktów końcowych jest skończona. Wszystkie są bowiem zdefiniowane przez abonenta usługi i udostępnione przez operatora telekomunikacyjnego.

Aby zapewnić iluzję bezpośredniego połączenia między urządzeniami, każdy punkt końcowy obwodu VC otrzymuje adres warstwy drugiej, nazywany **identyfikatorem połączenia w warstwie łącza danych** (DLCI — ang. *Data Link Connection Identifier*). Identyfikatory DLCI (i powią-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

zane z nimi dane) są przetwarzane jedynie w przełącznikach telekomunikacyjnych oraz routerach brzegowych. Inni użytkownicy przyłączeni do sieci Frame Relay nie mogą posługiwać się Twoimi identyfikatorami DLCI lub przetwarzać przesyłanych w obwodzie danych. Wartości DLCI określa operator telekomunikacyjny.

Obwody wirtualne można definiować w taki sposób, aby jeden identyfikator DLCI opisywał zakończenia kilku połączeń. Przykład takiego rozwiązania został pokazany na rysunku 23.3. Nic również nie stoi na przeszkodzie, aby każdy obwód VC dysponował niezależnymi wartościami DLCI na obydwu końcach. Sposób wykorzystania obwodów PVC zależy od konkretnych potrzeb administratora sieci.



Rysunek 23.3. Identyfikatory DLCI w sieci Frame Relay

W sieciach Frame Relay wartości DLCI są odwzorowywane na adresy IP w taki sam sposób jak adresy MAC w sieciach Ethernet. Jednak w przeciwieństwie do sieci Ethernet (w których pozyskiwanie adresów MAC ma charakter dynamiczny) obowiązek powiązania identyfikatora DLCI z odpowiednim adresem IP spoczywa na administratorze routera.

W sieciach Ethernet do ustalenia adresu MAC na podstawie adresu IP wykorzystuje się protokół ARP. W rozwiązaniach Frame Relay podobne zadanie jest realizowane przez odwrotny protokół ARP — **Inverse ARP** — który służy do odwzorowywania znanych wartości DLCI na odpowiadające im adresy IP. Przełączniki Frame Relay okresowo dostarczają bowiem informacji na temat statusu wszystkich skonfigurowanych identyfikatorów DLCI.

Konfigurując protokół Frame Relay, zachowaj szczególną ostrożność. W konfiguracji routera mogą być zapisane obwody PVC, które nie powinny zostać udostępnione. Mechanizm Inverse ARP może je jednak udostępnić bez Twojej wiedzy.

Największą zaletą połączeń Frame Relay jest ich niski koszt i duża elastyczność konfiguracji. Łącze punkt-punkt zrealizowane w technologii T1 jest znacznie droższe niż analogiczny obwód Frame Relay ustanowiony między dwoma ośrodkami, szczególnie jeśli nie są one umiejscowione w tym samym obszarze geograficznym. Ponadto właściwa połączeniu T1 szybkość transmisji danych 1,5 Mb/s musi zostać przydzielona każdemu punktowi końcowemu niezależnie od

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

tego, jakie będzie faktyczne obciążenie łącza. Rozwiązania Frame Relay pozwalają na współdzielenie zasobów — jeśli pewna część pasma nie jest wykorzystywana, inni użytkownicy sieci mogą z niej skorzystać.

Współdzielenie pasma zawsze rodzi pewne wątpliwości. Co się stanie, jeśli będzie Ci potrzebne pasmo, z którego w danej chwili korzysta ktoś inny? Jakie będzie działanie sieci, gdy ktoś zażąda pasma, z którego korzystasz? W protokole Frame Relay uwzględniono funkcję **gwarantowania szybkości transmisji** (CIR — ang. *Commited Information Rate*), która reguluje te kwestie.

Zamawianie usługi Frame Relay

Gdy zamawiasz łącze Frame Relay, musisz przekazać operatorowi telekomunikacyjnemu cztery informacje. Jedną z nich jest właśnie wartość CIR. Oto wszystkie potrzebne dane:

Adresy i numery telefoniczne punktów końcowych

Adresy pocztowe, a przede wszystkim numery telefonów, są najważniejszymi informacjami potrzebnymi do zamówienia łącza. Jeśli w danej lokalizacji nie są realizowane usługi telefoniczne, taki obszar nie występuje na mapie operatora telekomunikacyjnego.

Przepustowość portu

Przepustowość portu określa szybkość transmisji danych w łączu wykorzystywanym w każdym punkcie końcowym. Przepustowości poszczególnych zakończeń nie muszą być jednakowe. Nie mogą być natomiast mniejsze od wartości CIR zdefiniowanej w odniesieniu do danego połączenia Frame Relay. Dolną granicę wartości wyznacza obwód 56 kb/s. Maksymalna przepustowość przekracza wartości charakterystyczne dla połączeń DS3. Zazwyczaj łącze Frame Relay umożliwia przekazywanie danych z szybkością T1 (E1), choć stosuje się również przepustowości odpowiadające ułamkowym częściom szybkości T1 (zależnie od ceny łącza).

CIR

Parametr CIR wyznacza szybkość transmisji danych, którą operator powinien zapewnić. Zamawiając usługę Frame Relay, musisz określić szerokość stale dostępnego pasma. Wartość ta zobowiązuje operatora telekomunikacyjnego do przekazywania w obwodzie wirtualnym wszystkich ramek, które nie wyczerpują zamówionej przepustowości (nadmiarowe ramki mogą być jednak usuwane). Zdefiniowana przepływność odnosi się do całego połączenia między danym ośrodkiem i wskazanymi ośrodkami docelowymi. Zwiększanie wartości CIR podnosi koszt łącza. Parametr CIR jest definiowany w odniesieniu do każdego obwodu VC.

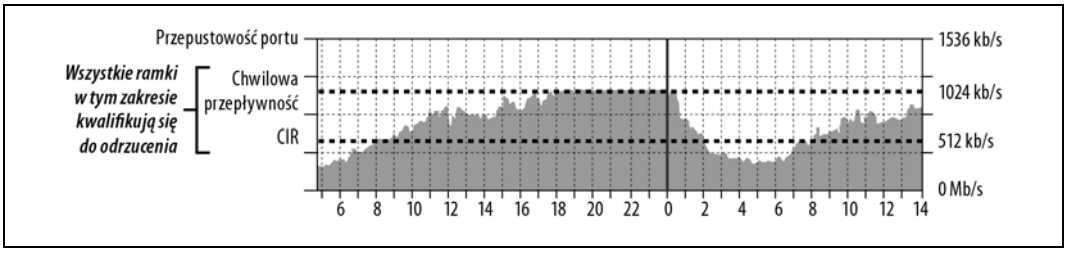
Chwilowa przepływność

Chwilowa przepływność (ang. *burst rate*) opisuje maksymalną szybkość transmisji w obwodzie wirtualnym sieci Frame Relay. Ramki, które przekraczają wartość CIR (ale nie powodują przekroczenia chwilowej przepływności), są oznaczane jako **kwalifikujące się do usunięcia** (DE — ang. *Discard Eligible*). Są one przekazywane między przełącznikami Frame Relay tylko w przypadku, gdy sieć dysponuje dostatecznie dużą przepustowością. Operator telekomunikacyjny może je jednak usunąć bez jakichkolwiek konsekwencji. Wyznaczenie dużej wartości chwilowej przepływności pozwala na uzyskanie szerokiego pasma transmisyjnego bez podnoszenia kosztów połączenia. Zazwyczaj wartość parametru jest wielokrotnością wartości CIR. Często jest ona dwukrotnie większa od gwaran-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

towanej przepustowości (2 × CIR). Oczywiście można również zamówić obwód wirtualny o zerowej wartości chwilowej przepływności. Parametr ten należy jednak zdefiniować niezależnie dla każdego obwodu VC.

Na rysunku 23.4 przedstawiono wykres wykorzystania pasma w łączu Frame Relay. Odnosi się on do łącza T1 pracującego z przepustowością portu 1,5 Mb/s. Parametr CIR ma wartość 512 kb/s, a chwilowa przepływność jest dwukrotnością CIR. Jak nietrudno zauważyć, przekroczenie dwukrotnej przepustowości CIR powoduje usunięcie ramek z sieci. Z wykresu wynika również, że łącze jest wysycone. Jego przepustowość powinna zostać zwiększona. Warto stosować zasadę, zgodnie z którą należy zwiększać przepustowość połączenia po przekroczeniu 70 procent nominalnej przepustowości. Oznacza to, że parametry opisywanego obwodu powinny zostać zmienione już znacznie wcześniej.



Rysunek 23.4. Parametry CIR i DE w sieci Frame Relay

W wielu rozwiązaniach korporacyjnych standard Frame Relay stracił wiodącą rolę. Agresywna polityka cenowa dostawców usług internetowych oraz doprowadzanie łączy w standardzie Ethernet sprawiły, że większość firm woli skorzystać z technologii MPLS, której wybór jest uzasadniony ekonomicznie. Niemniej rozwiązania Frame Relay nadal są stosowane. Trzeba je więc znać choćby po to, by zastąpić je połączeniami MPLS.

Projektowanie sieci Frame Relay

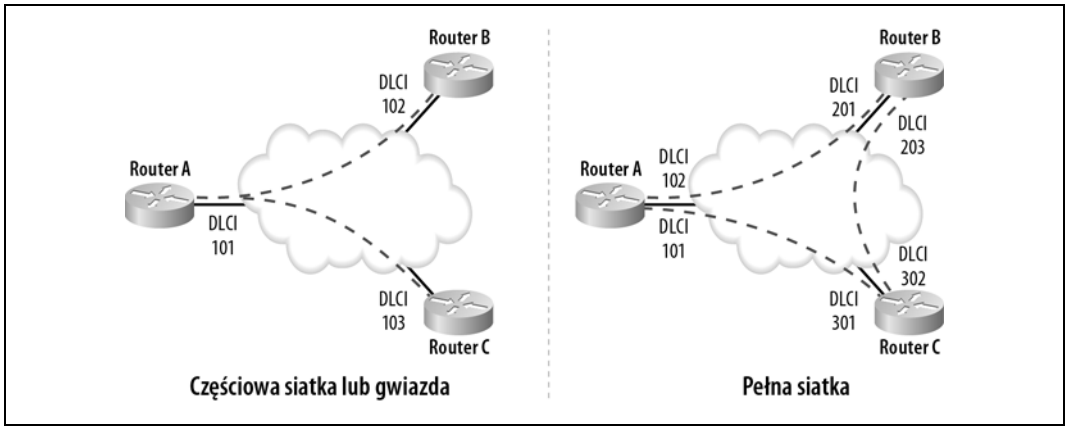
Połączenia Frame Relay cechują się większą elastycznością niż łącza punkt-punkt, ponieważ umożliwiają ustanawianie wielu obwodów za pomocą pojedynczego interfejsu routera. Dzięki temu można zrealizować połączenia z wieloma ośrodkami za cenę niższą niż w przypadku połączeń punkt-punkt.

Sposób połączenia trzech ośrodków za pomocą sieci Frame Relay został pokazany na rysunku 23.5. Widoczne po lewej stronie routery B i C są połączone z routerem A, ale nie są połączone ze sobą nawzajem. Taka konfiguracja jest często nazywana **częściową siatką** (ang. *partial mesh*) lub **siecią gwieździstą** (ang. *hub and spoke*). Routery B i C mogą się ze sobą komunikować jedynie za pośrednictwem routera A.

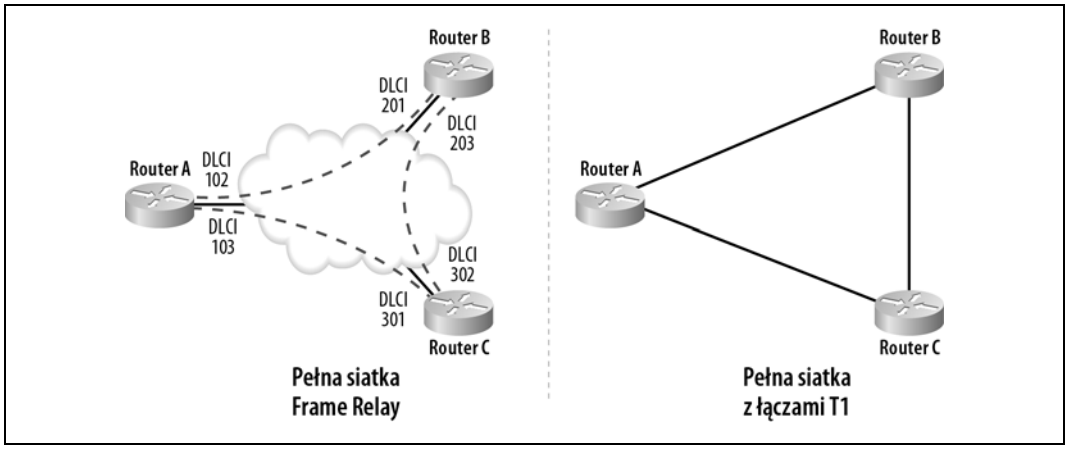
W prawej części rysunku 23.5 przedstawiono przykład **pełnej siatki** (ang. *full mesh*). W tej topologii każdy z ośrodków jest połączony ze wszystkimi pozostałymi ośrodkami. Router B może się więc komunikować bezpośrednio z routerem C.

Sieci siatkowe nie są rozwiązaniami charakterystycznymi jedynie dla standardu Frame Relay. Jak widać na rysunku 23.6, taki sam schemat połączeń można łatwo zrealizować za pomocą łączy punkt-punkt T1.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 23.5. Siatkowe sieci Frame Relay



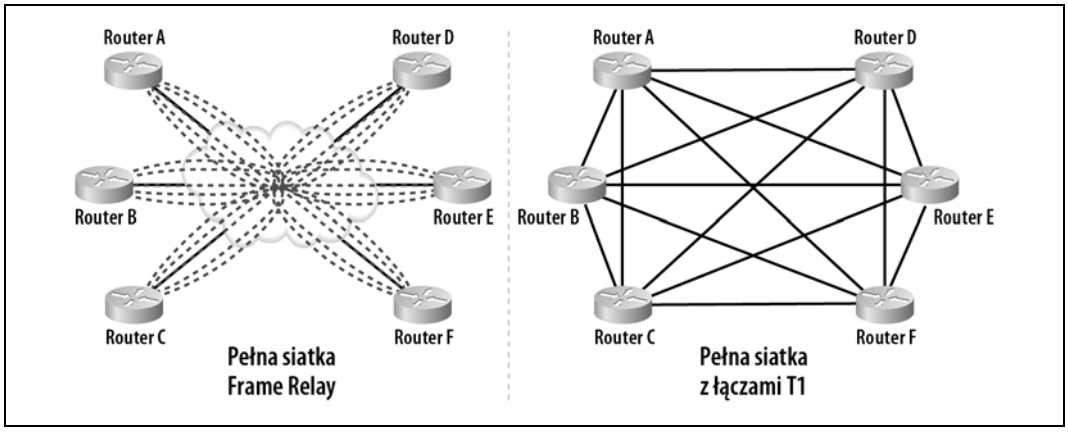
Rysunek 23.6. Sieci siatkowe Frame Relay i punkt-punkt

W sieci Frame Relay (przedstawionej w lewej części rysunku) każdy z routerów musi być wyposażony w pojedynczy interfejs T1. Poszczególne obwody PVC są bowiem konfigurowane jako niezależne interfejsy wirtualne nazywane **podinterfejsami** (ang. *subinterface*). Zastosowanie podinterfejsów umożliwia wydzielenie obwodów wirtualnych odpowiadających poszczególnym interfejsom logicznym jednego interfejsu fizycznego.

W przypadku połączeń punkt-punkt każdy router musiałby być wyposażony w dwa interfejsy T1. Routery obsługujące dwa połączenia T1 są droższe niż urządzenia przeznaczone do przesyłania danych w jednym łączu T1. Ponadto usługi transmisji danych w łączach T1 (szczególnie na dużych odległościach) są znacznie droższe niż dzierżawienie obwodów Frame Relay.

Przykład przedstawiony na rysunku 23.6 nie jest szczególnie skomplikowany, ale oddaje ideę obowiązującą również w większych sieciach. Na rysunku 23.7 pokazano dwie sieci o sześciu węzłach. W lewej części widać konfigurację siatkową zrealizowaną za pomocą technologii Frame Relay. Natomiast po prawej stronie przedstawiono tę samą topologię wykonaną na bazie połączeń punkt-punkt.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 23.7. Sieć siatkowa składająca się z sześciu węzłów

W przypadku sześciu węzłów sieci z każdego routera wychodzi pięć połączeń. W technologii Frame Relay do zbudowania takiej sieci wystarczy jeden interfejs T1 w każdym routerze (przy założeniu, że sumaryczna przepustowość łączy nie przekracza przepustowości T1). W przypadku połączeń punkt-punkt konieczne jest jednak doprowadzenie pięciu łączy T1 do każdego z routerów sieci. Oprócz wyższej ceny urządzeń trzeba więc uwzględnić również wyższy koszt usługi telekomunikacyjnej, szczególnie gdy wymiana danych jest realizowana na dużych odległościach.

Aby wyznaczyć liczbę łączy potrzebnych do zbudowania sieci siatkowej, możesz się posłużyć wzorem N(N–1)/2, w którym N odpowiada liczbie węzłów w sieci. W opisywanym przykładzie jest sześć węzłów, więc potrzeba 6(6–1)/2 = 15 łączy.

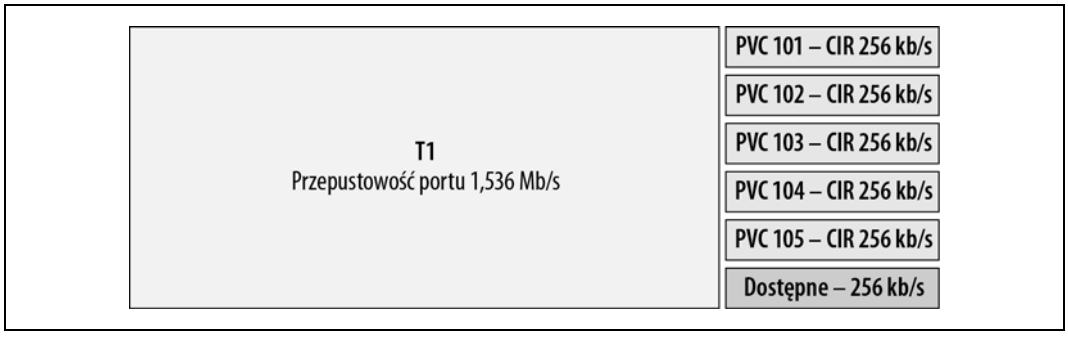
Nadsubskrypcja

Projektując sieci Frame Relay, trzeba zadbać o to, by sumaryczna wartość szerokości pasma wynikająca z parametrów CIR nie przekraczała przepustowości portu. Parametr CIR definiuje gwarantowaną szerokość pasma, a z technicznego punktu widzenia nie można zagwarantować szerokości pasma wykraczającej poza możliwości portu. Jeśli sumaryczna przepustowość obwodów wirtualnych danego łącza przekracza szybkość transmisji portu, mamy do czynienia z **nadsubskrypcją** (ang. *oversubscription*). Nie można na przykład utrzymywać czterech obwodów PVC o parametrze CIR równym 512 kb/s w łączu T1, gdyż suma wartości CIR wynosi 2048 kb/s. Niektórzy operatorzy telekomunikacyjni dopuszczają takie sytuacje, a inni nie.

Chwilowa przepływność nie jest uwzględniana przy ustalaniu, czy występuje nadsubskrypcja.

Zawsze należy starannie planować konfigurację sieci i pamiętać, że suma wartości CIR obwodów PVC nie może przekraczać przepustowości łącza fizycznego. Osobiście korzystam w tym celu z arkusza kalkulacyjnego, ale każda forma dokumentacji spełnia te same zadania. Często najefektywniejsze okazują się proste schematy graficzne, jak ten przedstawiony na rysunku 23.8.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 23.8. Wykorzystanie łącza T1 do ustanowienia obwodów PVC w sieci Frame Relay

Nie ma technicznych ograniczeń, które uniemożliwiałyby nadsubskrypcję. W czasie boomu internetowego niewielkie firmy ISP często debatowały na temat etycznych stron nadsubskrypcji, która była bardzo często stosowana w łączach Frame Relay w celu obniżenia kosztów i zwiększenia przychodów.

Rozwiązanie to nigdy nie jest korzystne dla odbiorców usług. Zawsze bowiem natężenie ruchu wzrasta do poziomu, powyżej którego pakiety są usuwane.

Uważaj na dostawców usług telekomunikacyjnych, którzy oferują łącza Frame Relay o zerowej wartości CIR. Takie postępowanie pozwala obniżyć koszt usługi, ale jednocześnie wiąże się z koniecznością oznaczenia wszystkich przekazywanych ramek jako zakwalifikowanych do odrzucenia. Zamawiając łącze, precyzyjnie określ jego parametry i zawsze sprawdzaj, co kupujesz.

Lokalny interfejs zarządzania

W roku 1990 firmy Cisco Systems, StrataCom, Northern Telecom oraz Digital Equipment Corporation opracowały zestaw rozszerzeń do protokołu Frame Relay, który nazwano **lokalnym interfejsem zarządzania** (LMI — ang. *Local Management Interface*). Mechanizm LMI zapewnia komunikację między urządzeniem terminala danych (czyli komponentem DTE; na przykład routerem) a urządzeniem transmisji danych (czyli komponentem DCE; na przykład przełącznikiem telekomunikacyjnym). Najbardziej użyteczną cechą interfejsów LMI jest to, że zapewniają przekazywanie danych statusowych o obwodach wirtualnych. Dzięki nim router uzyskuje informację, czy dany obwód PVC jest dostępny.

Komunikaty LMI są przesyłane w ramach wstępnie zdefiniowanych obwodów PVC. Aby określić rodzaj wykorzystywanych rozszerzeń LMI oraz odpowiadających im obwodów PVC, wystarczy wykonać instrukcję `show interfaces`:

```
Router-A#sho int s0/0 | inc LMI
  LMI enq sent 85, LMI stat recvd 86, LMI upd recvd 0, DTE LMI up
  LMI enq recvd 0, LMI stat sent 0, LMI upd sent 0
  LMI DLCI 1023 LMI type is CISCO frame relay DTE
```

Routery Cisco umożliwiają wybranie jednego z trzech rodzajów interfejsów LMI: `cisco`, `ansi` oraz `q933a` (aneks A). Urządzenie DCE (przełącznik telekomunikacyjny) zazwyczaj automa-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

tycznie wykrywa rodzaj stosowanych komunikatów LMI. Domyślne ustawienie w routerach Cisco to oczywiście `cisco`. Zmiana rodzaju LMI jest możliwa dzięki poleceniu `frame-relay lmi-type` wykonywanemu w ramach konfiguracji interfejsu sieciowego:

```
Router-A(config-if)#frame-relay lmi-type ?
  cisco
  ansi
  q933a
```

Unikanie przeciążenia w połączeniach Frame Relay

Standard Frame Relay definiuje mechanizmy unikania przeciążenia. Jego działanie bazuje na dwóch bitach nagłówka ramki, które pełnią funkcje **powiadamiania o przeciążeniu w kierunku transmisji** (FECN — ang. *Forward-Excplicit Congestion Notification*) oraz **wstecznego powiadamiania o przeciążeniu** (BECN — ang. *Backward-Explicit Congestion Notfication*). Obydwa znaczniki służą do informowania urządzeń DTE (routerów) o zaobserwowanym przeciążeniu sieci.

Bit FECN jest ustawiany przez moduł DCE (przełącznik telekomunikacyjny), gdy zarejestruje on przeciążenie sieci. Odbiorcze urządzenie DTE (router) może wykorzystać tę informację do uruchomienia mechanizmu sterowania przepływem danych (o ile jest odpowiednio skonfigurowane). Elementy chmury Frame Relay nie wdrażają żadnej polityki sterowania przepływem. Jest to zadaniem urządzeń końcowych.

Przełączniki Frame Relay ustawiają bit BECN wówczas, gdy przekazywane w przeciwnym kierunku ramki mają ustawiony bit FECN. Dzięki temu nadawcze urządzenie DTE jest informowane o przeciążeniu wynikającym z generowanych przez nie ramek.

Przykład występowania przeciążenia w sieci Frame Relay został przedstawiony na rysunku 23.9. Między routerami A i B jest ustanowiony obwód PVC, który przebiega przez górne przełączniki. Jeden z węzłów chmury wykrywa przeciążenie w transmisji kierowanej do routera B. Oznacza więc pakiety przekazywane do routera B za pomocą bitu FECN, a pakiety transmitowane w przeciwnym kierunku (do routera A) za pomocą bitu BECN.

Konfiguracja protokołu Frame Relay

Znając zasady działania protokołu Frame Relay, nie powinieneś mieć kłopotu z samą konfiguracją. Co prawda są w niej pewne nowe elementy, takie jak definiowanie podinterfejsów, ale nie są one szczególnie kłopotliwe (poza tym zostały szczegółowo opisane w dalszej części podrozdziału).

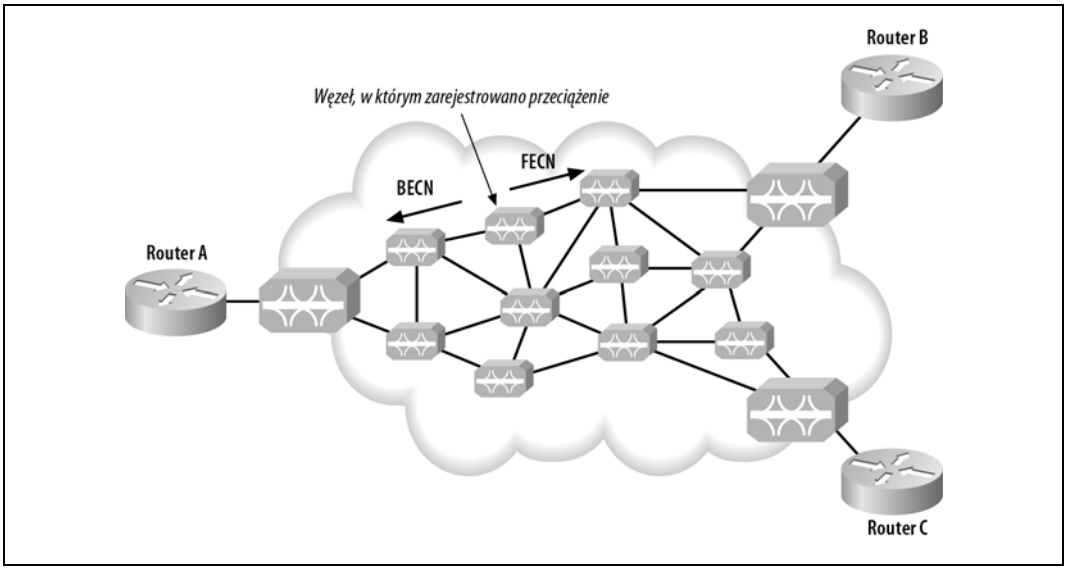
Podstawowa konfiguracja Frame Relay dla dwóch węzłów

Na rysunku 23.10 przedstawiono prostą sieć Frame Relay złożoną z dwóch węzłów. Router A jest połączony z routerem B za pomocą łącza T1. Przepustowość portu wynosi 1,536 Mb/s, wartość CIR to 512, a przepływność chwilowa jest równa dwukrotności CIR (1024 kb/s).

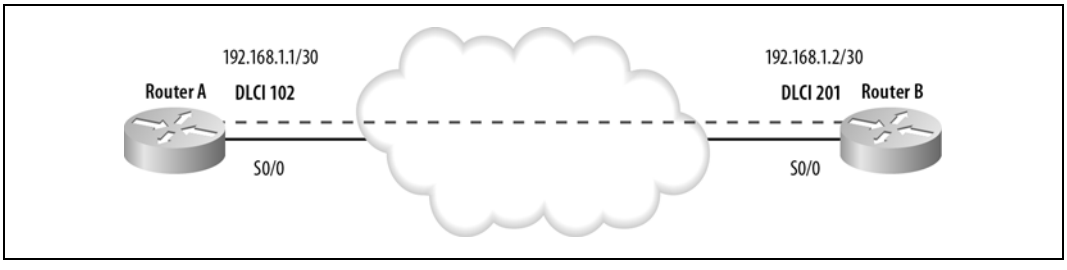
Pierwsza czynność polega na zdefiniowaniu enkapsulacji Frame Relay. Urządzenia Cisco umożliwiają wykorzystanie dwóch rodzajów enkapsulacji: `cisco` i `ietf`. Domyślnym ustawieniem jest `cisco`. Zostaje ono uwzględnione po wykonaniu polecenia `encapsulation frame-relay`:

```
interface Serial0/0
encapsulation frame-relay
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 23.9. Przykład użycia bitów FECN i BECN



Rysunek 23.10. Sieć Frame Relay złożona z dwóch węzłów

Aby ustawić enkapsulację `ietf`, należy wprowadzić instrukcję `encapsulation frame-relay ietf`. Ten sposób formowania ramek jest wykorzystywany zazwyczaj jedynie w połączeniach z urządzeniami innych dostawców niż Cisco.

Po skonfigurowaniu enkaspulacji i włączeniu interfejsu powinieneś zobaczyć komunikat statusowy interfejsu LMI. Wszelkie zdefiniowane obwody PVC są wyświetlane po wpisaniu polecenia `show frame-relay PVC`:

```
Router-A#sho frame pvc

PVC Statistics for interface Serial0/0 (Frame Relay DTE)

            Active   Inactive    Deleted     Static
  Local        0          0          0          0
  Switched     0          0          0          0
  Unused       0          1          0          0
DLCI = 102, DLCI USAGE = UNUSED, PVC STATUS = INACTIVE, INTERFACE = Serial0/0
  input pkts 0           output pkts 0        in bytes 0
  out bytes 0            dropped pkts 0       in pkts dropped 0
  out pkts dropped 0     out bytes dropped 0
  in FECN pkts 0         in BECN pkts 0       out FECN pkts 0
  out BECN pkts 0        in DE pkts 0         out DE pkts 0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  out bcast pkts 0       out bcast bytes 0
  switched pkts 0
  Detailed packet drop counters:
  no out intf 0          out intf down 0      no out PVC 0
  in PVC down 0          out PVC down 0       pkt too big 0
  shaping Q full 0       pkt above DE 0       policing drop 0
  pvc create time 00:19:12, last time pvc status changed 00:19:12
```

Wymieniane komunikaty dostarczają informacji związanych z lokalnym identyfikatorem DLCI, a nie z identyfikatorem zdalnym. Interfejs LMI przekazuje dane istotne dla routera lokalnego, a nie właściwe dla routera odległego. Wykonanie tego samego polecenia w routerze B spowoduje zwrócenie informacji na temat połączenia z perspektywy routera B:

```
Router-B#sho frame pvc

PVC Statistics for interface Serial0/0 (Frame Relay DTE)
            Active   Inactive    Deleted     Static
  Local        0          0          0          0
  Switched     0          0          0          0
  Unused       0          1          0          0
DLCI = 201, DLCI USAGE = UNUSED, PVC STATUS = ACTIVE, INTERFACE = Serial0/0
  input pkts 0           output pkts 0        in bytes 0
  out bytes 0            dropped pkts 0       in pkts dropped 0
  out pkts dropped 0     out bytes dropped 0
  in FECN pkts 0         in BECN pkts 0       out FECN pkts 0
  out BECN pkts 0        in DE pkts 0         out DE pkts 0
  out bcast pkts 0       out bcast bytes 0
  switched pkts 0
  Detailed packet drop counters:
  no out intf 0          out intf down 0      no out PVC 0
  in PVC down 0          out PVC down 0       pkt too big 0
  shaping Q full 0       pkt above DE 0       policing drop 0
  pvc create time 00:19:08, last time pvc status changed 00:19:08
```

Gdy interfejs LMI jest aktywny, możesz przypisać interfejsowi adres IP (w taki sam sposób jak w przypadku interfejsów innych typów). Oto konfiguracja adresu IP po stronie routera A:

```
Router-A(config-if)#int s0/0
Router-A(config-if)#ip address 192.168.1.1 255.255.255.252
```

Ustawienie adresu IP w routerze B:

```
Router-B(config-if)#int s0/0
Router-B(config-if)#ip address 192.168.1.2 255.255.255.252
```

Po tej operacji można wykonać polecenie `ping` w odniesieniu do urządzenia pracującego po drugiej stronie łącza:

```
Router-A#ping 192.168.1.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.1.2, timeout is 2 seconds:
.!!!!
```

Polecenie `ping` działa poprawnie, ponieważ router wykrywa połączenie punkt-punkt na podstawie maski podsieci.

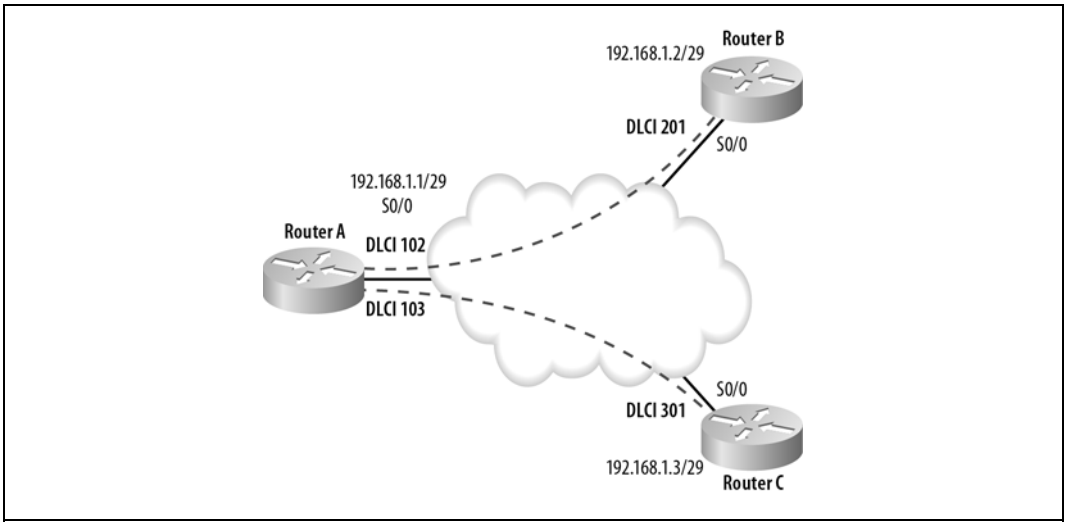
Aby sprawdzić status obwodu PVC i sposób odwzorowania adresów IP, wystarczy wykonać instrukcję `show frame-relay map`:

```
Router-A#sho frame map
Serial0/0 (up): ip 192.168.1.2 dlci 102(0×66,0×1860), dynamic,
              broadcast,, status defined, active
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Podstawowa konfiguracja Frame Relay dla większej liczby węzłów

Na rysunku 23.11 przedstawiono nieco bardziej skomplikowaną sieć Frame Relay, w której router A dysponuje obwodami PVC prowadzącymi do routerów B i C. Z kolei pomiędzy routerami B i C nie ma bezpośredniego połączenia.



Rysunek 23.11. Sieć Frame Relay złożona z trzech węzłów

Zadanie konfiguracji sieci rozpoczyna się, podobnie jak w poprzednim przykładzie, od zdefiniowania enkapsulacji Frame Relay. Należy je wykonać w jednakowy sposób we wszystkich routerach:

```
interface Serial0/0
encapsulation frame-relay
```

Również konfiguracja adresów IP nie odbiega szczególnie od wcześniejszego przypadku. Jedyna różnica wynika z wartości maski podsieci. Oto polecenia, które trzeba wprowadzić w poszczególnych routerach:

Router A:

```
Router-A(config)#int s0/0
Router-A(config-if)#ip address 192.168.1.1 255.255.255.248
```

Router B:

```
Router-B(config)#int s0/0
Router-B(config-if)#ip address 192.168.1.2 255.255.255.248
```

Router C:

```
Router-C(config)#int s0/0
Router-C(config-if)#ip address 192.168.1.3 255.255.255.248
```

Wykonanie powyższych instrukcji skutkuje uruchomieniem sieci Frame Relay, której działanie można przetestować za pomocą następujących poleceń:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Router-A#ping 192.168.1.2

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.1.2, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 56/57/60 ms

Router-A#ping 192.168.1.3

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.1.3, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 56/57/60 ms
```

Gdy sieć składała się jedynie z dwóch węzłów, routery mogły wyznaczyć adresy IP jednostki zdalnej na podstawie maski podsieci. W tym przypadku nie jest to możliwe, ponieważ konfiguracja obejmuje więcej niż dwa węzły.

Jeśli maska podsieci jest różna od `255.255.255.252`, routery wykorzystują protokół Inverse ARP do ustalenia adresu IP powiązanego z danym identyfikatorem DLCI.

Stosowanie protokołu Inverse ARP w rozbudowanych sieciach może się okazać problematyczne, ponieważ powoduje ono ujawnienie wszystkich obwodów PVC i odpowiadających im adresów (o ile zostały zdefiniowane). To z kolei może doprowadzić do udostępnienia niepotrzebnych lub ukrytych połączeń. Aby wyłączyć mechanizm Inverse ARP, należy użyć polecenia `no frame-relay inverse-arp`.

Lepszym sposobem odwzorowywania identyfikatorów DLCI na adresy IP jest własnoręczne zdefiniowanie odpowiednich powiązań. Służy do tego polecenie `frame-relay map`:

```
Router-A(config-if)#frame-relay map ip 192.168.1.2 102 broadcast
Router-A(config-if)#frame-relay map ip 192.168.1.3 103 broadcast
```

Odwzorowanie realizowane za pomocą protokołu Inverse ARP nazywa się odwzorowaniem dynamicznym. Natomiast trwałe skonfigurowanie powiązań jest opisywane jako odwzorowanie statyczne.

Pamiętaj, że każdy router operuje jedynie własnym końcem obwodu PVC. Zatem definiując powiązanie między adresem IP a identyfikatorem DLCI, użyj lokalnej wartości DLCI. Lokalny identyfikator DLCI można postrzegać jako wskaźnik zdalnego routera. Zależność ta staje się bardziej oczywista po przeanalizowaniu poleceń wykonywanych po stronie routerów B i C:

Router B:

```
Router-B(config-if)#frame-relay map ip 192.168.1.1 201 broadcast
```

Router C:

```
Router-C(config-if)#frame-relay map ip 192.168.1.1 301 broadcast
```

Na końcu instrukcji `frame-relay map` znajduje się słowo kluczowe `broadcast`. Należy je uwzględniać w każdej instrukcji tego typu. Zapewnia ono odwzorowanie również adresów rozgłoszeniowych i multiemisji, które są podstawą działania wielu protokołów routingu. Jeśli mimo uruchomionego połączenia Frame Relay nie możesz doprowadzić do ustanowienia relacji sąsiedzkich w protokołach EIGRP lub OSPF, sprawdź, czy uwzględniłeś w poleceniach `map` słowo kluczowe `broadcast`.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W odwzorowaniu IP-DLCI trzeba również uwzględnić pewien efekt uboczny. Mimo poprawnej konfiguracji sieci nie można wykonać polecenia `ping` w odniesieniu do własnego interfejsu:

```
Router-A#ping 192.168.1.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.1.1, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
```

Ograniczenie to bywa powodem irytacji, gdy podczas diagnozowania problemów nie można spingować własnego interfejsu tak jak w przypadku karty ethernetowej. Operacja taka nie jest możliwa z powodu braku wstępnie zdefiniowanego adresu warstwy drugiej. Interfejsy ethernetowe mają wstępnie określone adresy MAC, inaczej jest jednak w przypadku interfejsów Frame Relay. W sieciach Frame Relay wszystkie adresy warstwy drugiej są konfigurowane ręcznie.

Aby spingować własny interfejs, trzeba wcześniej powiązać jego adres IP ze zdalnym routerem. Choć rozwiązanie wydaje się niedorzeczne, działa poprawnie — gdy router zdalny odbierze dostarczone do niego pakiety, odeśle je z powrotem, bo ma zdefiniowane odwzorowanie adresu IP routera. Ponieważ nie ma lokalnego adresu warstwy drugiej, jest to jedyny sposób na wykonanie zadania (chmura Frame Relay rozgłasza jedynie identyfikator DLCI). Warto też pamiętać, że czasy dostarczania pakietów polecenia `ping` do własnego interfejsu odzwierciedlają czas transmisji w dwie strony w ramach zdefiniowanego obwodu PVC. Oto przykład:

```
Router-A(config-if)#frame-relay map ip 192.168.1.1 102

Router-A#ping 192.168.1.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 112/112/112 ms
```

Porównaj czas odpowiedzi z poprzednim przykładem. Jest on dwukrotnie dłuższy niż podczas pingowania routera zdalnego.

Odwzorowanie lokalnego adresu IP na identyfikator DLCI wskazujący zdalny węzeł nie gwarantuje dostępu do interfejsu w przypadku awarii zdalnego routera lub niedostępności obwodu PVC. Jeśli w przedstawionym przykładzie router B przestanie działać, router A nie będzie mógł pingować adresu IP własnego interfejsu S0/0 nawet w przypadku poprawnej komunikacji z routerem C.

Uważam za dobry nawyk własnoręczne definiowanie powiązań między identyfikatorami DLCI i adresami IP również w sytuacjach, w których router może zrobić to sam. Zapisywanie informacji o wartości DLCI w pliku konfiguracyjnym sprawia, że jest on łatwiejszy do analizowania i szybciej można w nim odnaleźć ewentualne błędy.

Podinterfejsy w sieciach Frame Relay

Często wyznaczanie jednego interfejsu jako punktu końcowego wielu obwodów PVC okazuje się rozwiązaniem niepożądanym. Jednocześnie przeznaczanie odrębnych interfejsów fizycznych poszczególnym obwodom podnosi koszty wdrożenia. Jako przykład rozważmy sieć przedstawioną na rysunku 23.11, w której zakończenia obydwu obwodów są zdefiniowane na tym samym interfejsie sieciowym. Jeśli w tak sformowanej sieci zostanie uruchomiony protokół

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

routingu, router B będzie dostarczał informacje do routera A, ale router A nie przekaże ich do routera C. Przyczyną jest działanie mechanizmu dzielonego horyzontu. Rozdzielenie obwodów PVC na odrębne interfejsy zapewnia poprawne rozgłaszanie aktualizacji w ramach protokołu routingu dynamicznego, ponieważ nie podlega ono ograniczeniom wynikającym ze stosowania techniki dzielonego horyzontu.

W urządzeniach Cisco funkcję tę realizują **podinterfejsy**. Ogólnie rzecz ujmując, podinterfejsy pozwalają na skonfigurowanie wirtualnych interfejsów odpowiadających poszczególnym obwodom PVC. Nazwy tych wirtualnych interfejsów pochodzą od nazwy interfejsu fizycznego, w którego ramach są definiowane. Na przykład podinterfejs działający w interfejsie S0/0 może mieć nazwę S0/0.100. Numer podinterfejsu jest określany przez administratora routera i może się zawierać w przedziale od 1 do 4 294 967 293. Ja nadaję podinterfejsom nazwy odzwierciedlające powiązane z nimi identyfikatory DLCI.

Podinterfejsy należą do jednej z dwóch kategorii: punkt-punkt (`point-to-point`) oraz wielopunkt (`multipoint`). Podinterfejsy typu punkt-punkt mogą mieć przypisany tylko jeden identyfikator DLCI. Natomiast interfejsy wielopunktowe umożliwiają zdefiniowanie większej ich liczby. Działają więc w sposób zbliżony do interfejsów fizycznych. Ponadto na jednym interfejsie fizycznym mogą działać zarówno podinterfejsy typu punkt-punkt, jak i podinterfejsy wielopunktowe. Nic nie stoi na przeszkodzie, aby pewne identyfikatory DLCI zostały przypisane do podinterfejsów, a inne do interfejsu fizycznego.

Zgodnie z wcześniejszymi informacjami jedną z największych zalet podinterfejsów Frame Relay jest możliwość użycia protokołów routingu bez problemu dzielonego horyzontu. Utworzenie wielu podinterfejsów typu punkt-punkt i przypisanie każdemu z nich innego obwodu PVC sprawia, że każdy obwód można utożsamiać z niezależnym interfejsem sieciowym. Do utworzenia podinterfejsów służy instrukcja konfiguracji ogólnej `interface`. Wystarczy określić nazwę tworzonego podinterfejsu i dopisać słowo kluczowe `point-to-point` lub `multipoint`.

```
Router-A(config)#int s0/0.102 point-to-point
Router-A(config-subif)#
```

Po wykonaniu tej operacji użytkownik pracuje w trybie konfiguracji podinterfejsu. Może więc korzystać z takich samych poleceń jak w przypadku definiowania parametrów interfejsu fizycznego.

Przemyśl wybór rodzaju podinterfejsu. Jeśli przez pomyłkę wybierzesz niewłaściwe ustawienia, jedynym sposobem przywrócenia odpowiedniej konfiguracji będzie zanegowanie instrukcji powołania podinterfejsu, zapisanie zmodyfikowanego pliku konfiguracyjnego i ponowne uruchomienie routera. Po usunięciu podinterfejsu Frame Relay na konsoli zostanie wyświetlony komunikat:

```
Not all config may be removed and may reappear after reactivating the sub-interface.
//Część konfiguracji mogła pozostać nieusunięta i może się uaktywnić po ponownym uruchomieniu
//podinterfejsu.
```

Następna czynność polega na przypisaniu obwodu wirtualnego do utworzonego podinterfejsu. Można w tym celu użyć polecenia `frame-relay interface-dlci` (wykonywanego w trybie konfiguracji podinterfejsu) lub instrukcji `frame-relay map` odpowiedzialnej za odwzorowanie adresu warstwy trzeciej na identyfikator DLCI. Jeśli przed dodaniem podinterfejsu identyfikator DLCI został przypisany do interfejsu fizycznego, operację trzeba poprzedzić usunięciem odwzorowań interfejsu fizycznego.

Powiązanie adresu IP z obwodem wirtualnym za pośrednictwem instrukcji `interface-dlci` jest zapisywane w nieco inny sposób, gdy dotyczy podinterfejsu:

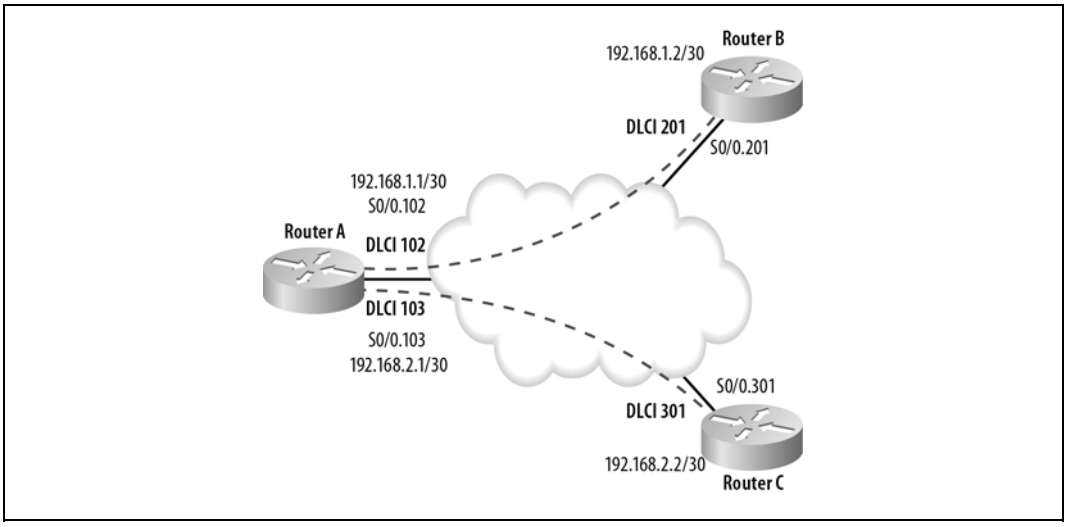
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
interface Serial0/0.102 point-to-point
frame-relay interface-dlci 102 protocol ip 192.168.1.2
```

Lubię tę metodę, ponieważ wprost z niej wynika, że do podinterfejsu został przypisany identyfikator DLCI oraz że identyfikator ten został powiązany z adresem IP. Użycie polecenia `map` nie dostarcza tak jednoznacznych informacji. Niemniej również spełnia swoje zadanie.

W przypadku podinterfejsów typu punkt-punkt nie trzeba odwzorowywać adresów IP na identyfikatory DLCI, ponieważ router wykryje, że zdalny węzeł korzysta z jedynego dostępnego wówczas adresu IP (zakładając, że maska podsieci ma wartość `255.255.255.252`). Pamiętaj, że definiując podinterfejsy typu punkt-punkt, musisz wyznaczyć osobną sieć IP dla każdego obwodu PVC.

Na rysunku 23.12 przedstawiono tę samą sieć, która widnieje na rysunku 23.11. Jednak tym razem każdy obwód PVC jest skojarzony z oddzielnym podinterfejsem Frame Relay.



Rysunek 23.12. Sieć Frame Relay złożona z trzech węzłów połączonych za pomocą podinterfejsów

Routery B i C w zasadzie nie potrzebują podinterfejsów, ale użycie interfejsu fizycznego wiązałoby się z koniecznością późniejszej zmiany konfiguracji po dodaniu obwodu PVC między węzłami B i C. Zastosowanie podinterfejsów ułatwi więc pracę w przyszłości.

Oto fragmenty plików konfiguracyjnych wszystkich trzech routerów.

Router A:

```
interface Serial0/0
 no ip address
 encapsulation frame-relay
!
interface Serial0/0.102 point-to-point
 ip address 192.168.1.1 255.255.255.252
 frame-relay interface-dlci 102 protocol ip 192.168.1.2
!
interface Serial0/0.103 point-to-point
 ip address 192.168.2.1 255.255.255.252
 frame-relay interface-dlci 103 protocol ip 192.168.2.2
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Router B:

```
interface Serial0/0
 no ip address
 encapsulation frame-relay
!
interface Serial0/0.201 point-to-point
 ip address 192.168.1.2 255.255.255.252
 frame-relay interface-dlci 201 protocol ip 192.168.1.1
```

Rotuer C:

```
interface Serial0/0
 no ip address
 encapsulation frame-relay
!
interface Serial0/0.301 point-to-point
 ip address 192.168.2.2 255.255.255.252
 frame-relay interface-dlci 301 protocol ip 192.168.2.1
```

Rozwiązywanie problemów z protokołem Frame Relay

Rozwiązywanie problemów w działaniu mechanizmu Frame Relay nie stanowi problemu dla osób, które wiedzą, jak on działa. Pamiętaj, że większość informacji dotyczących obwodów PVC pochodzi z urządzenia DCE będącego zazwyczaj przełącznikiem telekomunikacyjnym działającym po drugiej stronie łącza fizycznego.

Każda operacja poszukiwania błędów wymaga **wyizolowania problemu**. Ustalenie przyczyny problemu pozwala na wdrożenie odpowiednich środków naprawczych. Wykonując poniższą procedurę, będziesz mógł szybko zlokalizować błąd.

Najpierw warstwa fizyczna!

Czy kabel jest podłączony? Czy jest to właściwy, sprawdzony wcześniej kabel? Czy po drugiej stronie kabel również został podłączony? Choć takie rady mogą się wydawać niedorzeczne, warto je uwzględnić. Możesz poczuć się naprawdę głupio, gdy po wezwaniu techników Cisco okaże się, że przyczyną problemu był niepodłączony kabel.

Czy połączenie szeregowe działa?

Wykonaj polecenie `show interface` i upewnij się, że połączenie szeregowe działa poprawnie (ma status `up`). Pozostawienie interfejsu w stanie `shutdown` skutecznie zapobiegnie przekazywaniu jakiegokolwiek ruchu.

Czy odbierasz komunikaty LMI?

Pamiętaj, że komunikaty LMI są generowane przez urządzenie telekomunikacyjne podłączone do lokalnego routera. Jeśli nie odbierasz powiadomień LMI, nie otrzymasz informacji na temat statusu obwodów wirtualnych. Router nie będzie więc wiedział o ich istnieniu. Jest kilka sposobów sprawdzenia, czy komunikaty LMI docierają do routera:

`show interface`

Wynik wykonania instrukcji `show interface` w odniesieniu do interfejsu Frame Relay obejmuje liczniki komunikatów LMI. Powiadomienia LMI są generowane co 10 sekund. Dwukrotne wprowadzenie instrukcji w odstępie co najmniej 10 sekund pozwala na sprawdzenie, czy wspomniane wartości są zwiększane:

```
Router-A#sho int s0/0 | include LMI
  LMI enq sent 186, LMI stat recvd 186, LMI upd recvd 0, DTE LMI up
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  LMI enq recvd 0, LMI stat sent 0, LMI upd sent 0
  LMI DLCI 1023 LMI type is CISCO frame relay DTE
Router-A#
Router-A#
Router-A#sho int s0/0 | include LMI
  LMI enq sent 188, LMI stat recvd 188, LMI upd recvd 0, DTE LMI up
  LMI enq recvd 0, LMI stat sent 0, LMI upd sent 0
  LMI DLCI 1023 LMI type is CISCO frame relay DTE
```

```
debug frame-relay lmi
```

Polecenie `debug frame-relay lmi` włącza wyświetlanie każdego powiadomienia LMI dostarczanego do interfejsów Frame Relay. Jak zawsze, wykonując instrukcję z grupy `debug` w routerze użytkowym, trzeba zachować szczególną ostrożność. Wprowadzenie jej w mocno obciążonym routerze może spowodować wstrzymanie pracy urządzenia w wyniku nadmiernego wykorzystania procesora. Zatwierdzenie polecenia `debug frame-relay lmi` powinno spowodować wyświetlanie krótkich komunikatów statusowych (które same w sobie nie niosą zbyt wielu użytecznych informacji) w 10-sekundowych interwałach, a także generowanie co 30 sekund zestawienia wszystkich obwodów wirtualnych obsługiwanych w danym łączu. Komunikaty te można łatwo wyodrębnić na podstawie początkowego fragmentu (`PVC` lub `SVC`):

```
Router-A#debug frame lmi
Frame Relay LMI debugging is on
Displaying all Frame Relay LMI data
Router-A#
00:33:05: Serial0/0(out): StEnq, myseq 197, yourseen 196, DTE up
00:33:05: datagramstart = 0×3CE9B74, datagramsize = 13
00:33:05: FR encap = 0×FCF10309
00:33:05: 00 75 01 01 01 03 02 C5 C4
00:33:05:
404 | Chapter 23: Frame Relay
00:33:05: Serial0/0(in): Status, myseq 197, pak size 13
00:33:05: RT IE 1, length 1, type 1
00:33:05: KA IE 3, length 2, yourseq 197, myseq 197
00:33:15: Serial0/0(out): StEnq, myseq 198, yourseen 197, DTE up
00:33:15: datagramstart = 0×3CEA1B4, datagramsize = 13
00:33:15: FR encap = 0×FCF10309
00:33:15: 00 75 01 01 00 03 02 C6 C5
00:33:15:
00:33:15: Serial0/0(in): Status, myseq 198, pak size 53
00:33:15: RT IE 1, length 1, type 0
00:33:15: KA IE 3, length 2, yourseq 198, myseq 198
00:33:15: PVC IE 0×7 , length 0×6 , dlci 102, status 0×2 , bw 0
00:33:15: PVC IE 0×7 , length 0×6 , dlci 103, status 0×2 , bw 0
00:33:15: PVC IE 0×7 , length 0×6 , dlci 104, status 0×0 , bw 0
00:33:15: PVC IE 0×7 , length 0×6 , dlci 105, status 0×0 , bw 0
00:33:15: PVC IE 0×7 , length 0×6 , dlci 106, status 0×0 , bw 0
00:33:25: Serial0/0(out): StEnq, myseq 199, yourseen 198, DTE up
00:33:25: datagramstart = 0×3CEA574, datagramsize = 13
00:33:25: FR encap = 0×FCF10309
00:33:25: 00 75 01 01 01 03 02 C7 C6
```

W przedstawionym przykładzie przełącznik Frame Relay rozgłasza pięć obwodów PVC. Status poszczególnych z nich jest opisany jako `0x0` lub `0x2`. Wartość `0x0` oznacza, że obwód jest skonfigurowany, ale nie jest aktywny. Przyczyną takiego stanu jest zazwyczaj nieodpowiednia konfiguracja urządzenia zdalnego (innymi słowy, jest to Twój błąd, a nie operatora telekomunikacyjnego). Status `0x2` wskazuje, że obwód wirtualny jest skonfigurowany i aktywny. Jeśli określony obwód nie jest wymieniony w żadnym z wyświetlonych komunikatów, najprawdopodobniej nie został udostępniony przez operatora telekomunikacyjnego lub został włączony w niewłaściwym przełączniku bądź interfejsie.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Czy obwody wirtualne zostały uaktywnione w routerze?

Wykonanie instrukcji `show frame-relay pvc` pozwala na sprawdzenie statusu wszystkich zdefiniowanych w routerze obwodów PVC mechanizmu Frame Relay:

```
Router-A#sho frame pvc

PVC Statistics for interface Serial0/0 (Frame Relay DTE)

            Active   Inactive    Deleted     Static
  Local        2          0          0          0
  Switched     0          0          0          0
  Unused       0          3          0          0

DLCI = 102, DLCI USAGE = LOCAL, PVC STATUS = ACTIVE, INTERFACE = Serial0/0.102
  input pkts 46          output pkts 55       in bytes 11696
  out bytes 14761        dropped pkts 0       in pkts dropped 0
  out pkts dropped 0     out bytes dropped 0
  in FECN pkts 0         in BECN pkts 0       out FECN pkts 0
  out BECN pkts 0        in DE pkts 0         out DE pkts 0
  out bcast pkts 45      out bcast bytes 13721
  pvc create time 00:44:07, last time pvc status changed 00:44:07
DLCI = 103, DLCI USAGE = LOCAL, PVC STATUS = ACTIVE, INTERFACE = Serial0/0.103

  input pkts 39          output pkts 47       in bytes 11298
  out bytes 13330        dropped pkts 0       in pkts dropped 0
  out pkts dropped 0     out bytes dropped 0
  in FECN pkts 0         in BECN pkts 0       out FECN pkts 0
  out BECN pkts 0        in DE pkts 0         out DE pkts 0
  out bcast pkts 42      out bcast bytes 12810
  pvc create time 00:39:13, last time pvc status changed 00:39:13

DLCI = 104, DLCI USAGE = UNUSED, PVC STATUS = INACTIVE, INTERFACE = Serial0/0
  input pkts 0           output pkts 0        in bytes 0
  out bytes 0            dropped pkts 0       in pkts dropped 0
  out pkts dropped 0     out bytes dropped 0
  in FECN pkts 0         in BECN pkts 0       out FECN pkts 0
  out BECN pkts 0        in DE pkts 0         out DE pkts 0
  out bcast pkts 0       out bcast bytes 0
  switched pkts 0
  Detailed packet drop counters:
  no out intf 0          out intf down 0      no out PVC 0
  in PVC down 0          out PVC down 0       pkt too big 0
  shaping Q full 0       pkt above DE 0       policing drop 0
  pvc create time 00:44:01, last time pvc status changed 00:44:01
```

[fragment usunięty]

Każdy obwód PVC jest opisany w osobnej sekcji, która obejmuje informacje na temat statusu obwodu oraz interfejsu, na którym ten obwód został wykryty. Zwróć uwagę na fakt, że obwody PVC przypisane do podinterfejsów są wykazywane jako działające na tych podinterfejsach. Wszystkie pozostałe obwody PVC są skojarzone z interfejsami fizycznymi, na których zostały wykryte. Jeśli dany obwód PVC nie został wyświetlony w zestawieniu wynikowym, najprawdopodobniej router nie odbiera komunikatów LMI z nim związanych.

Każdy wpis zawiera informacje statusowe o następujących znaczeniach:

Active

Status ten informuje, że dany obwód PVC działa poprawnie na całym odcinku między urządzeniami końcowymi.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Inactive

Status ten informuje, że dany obwód wirtualny został zdefiniowany w przełączniku operatora telekomunikacyjnego, ale w routerze nie istnieje aktywne odwzorowanie jego identyfikatora DLCI. Jeśli jest to obwód, z którego starasz się skorzystać, prawdopodobnie zapomniałeś o jego odwzorowaniu lub jest ono niepoprawnie zdefiniowane.

Deleted

Status ten informuje, że korzystasz z urządzenia, które ma aktywne odwzorowanie, ale wskazywany obwód PVC nie istnieje. Przyczyną problemu mogą być niewłaściwie określone parametry powiązania.

Static

Status ten informuje, że interfejs routera został skonfigurowany z użyciem instrukcji `no keepalive`.

Czy obwód PVC jest powiązany z interfejsem?

Polecenie `show frame-relay map` generuje zwięzły raport na temat obwodów wirtualnych powiązanych z interfejsami:

```
Router-A#sho frame map
Serial0/0.103 (up): point-to-point dlci, dlci 103(0×67,0×1870), broadcast
          status defined, active
Serial0/0.102 (up): point-to-point dlci, dlci 102(0×66,0×1860), broadcast
          status defined, active
```

Jeśli obwód PVC nie jest wymieniony w zestawieniu, ale występuje na liście wynikowej polecenia `show frame-relay pvc`, router nie jest poprawnie skonfigurowany.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 24.

MPLS

Z mojego doświadczenia wynika, że w konfigurację mechanizmu MPLS (przełączania etykiet niezależnego od protokołu — ang. *Multiprotocol Label Switching*) są zaangażowane dwie grupy osób. Pierwsza z nich odpowiada za projektowanie i utrzymanie sieci MPLS. Są to zazwyczaj pracownicy takich firm jak AT&T, British Telecom itp. Drugą grupę stanowią administratorzy sieci, którzy korzystają z techniki MPLS w korporacyjnych sieciach WAN. Ten rozdział jest przeznaczony właśnie dla nich (dla osób, które muszą zastosować rozwiązania MPLS we własnych sieciach, a nie dla projektantów samego systemu MPLS). Dlaczego? Ponieważ odbiorcami tej książki są przede wszystkim administratorzy, którzy codziennie zmagają się z sieciami korporacyjnymi. Moim celem nie jest więc przedstawienie zasad budowy sieci MPLS, ale ułatwienie Ci wdrożenia rozwiązania MPLS przejętego od dostawcy usług internetowych. Dlatego nie spotkasz się tutaj z technicznym żargonem charakterystycznym dla mechanizmów MPLS. Znajomość wielu zagadnień nie jest bowiem potrzebna do uruchomienia sieci WAN w technologii MPLS. Moim celem jest natomiast przedstawienie sposobu na wykonanie określonego zadania.

Ogólnie rzecz ujmując, w sieciach MPLS pakiety są poprzedzane nagłówkiem MPLS (nazywanym **stosem etykiet**). Zawiera on co najmniej jedną **etykietę**. Każda etykieta z kolei przechowuje wartość etykiety, pole klasy ruchu (wykorzystywane do zapewnienia odpowiedniej jakości usługi (QoS)), znacznik końca stosu oraz 8-bitowe pole czasu życia pakietu (TTL). Brzmi ekscytująco, prawda? Dla większości osób skrót MPLS nic nie znaczy. Dlatego dalsze omówienie rozpoczniemy od zapoznania się z historią tego rozwiązania.

Stos etykiet jest jedynym obszarem pakietu analizowanym przez przełączniki MPLS. Jego zastosowanie nie wymaga więc analizowania tablic routingu, co teoretycznie wiąże się ze znacznie efektywniejszym przekazywaniem danych niż w tradycyjnych rozwiązaniach IP. Jednak od czasu opracowania techniki MPLS wydajność urządzeń znacząco wzrosła, więc argument o zwiększonej szybkości pracy nie jest aż tak istotnym czynnikiem w upowszechnianiu się tego rozwiązania.

W czasie prac nad systemem MPLS w sieciach WAN dominowały technologie ATM i Frame Relay, które miały wiele niedoskonałości. Na przykład największą wadą rozwiązania ATM był znaczny narzut transmisyjny, który czynił ten mechanizm mało atrakcyjnym. W systemie MPLS zaoferowano znacznie więcej niż tylko zwiększoną wydajność i szybkość transmisji.

Dodawanie do każdego pakietu nagłówka MPLS powoduje, że sieci tego typu mogą obsługiwać wiele połączeń jako wirtualne sieci w ramach chmury MPLS. Ponadto w urządzeniach końcowych

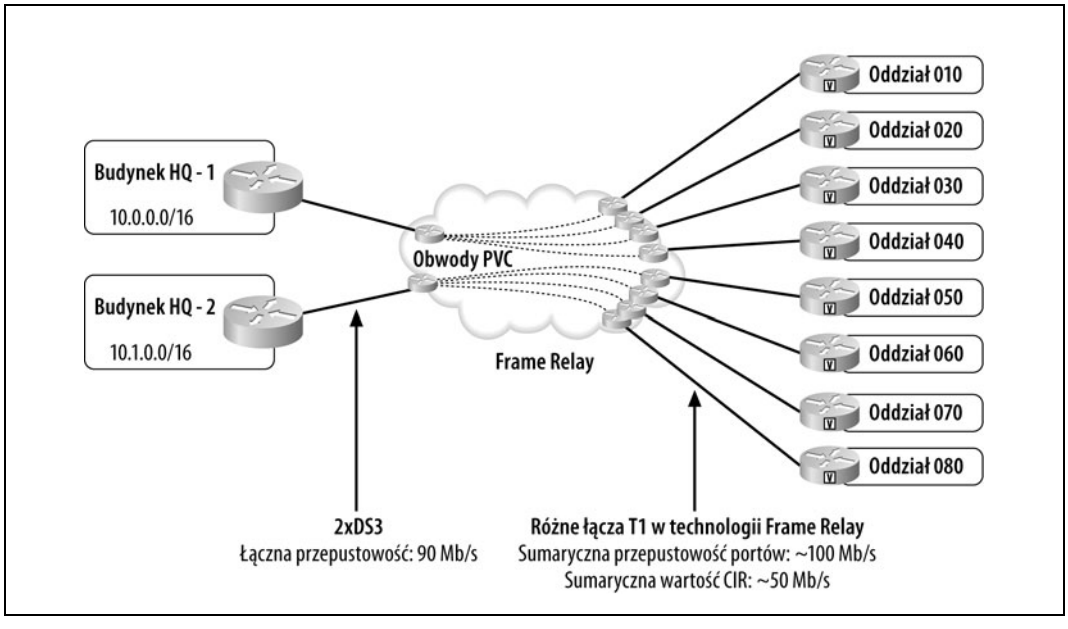
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

(zarządzanych przez użytkowników usługi) nie ma wprost wydzielonych obwodów wirtualnych. Korzystanie z rozwiązania MPLS przypomina więc w pewnym sensie korzystanie z prywatnego internetu, w którym dostępne są jedynie punkty końcowe wskazane przez abonenta usługi.

Przejdźmy do sedna zagadnienia. W czym rozwiązanie MPLS jest lepsze od Frame Relay lub ATM? Będę szczery — większość firm wdraża technikę MPLS z powodów, które nie mają nic wspólnego z technologią. Zazwyczaj główną przyczyną jest niższy koszt. W ostatnich latach systemy MPLS były bardzo intensywnie promowane przez wielu operatorów telekomunikacyjnych i stały się znacznie korzystniejsze cenowo od konkurencyjnych technologii. W przeszłości rozwiązania MPLS były uznawane za usługę specjalistyczną, z której korzystały firmy gotowe zapłacić dodatkową cenę za jej udostępnienie. Jednak obniżenie kosztów jej świadczenia spowodowało, że weszły do podstawowej oferty wielu dostawców usług internetowych. Osobiście odpowiadałem za wdrożenie techniki MPLS w sieci złożonej z 300 węzłów, gdy dobiegł końca kontrakt na usługę Frame Relay. Cena dostarczenia usługi T1 do każdego oddziału wyposażonego w dwa łącza OC3 była znacznie niższa niż cena za częściową przepustowość T1 zrealizowaną w technologii Frame Relay na łączu DS3. Firma zyskała więc znacznie szybsze połączenia za atrakcyjniejszą cenę. Kto nie skorzystałby z takiej oferty?

Szczególnym wyzwaniem okazało się zintegrowanie nowych rozwiązań z istniejącą siecią. Systemy MPLS wykorzystują zupełnie inną metodę definiowania połączeń logicznych niż Frame Relay. Z tego powodu zmiana technologii bywa dość kłopotliwa.

Połączenia Frame Relay bazują na obwodach wirtualnych. Podłączenie 300 zdalnych oddziałów do jednego z dwóch budynków siedziby firmy oznaczało więc konieczność konfigurowania i utrzymywania 300 obwodów wirtualnych. Topologia sieci obowiązująca w czasie działania systemu Frame Relay została przedstawiona na rysunku 24.1.

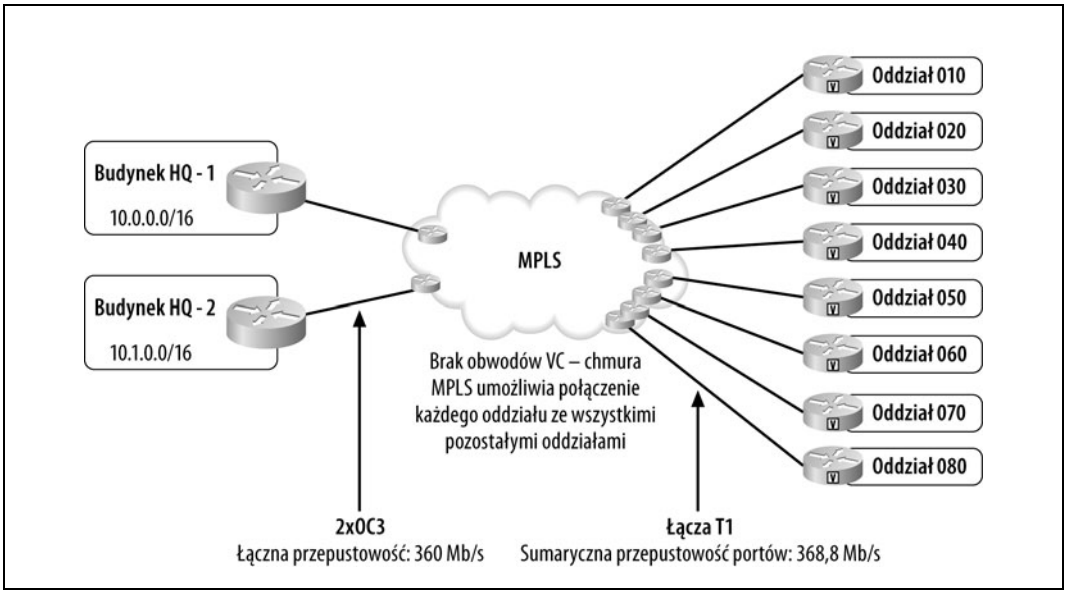


Rysunek 24.1. Połączenia Frame Relay z oddziałami zdalnymi

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Sieć działała poprawnie, ale miała kilka istotnych ograniczeń. Na przykład awaria systemu HQ-1 powodowała, że oddziały z nim połączone stawały się niedostępne. Nie było bowiem mechanizmu dynamicznego przełączania ruchu do systemu HQ-2. Gdyby każdy z oddziałów był połączony z obydwoma budynkami siedziby głównej, konieczne byłoby zdefiniowanie kolejnych obwodów PVC, a to podniosłoby koszt użytkowania sieci. Dlatego ryzyko utraty połączeń z połową zdalnych sieci zostało uznane za akceptowalne (cena nadmiarowości była zbyt wysoka). Jako łącza zapasowe wykorzystywano linie ISDN, które były przyczyną kolejnych problemów — łącza ISDN były (i są) bardzo kosztowne. Budowa sieci została zaproponowana przed upowszechnieniem się łączy szerokopasmowych, więc musieliśmy korzystać z kart, które były w owym czasie dostępne na rynku.

Zastosowanie techniki MPLS usprawniło działanie sieci, w czym niemałe znaczenie miało zwiększenie szybkości przesyłania danych. Jednak najważniejszą zmianą była dla nas możliwość bezpośredniego przyłączenia każdego z oddziałów do obydwu budynków głównej siedziby bez konieczności definiowania dodatkowych obwodów PVC. Ponadto każdy z oddziałów zyskał bezpośrednie połączenie z innymi oddziałami firmowymi (bez pośrednictwa urządzeń HQ), co w dużej mierze było zasługą technologii MPLS. To z kolei umożliwiło późniejsze wdrożenie systemu VoIP. Budowa sieci z zastosowaniem mechanizmów MPLS została przedstawiona na rysunku 24.2. Sama topologia nie zmieniła się istotnie. Każda sieć zdalna nadal jest podłączona do chmury. Różnica polega jedynie na usunięciu obwodów wirtualnych. Każdy oddział jest rzeczywiście przyłączony do chmury, a nie do wstępnie skonfigurowanego (lub dynamicznie narzuconego) połączenia wirtualnego, które prowadzi do innej sieci. Oczywiście trzeba pamiętać, że w chmurze są zdefiniowane trasy MPLS, co zwiększa złożoność rozwiązania, ale z perspektywy odbiorcy końcowego eliminuje obowiązek konfigurowania, utrzymywania i monitorowania obwodów wirtualnych.



Rysunek 24.2. System MPLS umożliwiający zdefiniowanie pełnej siatki połączeń

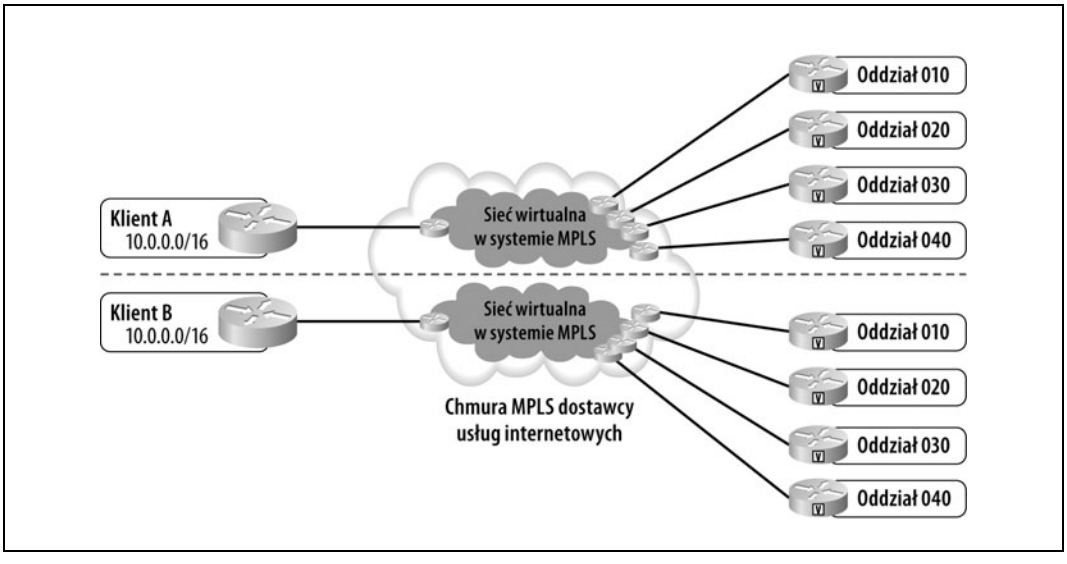
W jaki sposób poszczególne oddziały komunikują się ze sobą, skoro nie ma między nimi połączeń? Aby uzyskać odpowiedź na to pytanie, trzeba wiedzieć, w jaki sposób działają sieci MPLS

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

(poznać zasady przełączania pakietów na podstawie etykiet) oraz jakie informacje są potrzebne odbiorcy końcowemu. Z perspektywy każdego oddziału zdalnego niezbędne są jedynie dane na temat adresu kolejnego routera na trasie do sieci docelowej, które są dostarczane przez protokoły routingu IP.

Zaraz, czy nie stwierdziłem wcześniej, że tablice routingu IP nie są do niczego potrzebne? Tak. Na poziomie mechanizmów przekazywania pakietów w sieci operatora telekomunikacyjnego tablice routingu routerów końcowych nie mają żadnego znaczenia. Cała sieć WAN, protokół BGP i wszystkie związane z nim mechanizmy są elementami wirtualnej sieci, którą operator telekomunikacyjny wydziela w ramach własnej sieci MPLS. Urządzenia telekomunikacyjne działają jedynie na podstawie etykiet, a sieci abonentów usługi nie oznaczają pakietów za pomocą takich etykiet. Dlatego pakiety dostarczane do sieci operatora są na wejściu oznaczane odpowiednimi etykietami i przekazywane w chmurze na podstawie ich wartości. Zawartość pakietu nie ma żadnego znaczenia. Liczy się jedynie etykieta.

Sposób działania systemu został zilustrowany na rysunku 24.3. Przedstawiono na nim dwóch abonentów usługi MPLS przyłączonych do sieci jednego operatora. Obydwaj klienci wykorzystują taką samą przestrzeń adresową IP (10.0.0.0/16). Są przyłączeni do tego samego systemu MPLS. Jednak ich routery nie mogą wymieniać danych między sobą. Dlaczego? Każdemu z odbiorców odpowiadają inne etykiety. Działanie sieci jest zbliżone do rozwiązania VPN, w którym połączenia można ustanawiać jedynie z własnymi urządzeniami mimo współdzielenia zasobów sieciowych z innymi użytkownikami. Wiele osób opisuje nawet sieci MPLS jako systemy VPN. Nie można jednak nadużywać tego nazewnictwa, ponieważ w połączeniach MPLS nie stosuje się szyfrowania danych.



Rysunek 24.3. Sieci wirtualne w systemie MPLS

Dla użytkownika końcowego liczy się przede wszystkim to, że działanie mechanizmu jest przezroczyste. Abonent usługi musi wiedzieć o konieczności dodawania etykiet i wydzielania sieci wirtualnych. Często nie ma nawet świadomości, że to zadanie jest realizowane z wykorzystaniem technologii MPLS. Jedyną zauważalną zmianą w stosunku do klasycznych rozwiązań jest stosowanie protokołu BGP (opisane w dalszej części rozdziału).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Inną zaletą technologii MPLS jest wbudowana obsługa mechanizmów QoS. Dzięki temu, że jedna z etykiet stosu etykiet MPLS zawiera pole klasy ruchu, sieci MPLS umożliwiają rzeczywiste kategoryzowanie strumieni danych. Liczba klas zależy od dostawcy usługi. Każdy z operatorów stosuje też różne systemy klasyfikacji. Z kolei możliwość budowania określonej hierarchii priorytetów powoduje, że rozwiązanie MPLS staje się niezwykle atrakcyjne dla firm, które chcą skorzystać z technologii VoIP w sieciach WAN. Korporacje administrujące tysiącami węzłów mogą odczuć istotne usprawnienie funkcjonowania sieci, szczególnie jeśli wcześniej korzystały z protokołu Frame Relay.

Jak więc konfiguruje się połączenia MPLS? Krótko mówiąc, nie konfiguruje się ich wcale. W przeciwieństwie do operatora telekomunikacyjnego odbiorca usługi nie musi definiować żadnych ustawień związanych z działaniem systemu MPLS. Wszystkie operacje są realizowane w urządzeniach dostawcy usług internetowych. Jest to duże ułatwienie w porównaniu z połączeniami Frame Relay i ATM, które wymagają specjalistycznej wiedzy i znacznych nakładów pracy. W wielu implementacjach MPLS niezbędna jest znajomość protokołu BGP oraz mechanizmów QoS (jeśli konieczne jest ich wykorzystanie). Niektórzy dostawcy usług umożliwiają klientom zastosowanie protokołu EIGRP zamiast BGP, ale nie polecam takiego rozwiązania. Bez wątpienia jest ono łatwiejsze do wdrożenia, szczególnie dla osób, które nie korzystały wcześniej z mechanizmów BGP. Jednak w praktyce oznacza wyrażenie zgody na to, by operator dokonał redystrybucji tras między protokołami BGP i EIGRP, co z kolei pozbawia administratora kontroli nad routingiem. W środowiskach MPLS o małej złożoności często nie jest to problemem. Inaczej jest w bardziej rozbudowanych sieciach. Niedopracowana redystrybucja może być źródłem wielu zmartwień.

Poniżej została przedstawiona przykładowa konfiguracja routera w oddziale o numerze 10 (z omawianego wcześniej przykładu).

W pierwszym kroku zajmę się mechanizmem QoS (zachowując prostotę rozwiązania). Ruch VoIP w protokole RTP oraz sterowanie połączeniami głosowymi otrzymają wyższy priorytet. Pozostałe formy ruchu zostaną przekazane do domyślnej kolejki. Kolejka priorytetowa może zająć do 40 procent przepustowości łącza, natomiast sterowanie połączeniami otrzyma 5 procent pasma. Więcej informacji na temat technik QoS znajduje się w rozdziałach 31. i 32.

```
class-map match-any VoIP-RTP
 match ip dscp ef
class-map match-any VoIP-Call-Control
 match ip dscp cs3
 match ip dscp af31
!
policy-map MPLS-QoS
 class VoIP-RTP
  priority percent 40
 class VoIP-Call-Control
  bandwidth percent 5
 class class-default
  fair-queue
```

Oto konfiguracja połączenia MPLS (zauważ, że nie ma w niej żadnych elementów szczególnych dla sieci MPLS):

```
interface Serial0/3/0
 description [ Osrodek 10 MPLS ]
 ip address 10.255.10.2 255.255.255.252
 encapsulation ppp
 auto qos voip trust
 service-policy output MPLS-QoS
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Konfiguracja wewnętrznego interfejsu ethernetowego:

```
interface FastEthernet0/0
 description [ Osrodek 10 LAN ]
 ip address 10.10.10.1 255.255.255.128
 duplex auto
 speed auto
 auto qos voip trust
 service-policy output MPLS-QoS
```

Następnie zdefiniuję adres pętli zwrotnej potrzebny w procesie routingu. Przyda się do zapewnienia awaryjnego połączenia VPN (nieopisanego tutaj):

```
interface Loopback0
 description [ Osrodek 10 Loopback ]
 ip address 10.10.10.129 255.255.255.255
```

Na koniec trzeba wprowadzić pewne ustawienia protokołu BGP, dzięki którym możliwe będzie pozyskiwanie informacji o trasach w chmurze oraz wysyłanie tego typu powiadomień.

```
router bgp 65010
 no synchronization
 bgp log-neighbor-changes
 network 10.10.10.0 mask 255.255.255.128
 network 10.10.10.129 mask 255.255.255.255
 aggregate-address 10.10.10.0 255.255.255.0 summary-only
 neighbor 10.255.255.1 remote-as 65001
 neighbor 10.255.255.1 update-source Serial0/3/0
 neighbor 10.255.255.1 version 4
 neighbor 10.255.255.1 prefix-list Aggregate out
 no auto-summary
!
ip prefix-list Aggregate seq 5 permit 10.10.10.0/24
```

Jak nietrudno zauważyć, w całej konfiguracji nie ma żadnych ustawień związanych jedynie z mechanizmem MPLS. Nie ma tu wzmianek o etykietach lub o czymkolwiek innym niż routing, QoS i adresy IP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 25.

Listy dostępu

Formalna nazwa list dostępu to **listy kontroli dostępu** (ang. *access control list*). Z kolei pojedyncze elementy listy kontroli dostępu nazywają się **wpisami kontroli dostępu** (ACE — ang. *Access Control Entries*). Sam termin lista kontroli dostępu w praktyce często jest skracany do określenia „lista dostępu" lub „lista ACL".

Listy dostępu nie służą jedynie do regulowania praw dostępu. Są mechanizmami kategoryzowania i dopasowywania pakietów, które znajdują wiele zastosowań w pracy urządzeń firmy Cisco. Można je wykorzystać jako proste filtry przepuszczające ruch przez interfejsy. Nadają się również do wyodrębniania ruchu w ramach map interfejsów ISDN oraz dopasowywania pakietów w definicji map routingu.

Definiowanie list dostępu

Informacje zamieszczone w tym rozdziale w mniejszym stopniu dotyczą podstaw konstruowania list ACL. Mają na celu przede wszystkim uświadomienie Ci korzyści wynikających z ich stosowania oraz przedstawienie typowych problemów w definiowaniu ich treści. Kilka przedstawionych tutaj sztuczek powinno ułatwić Ci tworzenie poprawnych, wydajnych i efektywnych list dostępu.

Opracowując listy ACL (lub ustawienia z nimi związane), warto zapisywać je najpierw w edytorze tekstu, a później testować w środowisku laboratoryjnym. Wprowadzanie jakichkolwiek zmian w filtrach pakietów zawsze wiąże się z ryzykiem zablokowania dostępu do sieci.

Nazywanie czy numerowanie

Większość urządzeń Cisco obsługuje zarówno numerowane, jak i nazywane listy ACL. Listy nazywane są wzmiankowane w pliku konfiguracyjnym za pomocą przypisanych im nazw, takich jak `GAD` lub `ObronaSieci`. Numerowanie list jest starszym sposobem ich deklarowania i sprowadza się do przypisywania listom wartości liczbowych, takich jak `101`. Jeśli to tylko możliwe, warto więc stosować listy nazywane. W niektórych urządzeniach (takich jak przełączniki Nexus) numerowanie w ogóle nie jest dozwolone.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Maski wieloznaczne

Maski wieloznaczne (ang. *wildcard mask*) (nazywane również **maskami odwrotnymi**) są wykorzystywane w wielu urządzeniach do definiowania list dostępu. Ich zapis często bywa mylący, ponieważ stanowi zazwyczaj negację (binarną) standardowych masek podsieci. Innymi słowy, maska wieloznaczna o wartości `0.0.0.255` może posłużyć do wyodrębnienia adresów odpowiadających zastosowaniu maski podsieci o wartości `255.255.255.0`.

Poniższa zasada pozwala na rozwiązanie większości problemów z doborem masek podsieci i masek wieloznacznych:

> Zastępuj wszystkie zera wartościami `255`, a wszystkie wartości `255` zerami.

Zależności między maskami podsieci klas A, B i C oraz maskami wieloznacznymi zostały przedstawione w tabeli 25.1.

Tabela 25.1. Pełnoklasowe maski wieloznaczne

Maska podsieci	Maska wieloznaczna
`255.0.0.0`	`0.255.255.255`
`255.255.0.0`	`0.0.255.255`
`255.255.255.0`	`0.0.0.255`

Oczywiście w praktyce większość sieci nie zawiera się w pełnych klasach adresowych. Jako przykład przeanalizujmy maskę podsieci o wartości `255.255.255.224`. Odpowiadająca jej maska wieloznaczna ma wartość `0.0.0.31`.

W wyznaczaniu masek wieloznacznych bardzo pomocna jest pewna sztuczka, która sprawia, że całe zadanie staje się bardzo łatwe.

> Maskę wieloznaczną można wyliczyć, odejmując jeden od liczby stacji, która wynika z wartości maski podsieci.

Maska podsieci wspomniana w poprzednim przykładzie (`255.255.255.224`) wyznacza sieć z 32 stacjami (więcej informacji na temat sposobu obliczania liczby stacji w wydzielanych podsieciach zostało zamieszczonych w rozdziale 36.). Zatem obliczenie ostatniego oktetu sprowadza się do wykonania działania 32–1 = 31. Maska wieloznaczna ma więc wartość `0.0.0.31`. Tak, to naprawdę jest takie proste.

Chwili zastanowienia wymagają więc jedynie te oktety, które nie mają wartości `0` albo `255`. Jeśli maska kończy się na pozycjach innych niż ostatni oktet, można zastosować tę samą technikę — tak jakby oktet graniczny był ostatnim oktetem maski. Oto przykład obliczenia dla maski podsieci o wartości `255.240.0.0`:

1. Zastąp wszystkie oktety zerowe wartościami `255`, a wszystkie wartości `255` zerami. W wyniku otrzymasz `0.240.255.255`.
2. Gdyby wartość `240` była zapisana jako ostatni oktet maski (`255.255.255.240`), definiowałaby sieć złożoną z 16 stacji.
3. 16–1 = 15. Zapisz tę wartość zamiast oktetu niezamienionego w kroku 1. Wynikiem jest `0.15.255.255`.
4. Maska wieloznaczna ma wartość `0.15.255.255`.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Po pewnym czasie przekształcanie wartości okaże się znacznie łatwiejsze. Poćwicz trochę, by nabrać wprawy.

Sprawę komplikuje nieco to, że w urządzeniach Cisco Adaptive Security Appliances (ASA) zamiast odwrotnych masek stosowane są standardowe maski podsieci. Oto przykładowa lista dostępu z pliku konfiguracyjnego urządzenia ASA 5540:

```
access-list GAD extended permit tcp any 10.0.0.0 255.255.255.0 eq www
```

Poniżej znajduje się natomiast analogiczna lista dostępu zdefiniowana w przełączniku 3750:

```
ip access-list extended GAD
permit tcp any 10.10.10.0 0.0.0.255 eq www
```

Zwróć uwagę na wyróżnione fragmenty. W urządzeniu ASA zakres adresów został opisany za pomocą standardowej maski podsieci. Natomiast w przełączniku 3750 użyta została maska odwrócona. Na szczęście wykorzystanie niewłaściwego formatu zapisu w którymkolwiek z wymienionych systemów spowoduje wygenerowanie komunikatu o błędzie. Nie zmienia to jednak faktu, że taki brak spójności zapisu może być irytujący.

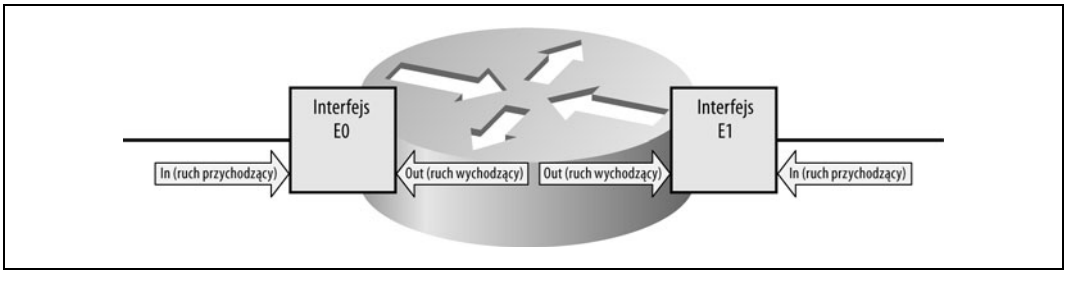
W systemach NX-OS zamiast masek podsieci lub masek wieloznacznych można stosować w listach dostępu prefiksy sieci. Osobiście uważam ten sposób zapisu za najkorzystniejszy i mam nadzieję, że zostanie on wprowadzony we wszystkich urządzeniach:

```
ip access-list GAD
10 permit tcp 10.10.10.0/24 any eq www
```

Zastosowanie list dostępu

Jednym z pytań, które najczęściej zadają mi początkujący administratorzy, jest: „czy powinienem skojarzyć listę z ruchem wchodzącym, czy wychodzącym?”. Niemal zawsze odpowiedzią jest **wychodzącym.**

Na rysunku 25.1 przedstawiono router z dwoma interfejsami (E0 i E1). Zaznaczono również punkty, w których lista dostępu może zostać zastosowana. Trzeba jednak pamiętać, że określając kierunek ruchu, należy spojrzeć na problem z perspektywy urządzenia.



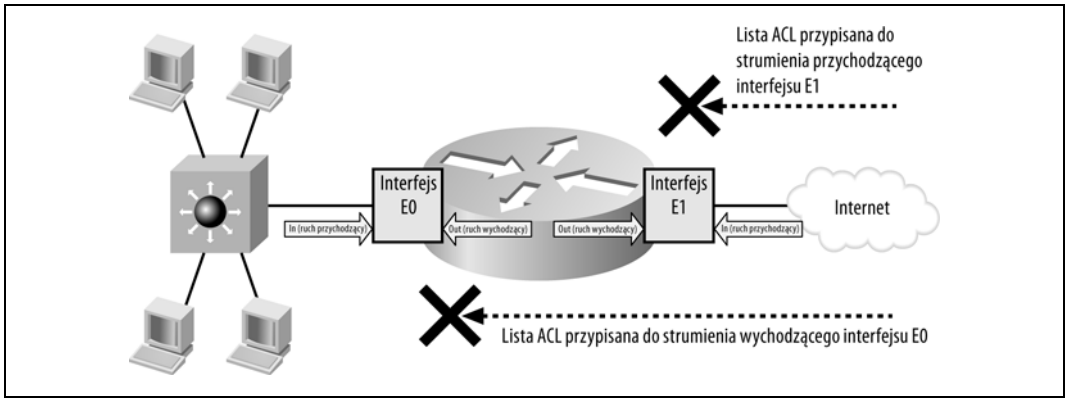
Rysunek 25.1. Punkty zastosowania listy dostępu

Zazwyczaj gdy filtrujemy ruch, chcemy uniemożliwić wprowadzenie go do sieci, a nawet do samego urządzenia. Powiązanie listy z kierunkiem ruchu przychodzącego na interfejsie zewnętrznym zapobiega dostarczeniu pakietów do urządzenia i oszczędza zasoby procesora. Jeśli pakiet zostanie wpuszczony do urządzenia, a następnie przekazany do innego interfejsu tylko po to, by tam został odrzucony przez filtr wyjściowy, marnuje tylko zasoby systemowe potrzebne do przełączania pakietów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Listy refleksyjne (opisane w dalszej części rozdziału) odnoszą się do obydwu kierunków przekazywania ruchu.

Na rysunku 25.2 pokazano niewielką sieć przyłączoną do internetu za pośrednictwem routera. Router filtruje ruch przychodzący z internetu, chroniąc jednostki wewnętrzne. Pakiety pochodzące z sieci zewnętrznych są dostarczane do interfejsu E1, a następnie zostają przekazane do interfejsu E0, z którego są wysyłane do sieci wewnętrznej. Powiązanie listy ACL ze strumieniem przychodzącym w ramach interfejsu E1 pozwala na usunięcie pakietów przed dalszym ich przetworzeniem. Gdyby lista została skojarzona ze strumieniem wychodzącym z interfejsu E0, router musiałby wykorzystać odpowiednie zasoby systemowe do przenoszenia pakietów tylko po to, by je później usunąć.



Rysunek 25.2. Użycie list dostępu

Usuwając listy dostępu, zachowaj szczególną ostrożność. Jeśli usuniesz listę, która jest powiązana z interfejsem routera, wszystkie dostarczane do niego pakiety zostaną odrzucone. Z tego powodu przed usunięciem listy zawsze usuń stosowne instrukcje `access-group`.

Nazywanie list dostępu

Systemy routerów Cisco umożliwiają definiowanie nazw list dostępu (a firma Cisco nawet zachęca administratorów do takiego działania) i zastępowanie nimi identyfikatorów numerycznych. Dzięki temu pliki konfiguracyjne są łatwiejsze do analizowania. Wadą wykorzystywania list nazywanych jest to, że nie można ich stosować w wielu przypadkach, w których użycie list numerowanych jest dopuszczalne. Na przykład są one wykorzystywane w konfigurowaniu map routingu, ale nie współdziałają z mapami interfejsów dzwoniących. Zapory sieciowe ASA i PIX obsługują jedynie listy nazywane. Nawet jeśli zostanie utworzona lista o identyfikatorze 10, będzie uznawana za listę nazywaną, a nie standardową (numerowaną). Przełączniki Nexus w ogóle nie obsługują list numerowanych.

Definiując taką listę, warto zastanowić się nad wyborem nazwy. Widziałem wiele konfiguracji zapór sieciowych PIX, w których wejściowa lista dostępu była określana jako „out" lub podobnie. Wyobraź sobie, że musisz ustalić przeznaczenie takiego polecenia:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
access-group out in interface outside
```

Osoby, które nie zajmują się na co dzień konfiguracją urządzeń ASA i PIX, mogą mieć trudności w interpretacji instrukcji. Gdyby lista była nazwana „RuchPrzychodzacy", polecenie byłoby znacznie łatwiejsze do zrozumienia:

```
access-group RuchPrzychodzacy in interface outside
```

Możliwość szybkiego przeanalizowania konfiguracji urządzenia pozwala zaoszczędzić czas w przypadku awarii sieci, a to może przekładać się wprost na zachowanie posady. Nazwy list dostępu zazwyczaj rozpoczynam wielką literą. Dzięki temu mogę szybciej wyszukać je w kodzie. Są to jednak tylko moje osobiste preferencje i nie wszyscy muszą się z nimi zgadzać — pracowałem z osobami, które unikały naciskania klawisza *Shift* podczas pisania na klawiaturze.

Przetwarzanie od góry do dołu

Listy dostępu są przetwarzane od góry do dołu, po jednym wierszu. Przetwarzanie ustaje w chwili odnalezienia pasującego wpisu. Warto o tym pamiętać, tworząc lub weryfikując listę dostępu. Częstym błędem administratorów jest bowiem dodawanie reguły przeznaczonej do wyodrębniania pakietów, które zostały już dopasowane na podstawie wcześniejszego, bardziej ogólnego wzorca. Przeanalizujmy przykład listy zapisanej w przełączniku 3750:

```
ip access-list extended GAD
 permit tcp any 10.10.10.0 0.0.0.255 eq www
 permit tcp any host 10.10.10.100 eq www
 permit tcp any host 10.10.10.100 eq domain
```

Drugi z wierszy `permit` nigdy nie zostanie wykorzystany, ponieważ wskazany w nim adres IP oraz protokół spełnią kryteria dopasowania wcześniejszego wiersza. Niemniej gdy pakiet nie będzie pasował do pierwszego wiersza, druga reguła i tak zostanie sprawdzona, co spowoduje marnowanie się czasu i zasobów procesora. Jest to częsty problem w sieciach korporacyjnych. Może być szczególnie uciążliwy w przypadku bardziej zaawansowanych zapór sieciowych, którymi administruje kilka osób. Bywa również bardzo trudny do wykrycia, ponieważ nie blokuje działania wymienionych protokołów. Często jest ujawniany dopiero w czasie audytu sieci.

Najczęstsze na początek

Listy dostępu powinny być budowane w taki sposób, aby reguły dopasowujące największą liczbę pakietów występowały na początku zestawienia. Pamiętaj, że przetwarzanie listy ACL trwa aż do napotkania pasującego wiersza. Jeśli pakiet spełnia kryteria dopasowania reguły, kolejne wpisy nie są interpretowane. W przypadku list o niewielu wpisach nie jest to szczególny problem. Jednak definicje ACL stosowane w routerach korporacyjnych często bywają bardzo rozbudowane. Zarządzałem kiedyś zaporą sieciową, w której zdefiniowanych było ponad 50 tys. reguł. Jak sądzisz, ile z tych wpisów było poprawnych?

Poniżej został zamieszczony przykład konfiguracji zapory sieciowej PIX. Budowanie tej listy sprowadzało się do tego, że członkowie zespołu dodawali poszczególne elementy w takiej kolejności, w jakiej o nich pomyśleli. Wbrew pozorom jest to często stosowany sposób postępowania. Zadanie polegało na sporządzeniu listy serwerów (`web1`, `lab`, `web2`) i zdefiniowaniu dozwolonych protokołów komunikacyjnych:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
access-list Inbound permit tcp any host web1.gad.net eq www
access-list Inbound permit tcp any host web1.gad.net eq ssh
access-list Inbound permit udp any host web1.gad.net eq domain
access-list Inbound permit tcp any host web1.gad.net eq smtp
access-list Inbound permit tcp any host web1.gad.net eq imap4
access-list Inbound permit tcp any host lab.gad.net eq telnet
access-list Inbound permit tcp any host lab.gad.net eq 8080
access-list Inbound permit udp any host web2.gad.net eq domain
access-list Inbound permit tcp any host web2.gad.net eq smtp
access-list Inbound permit tcp any host web2.gad.net eq imap4
```

Po kilku dniach funkcjonowania mechanizmu sprawdziliśmy, jakie są efekty jego działania. Wykorzystaliśmy do tego celu polecenie `show access-list`:

```
PIX#sho access-list
access-list cached ACL log flows: total 0, denied 0 (deny-flow-max 1024)
alert-interval 300
access-list Inbound; 15 elements
access-list Inbound permit tcp any host web1.gad.net eq www (hitcnt=42942)
access-list Inbound permit tcp any host web1.gad.net eq ssh (hitcnt=162)
access-list Inbound permit udp any host web1.gad.net eq domain (hitcnt=22600)
access-list Inbound permit tcp any host web1.gad.net eq smtp (hitcnt=4308)
access-list Inbound permit tcp any host web1.gad.net eq imap4 (hitcnt=100)
access-list Inbound permit tcp any host lab.gad.net eq telnet (hitcnt=0)
access-list Inbound permit tcp any host lab.gad.net eq 8080 (hitcnt=1)
access-list Inbound permit udp any host web2.gad.net eq domain (hitcnt=10029)
access-list Inbound permit tcp any host web2.gad.net eq smtp (hitcnt=2)
access-list Inbound permit tcp any host web2.gad.net eq imap4 (hitcnt=0)
```

Szczególnie istotnymi elementami zestawienia wynikowego są parametry `hitcnt`. Odpowiadają one liczbie udanych dopasowań. Ich wartości wskazują na to, że lista ACL nie została przygotowana optymalnie. Jej poprawienie nie stanowi problemu. Wystarczy posortować uzyskane wyniki względem parametru `hitcnt`, umieszczając na początku wpis o największej liczbie trafień. Wynik powinien być następujący:

```
access-list Inbound permit tcp any host web1.gad.net eq www (hitcnt=42942)
access-list Inbound permit udp any host web1.gad.net eq domain (hitcnt=22600)
access-list Inbound permit udp any host web2.gad.net eq domain (hitcnt=10029)
access-list Inbound permit tcp any host web1.gad.net eq smtp (hitcnt=4308)
access-list Inbound permit tcp any host web1.gad.net eq ssh (hitcnt=162)
access-list Inbound permit tcp any host web1.gad.net eq imap4 (hitcnt=100)
access-list Inbound permit tcp any host web2.gad.net eq smtp (hitcnt=2)
access-list Inbound permit tcp any host lab.gad.net eq 8080 (hitcnt=1)
access-list Inbound permit tcp any host lab.gad.net eq telnet (hitcnt=0)
access-list Inbound permit tcp any host web2.gad.net eq imap4 (hitcnt=0)
```

Jest to optymalne ustawienie tej niewątpliwie niezbyt rozbudowanej listy dostępu. Wpisy o największej liczbie dopasowań są teraz zapisane na początku zestawienia, a najrzadziej wykorzystywane znajdują się na końcu listy.

Weryfikuj założenia. Można by sądzić, że skoro dana zapora sieciowa chroni serwer pocztowy, to reguła odnosząca się do protokołu SMTP powinna występować na początku listy. Jednak analiza przedstawionego zestawienia wyraźnie wskazuje, że liczba odwołań do serwerów DNS znacznie przewyższa liczbę połączeń SMTP. Zawsze sprawdzaj, jaki jest rzeczywisty ruch sieciowy, i odpowiednio dostosowuj konfigurację.

Wadą opisanego rozwiązania jest zmniejszenie czytelności kodu. Analiza pierwotnej listy ACL okazuje się znacznie łatwiejsza. W zmodyfikowanej wersji wpis odnoszący się do serwera `web2` znajduje się pomiędzy regułami serwera `web1`. Można go więc łatwo pominąć, a tym samym wydłużyć czas poszukiwania ewentualnego problemu w funkcjonowaniu mechanizmu. Sam

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

musisz zdecydować, czy wprowadzone zmiany mają więcej wad, czy zalet. Budując krótkie listy ACL, można się zastanawiać nad zachowaniem ich czytelności, ale jeśli wykaz reguł zajmuje 17 stron, umieszczenie najczęściej wykorzystywanych elementów na początku znacznie przyspiesza działanie obciążonej zapory sieciowej.

Grupy reguł w urządzeniach ASA i PIX

Zapory sieciowe ASA i PIX umożliwiają definiowanie **grup** w ramach list dostępu. Mechanizm grupowania znacznie usprawnia proces tworzenia list ACL, ponieważ pozwala na uproszczenie rozbudowanych zestawień konfiguracyjnych. Zastosowanie grup umożliwia również wprowadzanie zmian w kilku listach ACL naraz — zmodyfikowanie ustawień grupy jest natychmiast odzwierciedlane we wszystkich listach, w których dana grupa została wymieniona. W przypadku bardziej rozbudowanych konfiguracji ACL stosowanie grup zapobiega powstawaniu błędów. Jeśli zapomnisz o pewnej regule, nie musisz dodawać jej w kilku miejscach. Wystarczy, że dopiszesz ją do grupy.

Przeanalizujmy zatem działanie wspomnianych grup. Oto pierwotna lista ACL:

```
object-group service CCIE-Rack tcp
  description [< Reverse telnet serwera terminali >]
  port-object range 2033 2050

access-list Inbound permit tcp any host gto eq www
access-list Inbound permit tcp any host gto eq ssh
access-list Inbound permit tcp any host meg eq ssh
access-list Inbound permit tcp any host meg eq www
access-list Inbound permit tcp any host lab eq telnet
access-list Inbound permit tcp any host lab object-group CCIE-Rack
access-list Inbound permit udp any host ASA-Outside eq 5060
access-list Inbound permit tcp any host lab eq 8080
access-list Inbound permit udp any host meg eq domain
access-list Inbound permit udp any host gto eq domain
access-list Inbound permit tcp any host gto eq smtp
access-list Inbound permit tcp any host meg eq smtp
access-list Inbound permit tcp any host gto eq imap4
access-list Inbound permit tcp any host meg eq imap4
access-list Inbound permit esp any any
access-list Inbound permit icmp any any unreachable
access-list Inbound permit icmp any any time-exceeded
access-list Inbound permit icmp any any echo-reply
```

Z analizy zestawienia wynika, że jest już zdefiniowana jedna grupa obiektów (`CCIE-Rack`). Jej użycie może się wydawać zbyteczne, ponieważ ten sam rezultat gwarantuje słowo kluczowe `range`:

```
access-list Inbound line 3 permit tcp any host lab range 2033 2050
```

Poza tym grupa obiektów i tak jest zastępowana instrukcjami o powyższej składni (co zostało opisane w dalszej części rozdziału). Część osób twierdzi, że jeśli zdefiniowanie grupy obiektów wymaga wprowadzenia większej liczby instrukcji niż liczy blok powstały po translacji, to nie należy takiej grupy wykorzystywać. Nie zgadzam się z tym poglądem. Cenię sobie możliwość definiowania opisu grupy oraz uzupełniania jej wpisów bez konieczności modyfikowania jakichkolwiek list dostępu.

Poniżej znajduje się listing grup, które utworzyłem na podstawie przedstawionego wcześniej zestawienia. Wszystkie instrukcje wspólne dla różnych serwerów zapisałem w grupie `Web ↪server-svcs`. Ponadto utworzyłem grupę `Webservers`, która obejmuje wszystkie serwery WWW,

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

grupę `Webserver-svcs-udp` uwzględniającą usługi bazujące na protokole UDP (na przykład DNS), a także grupę odnoszącą się do pakietów protokołu ICMP o nazwie `ICMP-Types`. Grupa `ICMP-Types` opisuje pakiety zwrotne z poleceń `ping` i `traceroute`. Nawiasy kwadratowe w polach opisu mogą się wydawać nieco dziwne, ale stosuję zasadę wyróżniania tego rodzaju treści.

```
object-group service CCIE-Rack tcp
  description [< Reverse telnet serwera terminali >]
  port-object range 2033 2050
object-group service Webserver-svcs tcp
  description [< Usługi TCP serwera WWW >]
  port-object eq www
  port-object eq ssh
  port-object eq domain
  port-object eq smtp
  port-object eq imap4
object-group service Webserver-svcs-udp udp
  description [< Usługi UDP serwera WWW >]
  port-object eq domain
object-group network Webservers
  description [< Serwery WWW >]
  network-object host gto
  network-object host meg
object-group icmp-type ICMP-Types
  description [< Dozwolone komunikaty ICMP >]
  icmp-object unreachable
  icmp-object time-exceeded
  icmp-object echo-reply
```

Gdy usługi i serwery są pogrupowane, można przystąpić do przepisywania list dostępu:

```
access-list Inbound permit udp any object-group Webservers object-group Webserversvcs-
udp
access-list Inbound permit tcp any object-group Webservers object-group Webserversvcs
access-list Inbound permit tcp any host lab eq telnet
access-list Inbound permit tcp any host lab object-group CCIE-Rack
access-list Inbound permit udp any host ASA-Outside eq 5060
access-list Inbound permit tcp any host lab eq 8080
access-list Inbound permit esp any any
access-list Inbound permit icmp any any object-group ICMP-Types
```

Liczba wierszy listy dostępu została zredukowana z 18 do 8. Oczywiście trzeba pamiętać, że są to wiersze widoczne dla użytkownika. W pamięci zapory sieciowej zostaną rozwinięte z powrotem do 18 wierszy.

Kolejność reguł może się okazać nieoptymalna, co z kolei bywa istotnym problemem w bardziej rozbudowanych konfiguracjach. Podobnie jak w wielu innych przypadkach, tak i w tym trzeba znaleźć pewien kompromis między wydajnością i elastycznością konfiguracji. Opracowując bardziej rozbudowane ustawienia firewalla PIX, warto skorzystać z list ACL Turbo (omówionych w kolejnym podrozdziale).

Tworzenie grup nie zawsze oznacza mniej pisania, wręcz przeciwnie. Mimo że sama lista dostępu została skrócona z 18 wierszy do 8, musiałem wpisać znacznie więcej instrukcji niż w poprzednim rozwiązaniu. Celem operacji jest jednak opracowanie zestawu reguł, które będą łatwiejsze do weryfikowania i korygowania. Każdy administrator musi więc sam ocenić, czy dodatkowy nakład pracy jest warty późniejszych korzyści.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Aby wyświetlić konfigurację w ostatecznej formie, wystarczy wpisać polecenie `show access-list`. Listing wynikowy zawiera zarówno wiersze `object-group`, jak i rzeczywiste wpisy (ACE), na które poszczególne grupy zostały przetłumaczone. Na prezentowanym listingu elementy `object-group` zostały odpowiednio wyróżnione.

```
GAD-PIX#sho access-list
access-list cached ACL log flows: total 0, denied 0 (deny-flow-max 1024)
            alert-interval 300
access-list Inbound; 20 elements
access-list Inbound line 1 permit udp any object-group Webservers object-group Webserver-svcs-udp
access-list Inbound line 1 permit udp any host gto eq domain (hitcnt=7265)
access-list Inbound line 1 permit udp any host meg eq domain (hitcnt=6943)
access-list Inbound line 2 permit tcp any object-group Webservers object-group Webserver-svcs
access-list Inbound line 2 permit tcp any host gto eq www (hitcnt=21335)
access-list Inbound line 2 permit tcp any host gto eq ssh (hitcnt=4428)
access-list Inbound line 2 permit tcp any host gto eq domain (hitcnt=0)
access-list Inbound line 2 permit tcp any host gto eq smtp (hitcnt=1901)
access-list Inbound line 2 permit tcp any host gto eq imap4 (hitcnt=116)
access-list Inbound line 2 permit tcp any host meg eq www (hitcnt=23)
access-list Inbound line 2 permit tcp any host meg eq ssh (hitcnt=15)
access-list Inbound line 2 permit tcp any host meg eq domain (hitcnt=0)
access-list Inbound line 2 permit tcp any host meg eq smtp (hitcnt=1)
access-list Inbound line 2 permit tcp any host meg eq imap4 (hitcnt=0)
access-list Inbound line 3 permit tcp any host lab eq telnet (hitcnt=0)
access-list Inbound line 4 permit tcp any host lab object-group CCIE-Rack
access-list Inbound line 4 permit tcp any host lab range 2033 2050 (hitcnt=0)
access-list Inbound line 5 permit udp any host ASA-Outside eq 5060 (hitcnt=0)
access-list Inbound line 6 permit tcp any host lab eq 8080 (hitcnt=0)
access-list Inbound line 7 permit esp any any (hitcnt=26256)
access-list Inbound line 8 permit icmp any any object-group ICMP-Types
access-list Inbound line 8 permit icmp any any unreachable (hitcnt=359)
access-list Inbound line 8 permit icmp any any time-exceeded (hitcnt=14)
access-list Inbound line 8 permit icmp any any echo-reply (hitcnt=822)
```

Usuwanie list ACL

W starych dobrych czasach do usunięcia listy o numerze 101 wystarczyło polecenie `no access-list 101`. W urządzeniach ASA instrukcja tego typu nie wykonuje żadnego zadania. Konieczne jest bowiem określenie wiersza przeznaczonego do usunięcia:

```
ASA-5540/GAD(config)#no access-list 101
ERROR: % Incomplete command
```

W praktyce okazuje się to użytecznym rozwiązaniem, ponieważ zapobiega przypadkowemu usunięciu całej listy, gdy użytkownik zamierzał usunąć tylko jeden wiersz (nie twierdzę, że mi się to kiedyś przydarzyło). Za usunięcie całej listy w systemie ASA odpowiada polecenie `clear configure access-list` *`nazwa-listy-dostępu`*:

```
ASA-5540/GAD(config)# clear configure access-list 101
```

Listy Turbo ACL

Klasyczne listy ACL są interpretowane za każdym razem, gdy występuje konieczność ich zastosowania. Taki sposób działania wiąże się ze znacznym obciążeniem procesora (szczególnie w rozwiązaniach uwzględniających rozbudowaną konfigurację ACL).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeden ze sposobów zwiększenia wydajności przetwarzania długich list ACL polega na ich wstępnym skompilowaniu. Skompilowana lista ACL jest nazywana listą **Turbo ACL**. Sam proces kompilacji polega na przekształceniu reguł ACL w kod maszynowy, który nie wymaga interpretacji przed zastosowaniem. Dzięki temu można uzyskać istotny wzrost wydajności przetwarzania pakietów.

Mechanizm Turbo ACL jest dostępny w zaporach sieciowych PIX i niektórych routerach. W urządzeniach PIX kompilacja wszystkich list dostępu następuje po wprowadzeniu polecenia `access-list compiled`. W przypadku routerów obsługa list Turbo została zaimplementowana jedynie w modelach 7100, 7200, 7500 oraz 12000 (z systemem 12.0(6)S lub późniejszym). Włączenie opcji kompilacji również wymaga wprowadzenia instrukcji `access-list compiled`. Najnowsze wersje urządzeń ASA są sprzedawane z domyślnie włączoną opcją list Turbo ACL (której nie można wyłączyć).

Jeśli w zaporze PIX funkcjonują listy Turbo ACL, wynik wykonania polecenia `show access-list` jest zmieniany w taki sposób, aby uwzględniał informacje o kompilacji reguł oraz o rozmiarze pamięci zajętej przez poszczególne listy. Zestawienie to nie jest generowane w najnowszych wersjach urządzeń ASA:

```
PIX(config)#access-list comp
PIX(config)#show access-list
TurboACL statistics:
ACL                     State       Memory(KB)
----------------------- ----------- ----------
Inbound                 Operational 2          Shared memory usage: 2056 KB
access-list compiled
access-list cached ACL log flows: total 0, denied 0 (deny-flow-max 1024)
            alert-interval 300
access-list Inbound turbo-configured; 20 elements
access-list Inbound line 1 permit udp any object-group Webservers object-group Webserver-svcs-udp
access-list Inbound line 1 permit udp any host gto eq domain (hitcnt=7611)
access-list Inbound line 1 permit udp any host meg eq domain (hitcnt=7244)
access-list Inbound line 2 permit tcp any object-group Webservers object-group Webserver-svcs
access-list Inbound line 2 permit tcp any host gto eq www (hitcnt=22578)
access-list Inbound line 2 permit tcp any host gto eq ssh (hitcnt=4430)
[fragment usunięty]
```

Dopuszczenie wychodzącego ruchu ping i traceroute

Jednym z najbardziej irytujących aspektów pracy zapór sieciowych jest niemożność wykorzystania polecenia `ping` lub `traceroute` po zainstalowaniu reguł filtrowania ruchu. Przeświadczenie o tym, że protokół ICMP stanowi zagrożenie, jest w pełni uzasadnione. Niemniej zrozumienie zasad działania mechanizmu ICMP pozwala na wyróżnienie jedynie określonych typów pakietów i zapewnienie poprawnego działania narzędzi `ping` i `traceroute`.

Zakładając, że zezwalasz na przekazywanie całego ruchu wychodzącego, możesz zastosować takie reguły filtrowania pakietów przychodzących, które zapewnią dostarczanie informacji zwrotnych wygenerowanych w wyniku zastosowania poleceń `ping` i `traceroute`. Dzięki temu wszystkie testy wykonywane z wnętrza sieci będą realizowane poprawnie. Jednocześnie wszelkie próby pingowania urządzeń spoza sieci chronionej zostaną odrzucone. Aby wdrożyć opisaną politykę, trzeba umożliwić dostarczanie z zewnątrz pakietów ICMP o następujących typach:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ICMP unreachable (niedostępność)
W protokole ICMP zdefiniowano wiele komunikatów o niedostępności (ang. *unreachable*) urządzeń lub sieci. Generalnie zezwalanie na ich dostarczenie nie stanowi zagrożenia, ponieważ są to pakiety odpowiedzi.

Time exceeded (przekroczenie czasu)
Komunikaty przekroczenia czasu są generowane przez każdy router na trasie do jednostki docelowej w wyniku wykonania polecenia `traceroute`.

Echo reply (odpowiedź na żądanie echa)
Ten rodzaj komunikatu jest odpowiedzią na pakiet żądania wygenerowany przez polecenie `ping`.

Wspomniane filtry są zazwyczaj dołączane na końcu zainstalowanej wcześniej wejściowej listy dostępu. W większości przypadków powinny być definiowane jako ostatnie elementy listy ACL, chyba że z sieci wewnętrznej generowany jest ruch ICMP o znacznym natężeniu. Poniżej znajduje się kilka przykładów implementacji wspomnianych filtrów w routerach i zaporach sieciowych PIX firmy Cisco.

IOS:

```
access-list 101 remark [< Dozwolone pakiety ping i traceroute >]
access-list 101 permit icmp any any unreachable
access-list 101 permit icmp any any time-exceeded
access-list 101 permit icmp any any echo-reply
!
interface Ethernet1
ip access-group 101 in
```

Zastosowanie nazywanych list dostępu:

```
ip access-list extended Allow-ICMP
 remark [< Dozwolone pakiety ping i traceroute >]
 permit icmp any any unreachable
 permit icmp any any time-exceeded
 permit icmp any any echo-reply
!
interface GigabitEthernet1/0/1
 ip access-group Allow-ICMP in
```

Zapory sieciowe:

```
object-group icmp-type ICMP-Types
 description [< Typy dozwolonych pakietów ICMP >]
 icmp-object unreachable
 icmp-object time-exceeded
 icmp-object echo-reply
!
access-list Inbound permit icmp any any object-group ICMP-Types
!
access-group Inbound in interface outside
```

Akceptowanie pakietów mechanizmu wykrywania MTU

Mechanizm wykrywania maksymalnego rozmiaru jednostki transmisyjnej (MTU — ang. *Maximum Transmission Unit*) służy do wykrywania dopuszczalnej wartości MTU na trasie do sieci zdalnej. Aby mógł działać poprawnie, filtry muszą zezwalać na dostarczanie dwóch dodatkowych

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

rodzajów pakietów ICMP — **source-quench** (tłumienie źródła) oraz **parameter-problem** (problem z parametrem). Aby włączyć je w routerze lub zaporze sieciowej PIX, wystarczy wpisać następujące polecenia:

```
access-list 101 remark [<Dozwolone pakiety ping i traceroute >]
access-list 101 permit icmp any any unreachable
access-list 101 permit icmp any any time-exceeded
access-list 101 permit icmp any any echo-reply
access-list 101 permit icmp any any parameter-problem
access-list 101 permit icmp any any source-quench
!
interface Ethernet1
ip access-group 101 in
```

Nazywane listy dostępu:

```
ip access-list extended Allow-ICMP
 remark [<Dozwolone pakiety ping i traceroute >]
 permit icmp any any unreachable
 permit icmp any any time-exceeded
 permit icmp any any echo-reply
 permit icmp any any parameter-problem
 permit icmp any any source-quench
!
interface GigabitEthernet1/0/1
 ip access-group Allow-ICMP in
```

Zapory sieciowe:

```
object-group icmp-type ICMP-Types
 description [<Dozwolone pakiety ping i traceroute >]
 icmp-object unreachable
 icmp-object time-exceeded
 icmp-object echo-reply
 icmp-object source-quench
 icmp-object parameter-problem
!
access-list Inbound permit icmp any any object-group ICMP-Types
!
access-group Inbound in interface outside
```

Listy ACL w przełącznikach wielowarstwowych

Wielowarstwowe przełączniki z założenia realizują pewne funkcje, które nie są dostępne w przełącznikach warstwy drugiej oraz w routerach.

Przełącznik 3750 obsługuje zarówno listy ACL protokołu IP, jak i ethernetowe listy ACL (operujące adresami MAC). Oto kilka sposobów ich wykorzystania:

Listy ACL portu

Listy ACL portu (ang. *port ACL*) są powiązane z interfejsami warstwy drugiej przełącznika. Nie można ich przypisywać do portów interfejsów EtherChannel, SVI lub innych interfejsów wirtualnych. Dozwolone jest natomiast stosowanie ich w portach trunkowych, w których filtrują wszystkie VLAN-y. Do portów przypisywane są standardowe listy IP, rozszerzone listy IP oraz listy ACL odnoszące się do adresów MAC. Filtrowanie zawsze dotyczy strumieni danych **przychodzących**.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Listy ACL routera

Listy ACL routera (ang. *router ACL*) są przypisywane do interfejsów warstwy trzeciej przełącznika, czyli do interfejsów SVI, fizycznych interfejsów warstwy trzeciej (skonfigurowanych za pomocą instrukcji `no switchport`) oraz do interfejsów EtherChannel warstwy trzeciej. Jako listy ACL routera można wykorzystywać standardowe i rozszerzone listy IP. Nie wolno jednak używać list MAC. Listy ACL routera są stosowane do filtrowania danych **przychodzących** i **wychodzących**.

Mapy VLAN-ów

Mapy VLAN-ów (ang. *VLAN map*) są zbliżone w działaniu do map routingu. Są one powiązane z poszczególnymi sieciami VLAN, a ich konfiguracja umożliwia przekazywanie lub odrzucanie pakietów w zależności od wyniku zdefiniowanego testu. Mapy VLAN-ów regulują cały ruch wprowadzany do sieci VLAN, wychodzący z tej sieci oraz przenoszony w jej ramach. Nie mają określonych kierunków transmisji.

Konfiguracja list ACL portu

Listy ACL portu są zbiorami reguł filtrowania skojarzonymi z określonym interfejsem fizycznym. Można ich używać do blokowania dostępu określonej stacji do sieci VLAN lub do innej jednostki w ramach danej sieci VLAN. Znajdują również zastosowanie w ograniczaniu ruchu wychodzącego poza sieć VLAN.

Wyobraź sobie, że w sieci VLAN o identyfikatorze 100 pracuje wiele stacji, w tym komputer A. Komputer A nie ma komunikować się z żadną inną stacją w tej sieci VLAN. Powinien jednak mieć możliwość wymieniania danych z bramą domyślną (a tym samym z jednostkami spoza własnej sieci). Załóżmy, że adres IP komputera A to `192.168.1.155/24`, adres IP bramy domyślnej to `192.168.1.1/24` oraz że komputer A jest przyłączony do portu G0/20 przełącznika.

Pierwsza część zadania związanego z ograniczeniem zakresu komunikacji jednostce A polega na utworzeniu listy ACL, która zapewni blokowanie ruchu do innych stacji w sieci oraz zezwoli na wymianę danych z komputerami spoza tej sieci. Wykorzystam do tego celu nazywaną listę ACL (ponieważ te listy są obsługiwane w przełącznikach):

```
ip access-list extended Permit-Bob
 permit ip any host 192.168.1.1
 deny ip any 192.168.1.0 0.0.0.255
 permit ip any any
```

Po utworzeniu listę ACL należy przypisać do interfejsu fizycznego:

```
IOS-3750(config)#int g1/0/2
IOS-3750 (config-if)#switchport
IOS-3750 (config-if)#ip access-group Permit-Bob in
```

Zwróć uwagę na fakt, że mimo iż operacja odnosi się do portu warstwy drugiej, można w niej wykorzystać listę IP (listę ACL warstwy trzeciej). Z tego też względu przedstawiona lista jest nazywana listą ACL portu.

Jako listę ACL portu można również wykorzystać listę MAC. Oto przykład nieskomplikowanego rozwiązania, które uniemożliwia przekazywanie pakietów AppleTalk, zezwalając jednocześnie na każdą inną formę transmisji:

```
mac access-list extended No-Appletalk
 deny any any appletalk
 permit any any
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przypisanie powyższej listy do interfejsu czyni z niej listę ACL portu:

```
IOS-3750 (config)#int g1/0/2
IOS-3750 (config-if)# mac access-group No-Appletalk in
```

Listy ACL MAC mogą współdziałać z listami IP w ramach jednego interfejsu. Z kolejnego przykładu wynika, że na jednym interfejsie działa zarówno lista MAC, jak i lista IP:

```
IOS-3750#sho run int g1/0/2
interface GigabitEthernet1/0/2
 ip access-group Permit-Bob in
 mac access-group No-Appletalk in
end
```

Konfiguracja list ACL routera

Większość osób, mówiąc o listach ACL, ma na myśli listy ACL routera, które są przypisywane do interfejsów warstwy trzeciej. Starsze modele routerów zawierają jedynie interfejsy warstwy trzeciej, więc wszystkie zdefiniowane w nich listy ACL mają charakter list ACL routera.

Gdybyśmy zmodyfikowali nieco poprzednie rozwiązania, przekształcając port warstwy drugiej w interfejs warstwy trzeciej, zdefiniowana wcześniej lista ACL stałaby się listą ACL routera:

```
IOS-3750(config-if)#int g1/0/2
IOS-3750(config-if)#no switchport
IOS-3750(config-if)#ip access-group Permit-Bob in
```

W charakterze list routera nie można wykorzystywać list MAC.

Listy ACL routera można przypisywać również do strumieni danych wychodzących (choć nie jestem zwolennikiem tego rozwiązania):

```
IOS-3750(config-if)#ip access-group Permit-Bob out
```

Pamiętaj, że powiązanie listy ACL z interfejsem warstwy trzeciej powoduje, że dana lista staje się listą ACL routera. Trzeba o tym pamiętać, jeśli przypisuje się do interfejsu jednocześnie listę ACL portu i listę ACL routera:

```
2w3d: %FM-3-CONFLICT: Input router ACL Permit-Bob conflicts with port ACLs
```
//Wejściowa lista ACL routera Permit-Bob koliduje z listą ACL portu

Powyższy komunikat informuje o tym, że zdefiniowano listę ACL portu i listę ACL routera, które odnoszą się do tych samych adresów IP (zakresy adresów się pokrywają). Ostrzeżenie zostało wygenerowane, ponieważ obydwie listy będą aktywne, a lista portu będzie w tym przypadku ważniejsza.

Zdefiniowanie jednocześnie listy ACL portu i listy routera może być przyczyną wielu nieporozumień, jeśli administrator nie uświadomi sobie, że lista portu jest aktywna.

Konfigurowanie map VLAN-ów

Mapy VLAN-ów pozwalają na łączenie list dostępu w ciekawy sposób. Służą one bowiem do filtrowania całego ruchu **wewnątrz** sieci VLAN.

Listy ACL portu filtrują jedynie pakiety przychodzące do pojedynczego interfejsu. Listy ACL routera filtrują tylko pakiety wchodzące do interfejsu warstwy trzeciej lub wychodzące z takiego interfejsu. Mapy VLAN-ów pozwalają natomiast na filtrowanie wszystkich pakietów z danej

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

sieci VLAN niezależnie od tego, przez jaki port są przekazywane. Na przykład jeśli utworzysz filtr uniemożliwiający wymianę danych między jednostkami o adresach MAC 1111.1111.1111 i 2222.2222.2222, przypiszesz go do interfejsu, a następnie przełączysz urządzenie do innego portu, ograniczenie przestanie obowiązywać. W przypadku mapy VLAN-ów filtr będzie działał niezależnie od tego, który interfejs będzie uczestniczył w wymianie danych (zakładając oczywiście, że należy do danej sieci VLAN).

W pierwszym z przedstawionych przykładów zdefiniujemy filtr, który uniemożliwi korzystanie z protokołu AppleTalk w sieci VLAN o identyfikatorze 100. Oto stosowna lista dostępu MAC:

```
IOS-3750(config)#mac access-list extended No-Appletalk
IOS-3750(config-ext-macl)#permit any any appletalk
```

Z treści listy wynika, że ruch AppleTalk jest dozwolony (ang. *permit*), mimo że naszym zadaniem jest zablokowanie go. Taki zapis wynika ze sposobu działania map VLAN-ów, o czym się przekonasz, czytając dalszą część rozdziału.

Aby zablokować ruch w protokole AppleTalk w ramach danej sieci VLAN, musimy opracować mapę VLAN-u. Mapy VLAN-ów składają się z klauzul analogicznych do stosowanych w mapach routingu. Poszczególne wpisy są numerowane, ale w przeciwieństwie do map routingu każda klauzula zawiera definicję akcji.

Pierwszy krok to zdefiniowanie mapy za pomocą instrukcji `vlan access-map`. Sama mapa będzie się składała z dwóch klauzul. Zadaniem pierwszej (o numerze 10) będzie dopasowanie pakietów do listy MAC `No-Appletalk` i odrzucenie pakietów spełniających kryterium dopasowania. Z tego względu w treści listy dostępu konieczne było zapisanie dyrektywy `permit appletalk`, a nie `deny appletalk`. Wpis `permit` zapewnia dopasowanie pakietów protokołu AppleTalk. Samo odrzucenie pakietów należy do zadań instrukcji `action` w mapie VLAN-u:

```
vlan access-map Limit-V100 10
 action drop
 match mac address No-Appletalk
```

Kolejna czynność to zdefiniowanie klauzuli, która zapewni przekazywanie wszystkich pozostałych pakietów. Brak instrukcji `match` w klauzuli oznacza, że wszystkie pakiety są uznawane za dopasowane:

```
vlan access-map Limit-V100 20
 action forward
```

Oto cała mapa VLAN-u:

```
vlan access-map Limit-V100 10
 action drop
 match mac address No-Appletalk
vlan access-map Limit-V100 20
 action forward
```

Po przygotowaniu mapy trzeba ją powiązać z odpowiednią siecią VLAN. Służy do tego polecenie konfiguracji ogólnej `vlan filter`:

```
IOS-3750(config)#vlan filter Limit-V100 vlan-list 100
```

Aby użyć tej samej mapy w kilku sieciach VLAN, wystarczy wymienić ich numery na końcu wiersza.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Prawdopodobnie po zapoznaniu się z zaproponowanym rozwiązaniem wiele osób zastanawia się, czy tego samego zadania nie można by zrealizować za pomocą standardowej listy dostępu (takiej, jaka została przedstawiona poniżej) przypisanej do określonych interfejsów.

```
mac access-list extended No-Appletalk
deny any any appletalk
permit any any
```

Odpowiedź jest twierdząca. Trzeba by jednak najpierw ustalić, przez które interfejsy będą przekazywane pakiety AppleTalk, a tym samym, do których interfejsów należy przypisać listę. Oczywiście można powiązać ją ze wszystkimi interfejsami w ramach sieci VLAN, ale wówczas trzeba pamiętać o dodaniu stosownego powiązania za każdym razem, gdy kolejny port będzie dołączany do tej samej sieci VLAN. Skojarzenie listy dostępu z samą siecią VLAN zapewnia odrzucanie wszystkich pakietów AppleTalk niezależnie od tego, jakie są ich porty źródłowe i docelowe.

Aby sprawdzić, które mapy VLAN-ów są w danej chwili wykorzystywane, należy wprowadzić polecenie `show vlan filter`:

```
IOS-3750#sho vlan filter
VLAN Map Limit-VlOO is filtering VLANs:
  100
```

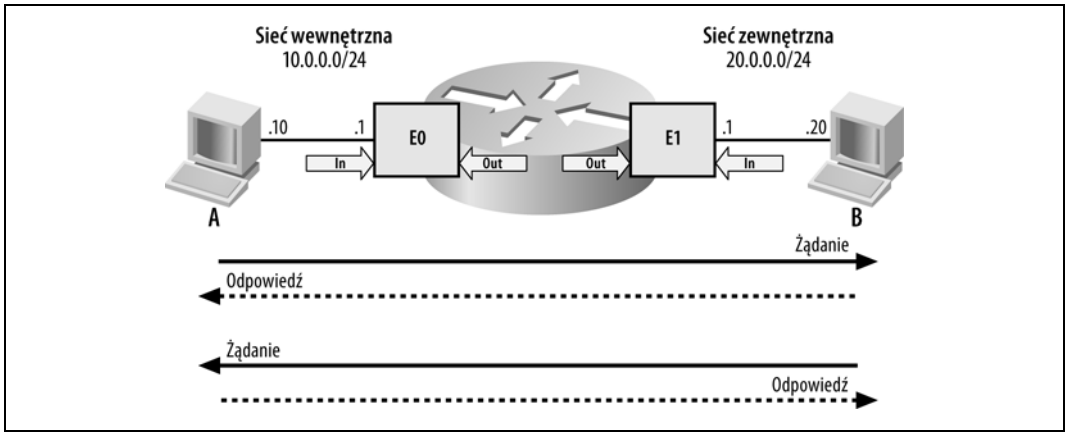
Refleksyjne listy dostępu

Refleksyjne listy dostępu są dynamicznymi filtrami, które decydują o przekazywaniu ruchu na podstawie informacji o ruchu przenoszonym w przeciwnym kierunku. Nadają się na przykład do zdefiniowania reguły typu „pozwalaj na dostarczanie pakietów telnet, jeśli wcześniej zainicjowałem wychodzący ruch usługi telnet". Gdy pierwszy raz tłumaczę zasadę działania tego mechanizmu technikom, często otrzymuję pytanie „a to nie działa zawsze w ten sposób?". Wiele osób myli reguły filtrowania z działaniem mechanizmu translacji adresów i portów (PAT — ang. *Port Address Translation*). Rozwiązanie PAT zapewnia wprowadzanie do sieci strumieni przychodzących tylko wówczas, gdy dostarczane pakiety są odpowiedzią na ruch wychodzący z danej sieci. Wynika to z natury mechanizmu PAT (przekazywanie danych jest możliwe jedynie po uprzednim zdefiniowaniu parametrów translacji). Refleksyjne listy dostępu są znacznie bardziej użyteczne i mogą być stosowane z wielu powodów.

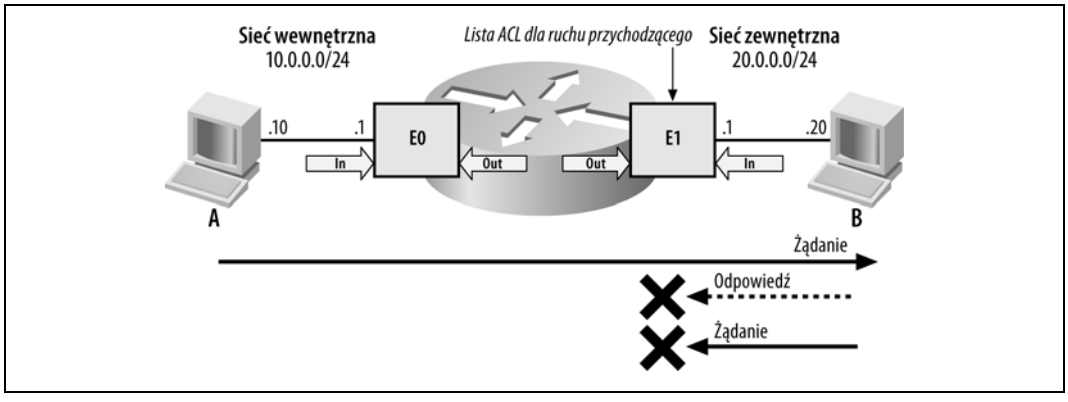
Przy braku mechanizmu PAT filtr ogranicza poszczególne formy ruchu niezależnie od innych rodzajów ruchu. Przeanalizujmy działanie sieci przedstawionej na rysunku 25.3. Składa się ona z dwóch stacji (A i B) połączonych za pomocą routera. W routerze nie zdefiniowano żadnych list dostępu. Odpowiedzi na żądania przesyłane ze stacji A do stacji B są poprawnie przenoszone. Nie ma również żadnych ograniczeń w dostarczaniu odpowiedzi na żądania jednostki B kierowane do komputera A.

Załóżmy, że stacja A powinna mieć możliwość ustanowienia połączenia telnet z jednostką B, ale nie chcemy, aby komputer B mógł nawiązywać tego typu połączenie z komputerem A. Gdybyśmy powiązali standardową wejściową listę dostępu z interfejsem E1 routera, moglibyśmy zezwolić jednostce A na kontakt ze stacją B i jednocześnie zabronić komunikacji stacji B z komputerem A. Niestety spowodowałoby to również odrzucenie odpowiedzi stacji B kierowanych do komputera A. Sytuację tę ilustruje rysunek 25.4.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 25.3. Prosta sieć bez list ACL



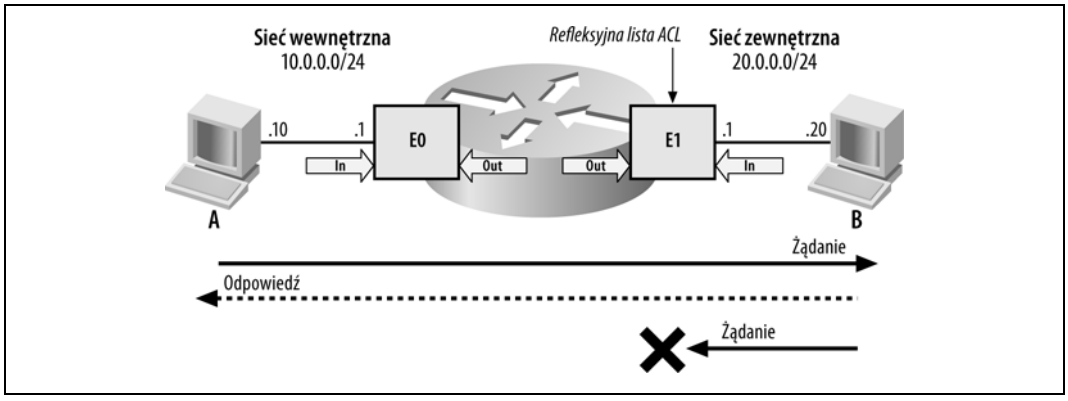
Rysunek 25.4. Standardowa lista dostępu powiązana z interfejsem E1

Takie rozwiązanie okazuje się zbyt restrykcyjne. Chroni komputer A przed próbami nawiązania połączeń ze strony stacji B, ale jednocześnie wyklucza wszelkie użyteczne formy komunikacji między tymi jednostkami. Praca routera powinna w tym zakresie przypominać działanie zapory sieciowej — router musi odrzucać żądania generowane przez komputer B, ale jednocześnie akceptować odpowiedzi jednostki B na żądania wysyłane przez stację A. Rozwiązaniem problemu jest zastosowanie list refleksyjnych.

Refleksyjne listy dostępu tworzą dynamicznie listy ACL, które umożliwiają przekazywanie odpowiedzi na żądania. W analizowanym przypadku chcemy zezwolić na dostarczanie pakietów z komputera B, ale pod warunkiem, że wcześniej zarejestrowano ruch ze stacji A. Komunikacja inicjowana po stronie jednostki B powinna zostać zablokowana. Zasada ta została przedstawiona graficznie na rysunku 25.5.

Refleksyjne listy dostępu generują tymczasowe instrukcje `permit`, które są odwrotną formą pierwotnych instrukcji. Na przykład jeśli zezwolimy na ruch wychodzący w protokole telnet, system utworzy tymczasową regułę `permit` dla przychodzącego strumienia telnetu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 25.5. Refleksyjna lista dostępu powiązana z interfejsem E1

Refleksyjne listy dostępu są bardzo użyteczne, ale mają wiele ograniczeń:

- Tymczasowy wpis jest zawsze instrukcją `permit` (nigdy `deny`).
- Tymczasowy wpis zawsze odnosi się do tego samego protokołu, który został zarejestrowany w strumieniu pierwotnym (TCP, UDP itd.).
- Źródłowy i docelowy adres IP we wpisie tymczasowym są zamienione miejscami względem wartości zarejestrowanych w strumieniu pierwotnym.
- Wpis tymczasowy uwzględnia takie same numery portów, jakie występują w ruchu pierwotnym, ale wartości portu źródłowego i docelowego są ze sobą zamienione (reguła ta odnosi się również do typów komunikatów ICMP, które nie wykorzystują numerów portów).
- Wpis tymczasowy jest usuwany po zarejestrowaniu ostatniego pakietu strumienia (zazwyczaj FIN lub RST).
- Wpis tymczasowy jest usuwany, jeśli w ustalonym czasie (domyślnie jest to pięć sekund) nie zostanie zaobserwowany żaden dodatkowy ruch.

Nie można utworzyć refleksyjnych list dostępu, które pozwolą na dostarczanie pakietów określonego protokołu, jeśli wykryty został inny protokół. Na przykład nie można skonfigurować routera w taki sposób, aby akceptował przychodzący ruch HTTP w reakcji na wychodzący ruch telnet. Aby mechanizm zadziałał w odniesieniu do wejściowego strumienia HTTP, lista musi zawierać regułę opisującą wyjściowy strumień HTTP.

Z uwagi na to, że numery portów tymczasowych wpisów są zawsze zamienionymi wartościami portów strumienia pierwotnego, opisywane rozwiązanie nie znajduje zastosowania w przekazywaniu danych takich protokołów jak RPC, w których numery portów źródłowych zmieniają się dynamicznie. Refleksyjne listy dostępu nie nadają się również do obsługi protokołów, które tworzą nowe połączenia, takich jak FTP.

Protokół FTP może być obsługiwany przez refleksyjne listy dostępu, jeśli działa w trybie **pasywnym**.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Konfiguracja refleksyjnych list dostępu

Definiowanie refleksyjnych list dostępu jest nieco bardziej skomplikowane niż tworzenie standardowych zestawień tego typu, ponieważ wymaga zagnieżdżenia jednej listy ACL w drugiej. Rozważmy operację testowania dwóch rodzajów ruchu — żądania oraz odpowiedzi na żądanie. Na potrzeby każdego testu musi zostać utworzona osobna lista ACL. Lista odnosząca się do odpowiedzi jest powoływana automatycznie po stwierdzeniu dopasowania strumienia do wpisu listy pierwotnej.

Firma Cisco nazywa sposób konfiguracji list refleksyjnych **zagnieżdżaniem**, choć sama konfiguracja nie przypomina w żaden sposób zagnieżdżania znanego większości programistów.

Kontynuując rozważania z poprzedniego przykładu, przygotujmy refleksyjną listę dostępu dla protokołu telnet. Chcemy, aby jednostka A mogła nawiązać połączenie telnet z komputerem B, ale nie pozwalamy na inne formy wymiany danych. Zaproponowane rozwiązanie jest nadmiernie restrykcyjne i uniemożliwia komunikację z większością aplikacji internetowych, ale doskonale nadaje się do zilustrowania zasad działania refleksyjnych list dostępu.

Aby skonfigurować mechanizm, musimy utworzyć jedną listę ACL opisującą ruch wychodzący oraz jedną powiązaną z ruchem przychodzącym.

Na początek przygotujmy listę o nazwie `TelnetOut`.

```
ip access-list extended TelnetOut
 permit tcp host 10.0.0.10 host 20.0.0.20 eq telnet reflect GAD
 deny ip any any
```

Listy refleksyjne można tworzyć jedynie na podstawie nazywanych list dostępu.

Treść reguł jest wyjątkowo prosta. Jedyna nowość polega na dodaniu słów `reflect GAD` na końcu wiersza `permit`. W ten sposób wyznaczona została nazwa tymczasowej listy dostępu, która zostanie utworzona przez router w reakcji na dopasowanie pakietu do klauzuli `permit`. Wpis `deny ip any any` nie jest potrzebny, ponieważ wszystkie listy są automatycznie wyposażane w taką regułę. Dodaję ją jawnie po to, aby zwiększyć czytelność kodu i mieć później możliwość sprawdzenia liczników odrzucanego ruchu.

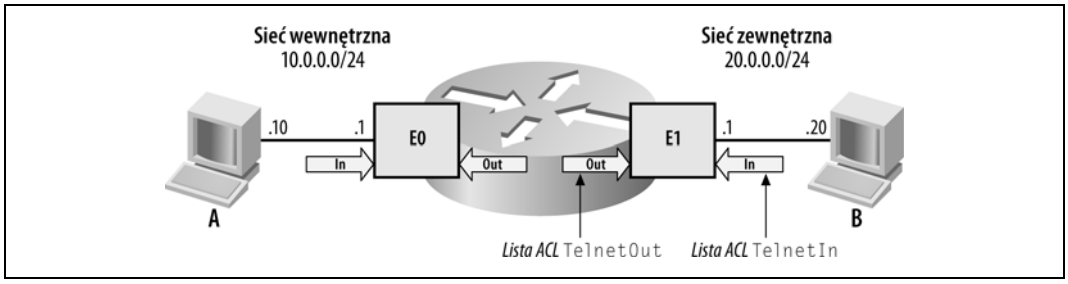
Następnie musimy utworzyć listę, którą nazwiemy `TelnetIn`:

```
ip access-list extended TelnetIn
 evaluate GAD
 deny ip any any
```

W tym zestawieniu nie występują żadne instrukcje `permit`. Jest natomiast dyrektywa `evalute GAD`. Wiersz ten odnosi się do klauzuli `reflect` z listy `TelnetOut`. Ciąg tekstowy `GAD` jest nazwą listy tworzonej w sposób automatyczny przez router.

Aby skonfigurowany mechanizm zadziałał, trzeba go skojarzyć z interfejsami routera. Lista `TelnetOut` powinna zostać przypisana do strumienia **wyjściowego** interfejsu E1, a lista `TelnetIn` do strumienia **wejściowego** tego samego interfejsu (zgodnie z rysunkiem 25.6).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 25.6. Zastosowanie refleksyjnych list dostępu

Powiązanie refleksyjnej listy dostępu z interfejsem należy do zadań instrukcji `access-group` (wykonywanej w trybie konfiguracji interfejsu):

```
interface Ethernet1
 ip access-group TelnetIn in
 ip access-group TelnetOut out
```

Cała konfiguracja opisywanego mechanizmu powinna mieć następującą postać:

```
interface Ethernet0
 ip address 10.0.0.1 255.255.255.0
!
interface Ethernet1
 ip address 20.0.0.1 255.255.255.0
 ip access-group TelnetIn in
 ip access-group TelnetOut out
!
ip access-list extended TelnetIn
 evaluate GAD
 deny ip any any
ip access-list extended TelnetOut
 permit tcp host 10.0.0.10 host 20.0.0.20 eq telnet reflect GAD
 deny ip any any
```

Wykonanie polecenia `show access-list` powoduje wyświetlenie obydwu list o dokładnie takiej treści, jaka została zdefiniowana:

```
Router#sho access-list
Reflexive IP access list GAD
Extended IP access list TelnetIn
    evaluate GAD
    deny ip any any
Extended IP access list TelnetOut
    permit tcp host 10.0.0.10 host 20.0.0.20 eq telnet reflect GAD
    deny ip any any (155 matches)
```

Z powyższego listingu wynika, że cały ruch wychodzący niezwiązany z telnetem jest blokowany. Nie ma również żadnych wpisów, które zezwalałyby na przetwarzanie strumieni przychodzących. To się jednak zmieni, gdy zostanie wyzwolona refleksyjna lista dostępu.

Po zainicjowaniu połączenia telnet z jednostki A do jednostki B wynik użycia polecenia ulega zmianie. Pojawia się dodatkowa lista dostępu o nazwie `GAD`:

```
Router#sho access-list
Reflexive IP access list GAD
    permit tcp host 20.0.0.20 eq telnet host 10.0.0.10 eq 11002 (12 matches)
Extended IP access list TelnetIn
    evaluate GAD
    deny ip any any
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Extended IP access list TelnetOut
    permit tcp host 10.0.0.10 host 20.0.0.20 eq telnet reflect GAD
    deny ip any any (155 matches)
```

Jest to tymczasowa lista dostępu, która została utworzona automatycznie w reakcji na zarejestrowanie ruchu wychodzącego spełniającego kryteria dopasowania instrukcji `permit` z klauzulą `reflect GAD`. Docelowy port ma wartość `11002`. Oznacza to, że takim portem posługuje się aplikacja generująca żądania telnet.

Po zakończeniu sesji lub po pewnym czasie nieaktywności lista jest usuwana. Aby zmienić dopuszczalny czas nieaktywności, należy wykonać w trybie konfiguracji ogólnej polecenie `ip refle`
`↪xive-list timeout` *`liczba_sekund`*. Instrukcja ta odnosi się do wszystkich refleksyjnych list dostępu skonfigurowanych w routerze. Domyślną wartością czasu nieaktywności jest pięć sekund.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 26.

Uwierzytelnianie w urządzeniach Cisco

Określenie **uwierzytelnianie** opisuje proces sprawdzania tożsamości użytkownika. Jego przykładem jest operacja, w której router żąda podania nazwy użytkownika i hasła.

W rozwiązaniach Cisco uwierzytelnianie ma dwie formy, są to: uwierzytelnianie standardowe oraz AAA (akronim od słów *Authentication*, *Authorization* i *Accounting* oznaczających uwierzytelnianie, autoryzację i rejestrowanie aktywności).

Uwierzytelnianie standardowe (nie AAA)

Standardowe uwierzytelnianie jest podstawową funkcją routerów oraz innych urządzeń z sieciowym systemem operacyjnym. Nie wymaga dostępności serwerów zewnętrznych. Jego konfiguracja i utrzymanie jest niezwykle łatwe, ale brakuje mu skalowalności i elastyczności konfiguracyjnej. Przykładem tych ograniczeń jest konieczność zapisywania nazw użytkowników i haseł w samym routerze. Jeśli pojedynczy użytkownik ma możliwość łączenia się wieloma urządzeniami (na przykład w sieci ISP), jego dane uwierzytelniające muszą zostać zapisane we wszystkich urządzeniach sieci, a trzeba pamiętać, że dostawcy usług internetowych obsługują niekiedy nawet dziesiątki tysięcy użytkowników. Zakładając, że zdefiniowanie konta użytkownika zajmuje jeden wiersz pliku konfiguracyjnego, cały plik musiałby składać się z setek stron.

Standardowe uwierzytelnianie może być stosowane przy niewielkiej liczbie kont lub jako mechanizm zapasowy w rozwiązaniu AAA.

Hasła na liniach dostępowych

Linie dostępowe są logicznymi lub fizycznymi interfejsami routera, które są wykorzystywane do zarządzania pracą urządzenia. Liniami dostępowymi są zarówno porty konsoli i AUX, jak też logiczne interfejsy VTY przeznaczone do połączeń telnet i SSH. Ustawienie hasła na linii dostępowej sprowadza się do wpisania polecenia `password`:

```
R1(config-line)#password hasło
```

Hasła mogą zawierać spacje. Ważna jest również wielkość użytych liter.

Brak hasła na liniach VTY powoduje wygenerowanie błędu podczas próby ustanowienia połączenia telnet z urządzeniem:

```
Password required, but none set //Hasło jest wymagane, ale nie zostało ustawione
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

VTY jest akronimem pochodzącym od angielskich słów *Virtual Teletype Terminal*, które oznaczają wirtualny terminal alfanumeryczny. W dawnych czasach terminale miały formę urządzeń o rozmiarach dużej maszyny do pisania i były połączone z komputerami typu mainframe za pomocą portów szeregowych. Gdy komputery stały się nieco bardziej wyrafinowane i opracowano protokół telnet, do obsługi połączeń konsolowych zaczęto wykorzystywać terminale wirtualne.

Zgodnie z domyślnymi ustawieniami hasła wprowadzane w postaci jawnego tekstu są w taki sam sposób zapisywane w pliku konfiguracyjnym. Aby włączyć funkcję szyfrowania wszystkich haseł odnotowywanych w pliku konfiguracyjnym, należy uruchomić usługę `password-encryp ↪tion`. Służy do tego polecenie `service password-encryption`. Oto przykład hasła zapisanego w pliku konfiguracyjnym za pomocą jawnego tekstu:

```
R1#sho run | include password
password Secret1
```

Włączenie usługi szyfrowania haseł i utajnienie wszystkich haseł w pliku sprowadza się do wykonania następujących instrukcji:

```
R1#conf t
Enter configuration commands, one per line. End with CNTL/Z.
R1(config)#service password-encryption
R1(config)#exit
```

A oto fragment pliku konfiguracyjnego, który został wygenerowany w czasie pracy usługi szyfrowania haseł:

```
R1#sho run | include password
password 7 073C244F5C0C0D54
```

Hasła szyfrowane za pomocą mechanizmów systemu IOS nie są całkowicie bezpieczne. Można je łatwo rozszyfrować, używając ogólnie dostępnych narzędzi.

Jeśli chcesz umożliwić ustanawianie połączeń z urządzeniem w protokole telnet bez obowiązku podawania hasła, możesz użyć polecenia `no login`:

```
R1(config-line)#no login
```

To oczywiście nie jest najlepsze rozwiązanie i powinno być stosowane jedynie w warunkach laboratoryjnych.

Polecenie `no login` jest dostępne tylko wówczas, gdy funkcja `aaa new-model` nie została wcześniej włączona. Wprowadzenie instrukcji `aaa new-model` powoduje zmianę składni i znaczenia polecenia `login`. Mechanizm ten został szczegółowo opisany w podrozdziale „Włączanie mechanizmu AAA” na stronie 437.

Przełączniki pracujące pod kontrolą systemów NX-OS dokonują uwierzytelnienia w nieco inny sposób. Nie posługują się hasłami terminali VTY, tak jak urządzenia z systemem IOS. Uwierzytelnianie jest realizowane z wykorzystaniem lokalnych lub zdalnych kont użytkowników.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Konfiguracja lokalnych kont użytkowników

Jako administrator możesz tworzyć lokalne konta użytkowników urządzenia. Wprowadzone w ten sposób nazwy kont posłużą do uwierzytelniania, w chwili gdy dana osoba będzie chciała się zalogować w urządzeniu. Taki sposób postępowania sprawdza się w sieciach o niewielkiej liczbie użytkowników lub gdy wykorzystuje się zewnętrzny serwer uwierzytelniania (zagadnienie to zostało omówione w podrozdziale „Uwierzytelnianie AAA" na stronie 437). Funkcja ta jest przydatna jako zapasowy mechanizm weryfikacji użytkowników wykorzystywany w czasie niedostępności zewnętrznego serwera.

Dodawanie i modyfikacja parametrów konta użytkownika należy do zadań instrukcji `username`. Towarzyszy jej wiele opcji. Skupimy się jednak tylko na tych, które są wykorzystywane podczas nawiązywania połączenia telnet lub SSH z urządzeniem.

Pierwszy etap prac polega na zdefiniowaniu nazwy użytkownika. W przedstawionym przykładzie jest nią ciąg tekstowy `GAD`:

```
R1(config)#username GAD ?
  access-class          Restrict access by access-class //Ogranicz dostęp na podstawie klasy dostępu
  autocommand           Automatically issue a command after the user logs in
                                                       //Po zalogowaniu wykonaj automatycznie podane polecenie
  callback-dialstring   Callback dialstring //Ciąg połączenia zwrotnego
  callback-line         Associate a specific line with this callback
                                                       //Przypisz połączeniu zwrotnemu określoną linię
  callback-rotary       Associate a rotary group with this callback
                                                       //Przypisz połączeniu zwrotnemu grupę cykliczną
  dnis                  Do not require password when obtained via DNIS
                                  //Nie wymagaj hasła, jeśli zostało pozyskane za pomocą mechanizmu DNIS
  nocallback-verify     Do not require authentication after callback
                                  //Nie wymagaj uwierzytelnienia po ustanowieniu połączenia zwrotnego
  noescape              Prevent the user from using an escape character //Zabroń użycia znaku specjalnego
  nohangup              Do not disconnect after an automatic command
                                  //Nie rozłączaj połączenia po wykonaniu automatycznego polecenia
  nopassword            No password is required for the user to log in //Nie wymagaj hasła do zalogowania
  password              Specify the password for the user //Określ hasło użytkownika
  privilege             Set user privilege level //Ustal poziom uprawnień użytkownika
  secret                Specify the secret for the user //Określ tajne hasło użytkownika
  user-maxlinks         Limit the user's number of inbound links
                                  //Ogranicz liczbę połączeń przychodzących danego użytkownika
  view                  Set view name //Ustal nazwę widoku
<cr>
```

Wprowadzenie instrukcji `username GAD` spowoduje utworzenie konta o nazwie `GAD` bez odpowiadającego mu hasła. Aby dodać hasło, trzeba dołączyć słowo kluczowe `password` oraz odpowiedni ciąg hasła:

```
R1(config)#username GAD password Secret1
```

W hasłach mogą występować spacje. Rozróżniana jest również wielkość liter. Jeśli usługa szyfrowania haseł jest wyłączona, wprowadzone wartości są zapisywane jawnym tekstem. Jeśli chcesz zapisać hasło w zaszyfrowanej formie bez użycia usługi, możesz zastąpić słowo kluczowe `password` dyrektywą `secret`.

```
R1(config)#username GAD secret Secret1
```

Wynikiem tej operacji jest zapisanie w pliku konfiguracyjnym następującego wiersza:

```
username GAD secret 5 $1$uyU6$6iZp6GLI1WGE1hxGDfQxc/
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Z każdym poleceniem i kontem użytkownika skojarzony jest pewien **poziom uprawnień** (ang. *privilege level*). Wartość poziomu uprawnień zawiera się w przedziale od 0 do 15. Standardowemu trybowi pracy użytkownika (EXEC) odpowiada poziom 1. Po wpisaniu polecenia `enable` i uwierzytelnieniu (z użyciem hasła określonego za pomocą instrukcji `enable password` lub `enable secret`) użytkownik rozpoczyna pracę w **trybie uprzywilejowanym**, któremu odpowiada wartość `15`. Jeśli chcesz, aby dany użytkownik mógł korzystać z poleceń trybu uprzywilejowanego bez konieczności wprowadzania hasła instrukcji `enable`, możesz przypisać mu wyższy poziom uprawnień. Innymi słowy, przypisanie użytkownikowi poziomu uprawnień o wartości `15` oznacza, że nie będzie musiał wprowadzać hasła (właściwego dla instrukcji `enable`), aby posługiwać się poleceniami trybu uprzywilejowanego:

```
R1(config)#username GAD privilege 15
```

Wykonanie kilku instrukcji w odniesieniu do tego samego konta użytkownika (w jednym wierszu lub w kilku wierszach) powoduje, że interpreter poleceń łączy instrukcje w jeden wpis konfiguracyjny. Oto wynik wpisania poleceń z dyrektywami `secret` i `privilege`:

```
username GAD privilege 15 secret 5 $1$uyU6$6iZp6GLI1WGE1hxGDfQxc/
```

Gdy użytkownik zaloguje się w systemie routera, zobaczy zgłoszenie typowe dla trybu uprzywilejowanego (symbolizowanego za pomocą znaku `#`) zamiast standardowego:

```
User Access Verification
Username: GAD
Password:
GAD-R1#
```

Inną ciekawą opcją instrukcji `username` jest możliwość określenia polecenia, które zostanie wykonane automatycznie po zalogowaniu użytkownika. Poniżej został pokazany przykład wykonania w ten sposób polecenia `show ip interface brief`:

```
R1(config)#username GAD autocommand show ip interface brief
```

Gdy użytkownik się zaloguje, zobaczy poniższe zestawienia, a następnie zostanie wylogowany:

```
User Access Verification
Username: GAD
Password:
Interface        IP-Address    OK?    Method  Status                 Protocol
FastEthernet0/0  10.100.100.1  YES    NVRAM   up                     up
FastEthernet0/1  unassigned    YES    NVRAM   administratively down  down
Serial0/0/0:0    unassigned    YES    unset   down                   down
```

Początkowo opisana funkcja nie wydaje się szczególnie użyteczna, ale jej ocena zmienia się, gdy uświadomimy sobie, że można w ten sposób skonfigurować konto technika, który będzie wykonywał tylko jedną instrukcję. Po co miałby mieć dostęp do innych poleceń?

Inną opcją dyrektywy `autocommand` jest wyznaczenie wstępnie zdefiniowanego menu poleceń (konfigurowanie zestawów instrukcji wykracza poza ramy tematyczne książki):

```
R1(config)#username GAD autocommand menu root-menu
```

Jeśli polecenie lub nazwa menu zostaną błędnie zdefiniowane, użytkownik nie będzie się mógł zalogować nawet po poprawnym uwierzytelnieniu.

Słowo kluczowe `nohangup` wyłącza funkcję automatycznego rozłączania po realizacji polecenia `autocommand`:

```
R1(config)#username GAD nohangup autocommand menu root-menu
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Polecenie `username` służy również do tworzenia kont użytkowników w systemach NX-OS, ale wymaga dodatkowo określenia roli użytkownika. Przed jego wykonaniem (tak naprawdę przed przystąpieniem do konfigurowania przełącznika) musisz ustalić, czy chcesz stosować silne hasła, czy nie.

Po pierwszym włączeniu przełącznika Nexus lub po pierwszym połączeniu z nowo utworzonym urządzeniem VDC zobaczysz na ekranie następujący dialog:

```
Daisy# switchto vdc Cozy
---- System Admin Account Setup ----
Do you want to enforce secure password standard (yes/no) [y]: //Czy chcesz wymusić stosowanie bezpiecznych
                                                              // haseł?
```

Wymuszanie użycia standardu bezpiecznych haseł jest dobrym rozwiązaniem i powinno być stosowane zawsze, gdy tylko jest to możliwe. Niekiedy jest to nawet konieczne, szczególnie jeśli musisz spełniać zalecenia PCI lub SOX. Włączenie opisywanej funkcji wiąże się z pewnymi utrudnieniami, ponieważ (podobnie jak w Uniksie) system będzie wymagał wprowadzenia 169 małych liter, dużych liter, cyfr, znaków alfabetu greckiego i sumeryjskiego. Tak naprawdę hasło musi się składać z co najmniej ośmiu znaków i nie musi uwzględniać symboli greckich i sumeryjskich. Niemniej często odnoszę wrażenie, że do prawidłowego zdefiniowania hasła muszę poznać jakiś starożytny język. Zgodnie z dokumentacją Cisco silne hasła mają następujące cechy:

- Składają się z co najmniej ośmiu znaków.
- Nie zawierają ciągów kolejnych znaków (na przykład „abcd”).
- Nie zawierają ciągów jednakowych znaków (na przykład „aaabbb”).
- Nie zawierają wyrazów słownikowych.
- Nie zawierają nazw własnych.
- Składają się z małych i dużych liter.
- Zawierają cyfry.

W tej samej dokumentacji podano pewne przykłady silnych haseł:

- If2CoM18,
- 2004AsdfLkj30,
- Cb1955S21.

Jeśli postępujesz tak jak ja, spędzisz wiele czasu, wymyślając hasło, które spełni wymagania, a jednocześnie będzie możliwe do zapamiętania. Czasami myślę, że nauczenie się języka sumeryjskiego byłoby łatwiejsze. Niemniej gdy pracuje się w rzeczywistym środowisku sieciowym, w którym istnieją realne zagrożenia, stosowanie silnych haseł jest koniecznością.

Ponieważ prezentowane przykłady odnoszą się do ćwiczeń laboratoryjnych, które nie są nadzorowane przez audytorów PCI, można w nich użyć mniej skomplikowanych wartości.

```
Do you want to enforce secure password standard (yes/no) [y]: no
Enter the password for "admin":   //Wprowadź hasło konta admin
```

W ramach buntu wpisałem hasło `admin` i postanowiłem kontynuować pracę. W rezultacie do pliku konfiguracyjnego został dodany wiersz:

```
username admin password 5 $1$uioc3i3ii$b071x923gsU09Ebu3/ role vdc-admin
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Role użytkowników są istotnym elementem systemu NX-OS. Konto założone w przedstawionym przykładzie obowiązuje tylko w danym urządzeniu logicznym (VDC). Więcej informacji na temat koncepcji VDC znajduje się w rozdziale 18. Rola `vdc-admin` jest przypisywana użytkownikowi, który ma pełne prawa administratora danej instancji VDC. Domyślnie w przełącznikach Nexus 7000 zdefiniowane są cztery role:

`network-admin`
: Użytkownik ma prawa odczytu i zapisu konfiguracji całego urządzenia NX-OS (konto jest dostępne jedynie w domyślnej instancji VDC).

`network-operator`
: Użytkownik ma prawa odczytu konfiguracji całego urządzenia NX-OS (konto jest dostępne jedynie w domyślnej instancji VDC).

`vdc-admin`
: Użytkownik ma prawa zapisu i odczytu, ale ograniczone do wskazanej instancji VDC.

`vdc-operator`
: Użytkownik ma prawa odczytu konfiguracji danej instancji VDC.

System umożliwia również definiowanie innych ról, a także dodawanie oraz odejmowanie uprawnień w ramach istniejących ról. Więcej informacji na ten temat znajduje się w dokumentacji urządzeń Cisco.

Uwierzytelnianie PPP

Uwierzytelnianie w ramach połączeń PPP jest realizowane z wykorzystaniem jednego z dwóch protokołów: PAP lub CHAP. Rozwiązanie PAP jest łatwiejsze do zrozumienia i zaimplementowania, ale ma pewne istotne ograniczenia — przekazuje hasła w formie jawnego tekstu. Mechanizm CHAP bazuje na bezpieczniejszym algorytmie, który nie wymaga przesyłania samego hasła. Obydwie metody zostały opisane w dokumencie RFC 1334. W tym samym dokumencie jest następująca uwaga odnosząca się do protokołu PAP:

> PAP nie jest bezpiecznym mechanizmem uwierzytelniania. Hasła są przekazywane w sieci za pomocą zwykłego tekstu. Protokół nie ma również zabezpieczeń przed powtórzeniem uwierzytelnienia oraz przed łamaniem haseł metodą prób i błędów. Zadanie monitorowania częstotliwości uwierzytelniania oraz liczby prób należy do jednostki końcowej.
>
> Jeśli w routerze zostały zdefiniowane mechanizmy bezpieczniejsze niż PAP (na przykład opisana poniżej metoda CHAP), MUSZĄ być proponowane do wykorzystania w pierwszej kolejności.

Konfigurując połączenia PPP, musisz zdecydować, który mechanizm uwierzytelniania będzie najlepszy w danych warunkach. Zazwyczaj CHAP okazuje się najkorzystniejszym rozwiązaniem (w związku z większym bezpieczeństwem operacji). Technika PAP znajduje zastosowanie tylko w implementacjach niewymagających wysokiego poziomu zabezpieczeń oraz w połączeniach, w których ryzyko podsłuchania transmisji jest minimalne. Zastosowanie którejkolwiek z wymienionych metod musi zostać poprzedzone utworzeniem lokalnego konta użytkownika (przynajmniej po stronie routera odbiorczego (wywoływanego)). Wymagania te są jednak różne w zależności od konkretnej konfiguracji.

Uwierzytelnianie w protokole PPP za pomocą standardowych kont (bez użycia mechanizmu AAA) zostało opisane w tym podrozdziale. Natomiast informacje na temat uwierzytelniania PPP z zastosowaniem mechanizmu AAA i protokołów PAP lub CHAP znajdują się w podrozdziale „Przypisywanie listy metod uwierzytelniania” na stronie 443.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

PAP

Protokół PAP można wykorzystywać do jednostronnego lub dwustronnego uwierzytelniania. Uwierzytelnianie jednostronne polega na tym, że tylko jedna strona inicjuje procedurę weryfikacji tożsamości. Operacja dwustronna pozwala natomiast na wzajemne uwierzytelnienie obydwu urządzeń końcowych.

Jednostronne uwierzytelnienie

W procedurze jednostronnej router wywołujący wysyła nazwę użytkownika i hasło, które muszą odpowiadać nazwie użytkownika i hasłu zdefiniowanym po stronie routera wywoływanego. Router wywołujący musi zawierać w pliku konfiguracyjnym instrukcję `ppp authentication pap callin` oraz polecenie `ppp pap sent user-name`, które wyznacza nazwę użytkownika i hasło. Poniżej znajdują się przykłady konfiguracji obydwu routerów. Instrukcje niezwiązane z uwierzytelnianiem PPP zostały pominięte.

Strona wywołująca:

```
interface BRI1/0
 encapsulation ppp
 ppp authentication pap callin
 ppp pap sent-username Bob password 0 ILikePie
```

Strona wywoływana:

```
username Bob password 0 ILikePie
!
 interface BRI1/0
 encapsulation ppp
 ppp authentication pap
```

Zastanawiałem się nad usunięciem z tekstu odniesień do interfejsów ISDN. Jeśli mieszkasz w dużym mieście, prawdopodobnie Twoje łącze ISDN już dawno zostało zastąpione połączeniem szerokopasmowym. Poza tym zagadnienia związane z systemami ISDN zostały usunięte nawet z laboratoriów przygotowujących do egzaminu Cisco CCIE. Okazuje się jednak, że w roku 2011 są jeszcze takie regiony kraju, w których nie ma szerokopasmowego dostępu do internetu. Pracowałem ostatnio w ogólnokrajowej sieci MPLS, w której musieliśmy utrzymywać infrastrukturę ISDN, ponieważ w pewnych lokalizacjach nie było łączy DSL ani internetu w sieciach kablowych. Mimo że rolę łączy zapasowych coraz częściej przejmują połączenia VPN, technologia ISDN wciąż jest w użyciu.

Dwustronne uwierzytelnienie

W przypadku dwustronnego uwierzytelniania słowo kluczowe `callin` nie jest potrzebne. Wymiana informacji uwierzytelniających jest dwukierunkowa, więc obydwa routery muszą dysponować nazwami użytkowników i hasłami. Do ich ustawienia służy polecenie `ppp pap sent username`. Po zakończeniu konfiguracji ustawienia w obydwu routerach muszą być zdefiniowane w taki sam sposób:

Strona wywołująca:

```
username Bob password 0 ILikePie
!
interface BRI1/0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
 encapsulation ppp
 ppp authentication pap
 ppp pap sent-username Bob password 0 ILikePie
```

Strona wywoływana:

```
username Bob password 0 ILikePie
!
interface BRI1/0
 encapsulation ppp
 ppp authentication pap
 ppp pap sent-username Bob password 0 ILikePie
```

Debugowanie procedury uwierzytelnienia w protokole PPP

Debugowanie uwierzytelniania w protokole PPP jest możliwe po wpisaniu polecenia `debug ppp authentication`. Prezentowane informacje są jednak dość trudne w interpretacji. Poniższy listing jest wynikiem uruchomienia procesu debugowania w routerze wywoływanym i przedstawia zapis nieudanego uwierzytelnienia wynikającego z błędnie wprowadzonego hasła.

```
8w4d: BR1/0:1 PPP: Using dialer call direction
8w4d: BR1/0:1 PPP: Treating connection as a callin
8w4d: BR1/0:1 PAP: I AUTH-REQ id 4 len 18 from "Bob"
8w4d: BR1/0:1 PAP: Authenticating peer Bob
8w4d: BR1/0:1 PAP: O AUTH-NAK id 4 len 27 msg is "Authentication failure" Bad password defined for
↪username Bob
```

Kolejny przykład pochodzi z komunikacji, w której dwustronne uwierzytelnienie zostało zakończone poprawnie. Zwróć uwagę na dodatkowe żądania. Pierwsze z nich pochodzi z routera wywołującego, który dostarcza do routera wywoływanego żądanie uwierzytelnienia (`AUTH-REQ`). Router wywoływany odsyła potwierdzenie (`AUTH-ACK`) oraz własne żądanie kierowane do routera wywołującego (`AUTH-REQ`). Ostatni z wyróżnionych wierszy świadczy o przesłaniu odpowiedzi (`AUTH-ACK`) z routera wywołującego i zakończeniu operacji. Litery `I` i `O` poprzedzające komunikaty `AUTH-ACK` i `AUTH-REQ` określają kierunek transmisji — przychodzący (**In**) lub wychodzący (**Out**).

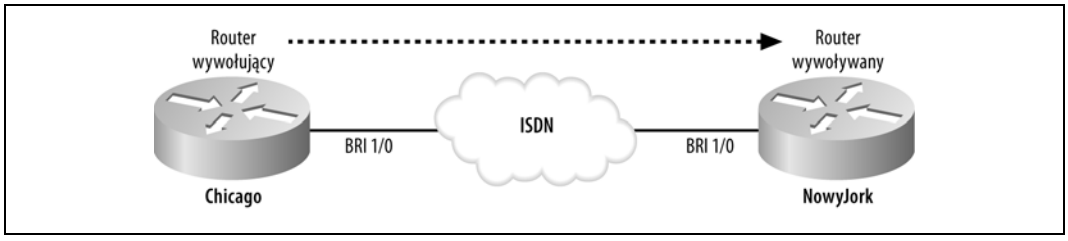
```
00:00:41: %LINK-3-UPDOWN: Interface BRI1/0:1, changed state to up
00:00:41: BR1/0:1 PPP: Using dialer call direction
00:00:41: BR1/0:1 PPP: Treating connection as a callin
00:00:43: %ISDN-6-LAYER2UP: Layer 2 for Interface BR1/0, TEI 68 changed to up
00:00:45: BR1/0:1 AUTH: Started process 0 pid 62
00:00:45: BR1/0:1 PAP: I AUTH-REQ id 2 len 17 from "Bob"
00:00:45: BR1/0:1 PAP: Authenticating peer Bob
00:00:45: BR1/0:1 PAP: O AUTH-ACK id 2 len 5
00:00:45: BR1/0:1 PAP: O AUTH-REQ id 1 len 17 from "Bob"
00:00:45: BR1/0:1 PAP: I AUTH-ACK id 1 len 5
00:00:46: %LINEPROTO-5-UPDOWN: Line protocol on Interface BRI1/0:1, changed state
to up
00:00:47: %ISDN-6-CONNECT: Interface BRI1/0:1 is now connected to 7802000 Bob
```

CHAP

Protokół CHAP gwarantuje wyższy poziom bezpieczeństwa niż PAP, ponieważ nigdy nie przesyła treści haseł. Generuje natomiast wartość skrótu, którą wyznacza na podstawie nazwy użytkownika i hasła. Uwierzytelnienie polega na sprawdzeniu, czy dostarczona wartość skrótu odpowiada wartości wygenerowanej lokalnie po stronie routera odbiorczego.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Na rysunku 26.1 została przedstawiona sieć złożona z dwóch routerów (`Chicago` i `NowyJork`) komunikujących się w protokole PPP. Podobnie jak w przypadku mechanizmu PAP uwierzytelnianie CHAP może mieć formę jednostronną lub dwustronną.



Rysunek 26.1. Połączenie ISDN z uwierzytelnianiem CHAP

Sposób działania protokołu CHAP sprawia, że jest on nieco trudniejszy do zrozumienia niż mechanizm PAP. Uwierzytelnianie metodą CHAP w routerach Cisco nie wymaga określenia nazwy użytkownika. Domyślnie jako nazwę użytkownika router wykorzystuje nazwę urządzenia (nazwę zdefiniowaną za pomocą instrukcji `hostname`). To rozwiązanie nie wydaje się szczególnie skomplikowane. Znacznie więcej kłopotów sprawia jednak zrozumienie zasad przetwarzania haseł.

Mechanizm PAP wymaga zdefiniowania co najmniej jednej pary wartości użytkownik-hasło po stronie routera wywoływanego. Gdy router wywołujący próbuje się uwierzytelnić, musi przesłać nazwę użytkownika i hasło odpowiadające parze tych wartości po stronie routera wywoływanego.

Również w rozwiązaniu CHAP każdy router musi dysponować nazwą użytkownika i hasłem, ale nazwa użytkownika musi odpowiadać nazwie routera pracującego po drugiej stronie łącza. Hasło z kolei musi być jednakowe w obydwu routerach. Zarówno w nazwach użytkowników, jak i w hasłach rozróżniana jest wielkość liter.

Konfigurując hasła, uważaj, aby nie popełnić błędu. Często zdarza się bowiem, że na końcu ciągu wpisywana jest spacja albo inny znak specjalny. Taką pomyłkę trudno później wykryć, ponieważ zapis w pliku konfiguracyjnym wygląda poprawnie. Jeśli masz pewność, że ustawienia są właściwe, a jednak połączenie nie funkcjonuje, usuń wiersze z hasłami i wpisz je ponownie (technika kopiuj-wklej zazwyczaj się nie sprawdza). Choć taki problem może wystąpić podczas definiowania różnych haseł, w konfiguracji mechanizmu CHAP okazuje się szczególnie uciążliwy.

Jednostronne uwierzytelnienie

Testowanie mechanizmu rozpoczniemy od zdefiniowania jednostronnego uwierzytelniania CHAP. Zwróć uwagę na to, że nazwa użytkownika zapisana w danym routerze musi odpowiadać nazwie routera zdalnego. Pamiętaj także, że hasła muszą mieć jednakową treść w obydwu urządzeniach.

Strona wywołująca (`Chicago`):

```
Hostname Chicago
!
username NowyJork
 password 0 Secret2
!
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
interface BRI1/0
 encapsulation ppp
 ppp authentication chap callin
```

Strona wywoływana (NowyJork):

```
Hostname NowyJork
!
username Chicago
 password 0 Secret2
!
interface BRI1/0
 encapsulation ppp
 ppp authentication chap
```

Przyjrzyjmy się teraz wynikom operacji debugowania zarejestrowanym podczas ustanawiania połączenia (zakończonego powodzeniem). Połączenie zostało zainicjowane przez router Chicago. Uważnie analizując listing, można zobaczyć, że router Chicago odbiera żądanie uwierzytelnienia (CHALLENGE) od routera NowyJork. Oto wyniki operacji debugowania mechanizmu po obydwu stronach:

Strona wywołująca:

```
20:08:11: %LINK-3-UPDOWN: Interface BRI1/0:1, changed state to up
20:08:11: BR1/0:1 PPP: Using dialer call direction
20:08:11: BR1/0:1 PPP: Treating connection as a callout
20:08:11: BR1/0:1 CHAP: I CHALLENGE id 3 len 28 from "NowyJork"
20:08:11: BR1/0:1 CHAP: O RESPONSE id 3 len 28 from "Chicago"
20:08:11: BR1/0:1 CHAP: I SUCCESS id 3 len 4
20:08:12: %LINEPROTO-5-UPDOWN: Line protocol on Interface BRI1/0:1, changed state to up
20:08:17: %ISDN-6-CONNECT: Interface BRI1/0:1 is now connected to 7802000
```

Strona wywoływana:

```
20:15:01: %LINK-3-UPDOWN: Interface BRI1/0:1, changed state to up
20:15:01: BR1/0:1 PPP: Using dialer call direction
20:15:01: BR1/0:1 PPP: Treating connection as a callin
20:15:02: BR1/0:1 CHAP: O CHALLENGE id 3 len 28 from "NowyJork"
20:15:02: BR1/0:1 CHAP: I RESPONSE id 3 len 28 from "Chicago"
20:15:02: BR1/0:1 CHAP: O SUCCESS id 3 len 4
20:15:03: %LINEPROTO-5-UPDOWN: Line protocol on Interface BRI1/0:1, changed state to up
20:15:07: %ISDN-6-CONNECT: Interface BRI1/0:1 is now connected to 7801000 Chicago
NowyJork#
```

Dwustronne uwierzytelnienie

Podobnie jak w przypadku mechanizmu PAP włączenie dwustronnego uwierzytelniania CHAP sprowadza się do usunięcia słowa kluczowego callin z instrukcji ppp authentication chap po stronie routera wywołującego:

Strona wywołująca:

```
Hostname Chicago
!
username NowyJork
 password 0 Secret2
!
interface BRI1/0
 encapsulation ppp
 ppp authentication chap
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Strona wywoływana:

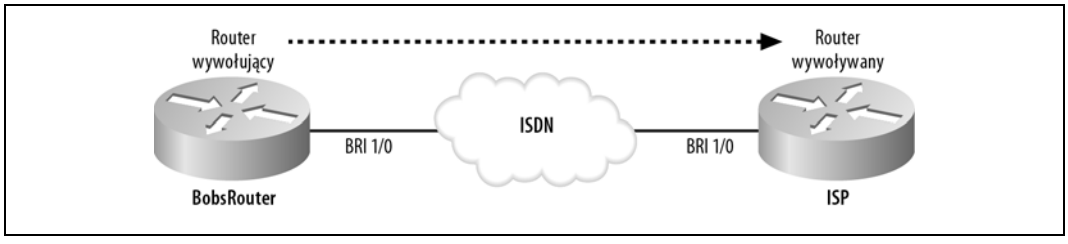
```
Hostname NowyJork
!
username Chicago
 password 0 Secret2
!
interface BRI1/0
 encapsulation ppp
 ppp authentication chap
```

Tym razem wynik polecenia `debug ppp authentication` jest bardziej rozbudowany, ponieważ uwierzytelnienie jest realizowane w obydwu kierunkach. Poniżej został przedstawiony jedynie listing pochodzący z routera wywoływanego:

```
20:01:59: %LINK-3-UPDOWN: Interface BRI1/0:1, changed state to up
20:01:59: BR1/0:1 PPP: Using dialer call direction
20:01:59: BR1/0:1 PPP: Treating connection as a callin
20:02:00: %ISDN-6-LAYER2UP: Layer 2 for Interface BR1/0, TEI 66 changed to up
20:02:00: BR1/0:1 CHAP: O CHALLENGE id 2 len 28 from "NowyJork"
20:02:00: BR1/0:1 CHAP: I CHALLENGE id 2 len 28 from "Chicago"
20:02:00: BR1/0:1 CHAP: Waiting for peer to authenticate first
20:02:00: BR1/0:1 CHAP: I RESPONSE id 2 len 28 from "Chicago"
20:02:00: BR1/0:1 CHAP: O SUCCESS id 2 len 4
20:02:00: BR1/0:1 CHAP: Processing saved Challenge, id 2
20:02:00: BR1/0:1 CHAP: O RESPONSE id 2 len 28 from "NowyJork"
20:02:00: BR1/0:1 CHAP: I SUCCESS id 2 len 4
20:02:01: %LINEPROTO-5-UPDOWN: Line protocol on Interface BRI1/0:1, changed state to up
20:02:05: %ISDN-6-CONNECT: Interface BRI1/0:1 is now connected to 7801000 Chicago
```

Zmiana wysyłanej nazwy użytkownika

Czasami użycie nazwy routera wywołującego nie jest możliwe. Przykładem takiej sytuacji może być nawiązywanie połączenia z routerem dostawcy usług internetowych. Na rysunku 26.2 przedstawiono kolejną sieć złożoną z dwóch routerów. W tym przypadku router `BobsRouter` ustanawia połączenia z routerem o nazwie ISP. Załóżmy, że dostawca usług internetowych (zarządzający routerem ISP) wygenerował konta i hasła dla swoich klientów. Nazwy kont nie odpowiadają nazwom urządzeń, które pracują w firmach korzystających z usługi.



Rysunek 26.2. Uwierzytelnianie CHAP z określoną nazwą użytkownika

Klient używający routera `BobsRouter` otrzymał konto o nazwie `Bob-01` oraz hasło `SuperSecret1`. Jeśli po stronie wywołującej zostanie wprowadzone dodatkowe polecenie `ppp chap hostname`, router będzie mógł przesłać nazwę `Bob-01` zamiast nazwy urządzenia (zdefiniowanej za pomocą instrukcji `hostname`). Po stronie wywoływanej wymieniona jest jedynie nazwa `Bob-01` i nie ma żadnego odniesienia do nazwy routera wywołującego.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Strona wywołująca:

```
Hostname BobsRouter
!
username ISP
 password 0 SuperSecret1
!
interface BRI1/0
 encapsulation ppp
 ppp authentication chap callin
 ppp chap hostname Bob-01
```

Strona wywoływana:

```
Hostname ISP
!
username Bob-01
 password 0 SuperSecret1
!
interface BRI1/0
 encapsulation ppp
 ppp authentication chap
```

Mimo że takie rozwiązanie działa poprawnie, prawdopodobnie jedyną sytuacją, w której się z nim spotkasz, będzie egzamin certyfikacyjny. Znacznie bardziej logiczne wydaje się zapisanie wysyłanej nazwy użytkownika i hasła jednocześnie w konfiguracji interfejsu. Nie ma wówczas potrzeby tworzenia konta, które będzie odpowiadało nazwie routera zdalnego.

Strona wywołująca:

```
Hostname BobsRouter
!
interface BRI1/0
 encapsulation ppp
 ppp authentication chap callin
 ppp chap hostname Bob-01
 ppp chap password 0 SuperSecret1
```

Strona wywoływana:

```
Hostname ISP
!
username Bob-01
 password 0 SuperSecret1
!
interface BRI1/0
 encapsulation ppp
 ppp authentication chap
```

Zastosowanie opisanej metody eliminuje wszelkie niejednoznaczności związane z dopasowywaniem nazw urządzeń i haseł. Cała operacja sprowadza się do zdefiniowania nazwy użytkownika i hasła po stronie routera wywołującego, do czego służy polecenie `ppp chap` wykonywane w trybie konfiguracji interfejsu. Po stronie wywoływanej wystarczy określić odpowiednią nazwę użytkownika i hasło. Nazwy urządzeń (instrukcje `hostname`) nie są potrzebne, mimo że zostały uwzględnione w prezentowanym przykładzie (oczywiście we wcześniej omawianych przykładach konfiguracji CHAP są niezbędne).

Ze wszystkich standardowych (poza rozwiązaniami AAA) metod uwierzytelniania połączeń PPP ta wydaje się najbezpieczniejsza i najłatwiejsza do zapamiętania.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Uwierzytelnianie AAA

Zgodnie z wcześniejszymi informacjami skrót AAA oznacza uwierzytelnianie, autoryzację i rejestrowanie aktywności. **Uwierzytelnianie** (ang. *authentication*) jest procesem potwierdzania tożsamości użytkownika w celu ustalenia, czy ma on prawo korzystania z urządzenia. **Autoryzacja** (ang. *authorization*) jest operacją blokowania dostępu lub zezwalania na dostęp do określonych funkcji urządzenia (po poprawnym uwierzytelnieniu). **Rejestrowanie aktywności** (ang. *accounting*) polega natomiast na zapisywaniu informacji o poczynaniach użytkownika, który został wcześniej uwierzytelniony i pomyślnie zautoryzowany. W dalszej części rozdziału przedstawione zostaną jedynie zagadnienia związane z uwierzytelnieniem jako najczęściej realizowaną czynnością z grupy AAA.

Aby wykorzystać mechanizm AAA w systemie IOS przełącznika lub routera, musisz wykonać następujące zadania:

- Włącz mechanizm AAA, wpisując polecenie `aaa new-model`.
- Skonfiguruj serwer zabezpieczeń (jeśli z niego korzystasz). Czynność ta polega na wprowadzeniu informacji o serwerze TACACS+ lub RADIUS.
- Utwórz listę metod uwierzytelniania za pomocą polecenia `aaa authentication`.
- Przypisz listę metod do interfejsu lub linii dostępowej.

W systemie NX-OS polecenia są zbliżone, choć nie ma w nim instrukcji `aaa new-model`.

Włączanie mechanizmu AAA

Aby włączyć mechanizm AAA, musisz wykonać polecenie `aaa new-model`:

```
Router(config)#aaa new-model
```

Jeśli tego nie zrobisz, instrukcje przedstawiane w dalszej części rozdziału nie będą dostępne. W systemach NX-OS tej czynności nie trzeba wykonywać.

Konfigurując system AAA po raz pierwszy, zachowaj szczególną ostrożność. Możesz bowiem w łatwy sposób zablokować sobie dostęp do routera. Wystarczy, że uruchomisz mechanizm AAA bez uprzedniego założenia konta użytkownika.

Konfiguracja serwera zabezpieczeń

Jedną z zalet systemu AAA jest możliwość wykorzystania zewnętrznego serwera do uwierzytelniania, autoryzacji i rejestrowania aktywności. Zastosowanie takiego serwera pozwala na przechowywanie danych o użytkownikach poza samym urządzeniem sieciowym. Zyskujemy w ten sposób scentralizowany mechanizm zarządzania profilami użytkowników, który dla samych użytkowników oznacza możliwość korzystania z wielu urządzeń, a dla administratorów stanowi ułatwienie w zarządzaniu prawami dostępu.

Protokoły RADIUS i TACACS+ są rozwiązaniami wykorzystywanymi w komunikacji z aplikacjami odpowiedzialnymi za uwierzytelnianie i autoryzację. Każdy z nich oprócz weryfikowania tożsamości użytkowników realizuje również wiele funkcji dodatkowych (rejestrowanie

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

wprowadzanych poleceń, zapisywanie przebiegu połączenia itp.). Obydwa są bardzo popularne. Oto krótka charakterystyka każdego z protokołów.

RADIUS

Usługa zdalnego uwierzytelniania użytkowników dzwoniących (RADIUS — ang. *Remote Authentication Dial-In User Service*) została opracowana przez firmę Livingston Enterprises (dzisiaj Lucent Technologies) z myślą o wykorzystaniu jej w serwerach sieciowych serii PortMaster. Urządzenia te były powszechnie stosowane przez firmy ISP w czasach, gdy dostęp do internetu zapewniały modemy 33,6 kb/s. Mechanizm RADIUS został później opisany w dokumentach RFC 2058 oraz 2059. Obecnie jest implementowany w wielu aplikacjach typu open source. System RADIUS zapewnia uwierzytelnienie i autoryzację użytkownika na podstawie jego profilu. Pracuje w protokole UDP i udostępnia funkcję uwierzytelniania na portach 1812 lub 1645, a funkcję rejestracji aktywności na portach 1813 i 1646.

TACACS+

System sterowania dostępem przez kontroler dostępu terminalowego (TACACS — ang. *Terminal Access Controller Access-Control System*) został opracowany na potrzeby zdalnego uwierzytelniania użytkowników serwerów Unix. Pierwotne rozwiązanie zostało jednak zastąpione przez protokół TACACS+, który jest rozwiązaniem firmy Cisco niezgodnym z pierwowzorem. Nowa wersja mechanizmu znajduje zastosowanie w uwierzytelnianiu i autoryzowaniu użytkowników urządzeń sieciowych. Funkcje te są jednak od siebie oddzielone. Działanie protokołu zostało opisane w propozycji zalecenia RFC (dostępnej pod adresem *http://tools.ietf.org/html/draft-grant-tacacs-02*). Bazuje ono na protokole TCP i porcie 49.

Firma Cisco zaleca zazwyczaj stosowanie protokołu TACACS+ zamiast RADIUS, choć obydwa umożliwiają zrealizowanie tego samego zadania. Jedną z wad mechanizmu RADIUS jest to, że nie pozwala on na ograniczenie zestawu poleceń wykonywanych przez użytkownika. Jeśli więc dostępność takiej funkcji jest niezbędna, wybierz TACACS+. Więcej informacji na temat różnic między systemami TACACS+ i RADIUS opisano w dokumencie Cisco o identyfikatorze 13838 (*http://www.cisco.com/en/US/tech/tk59/technologies_tech_note09186a0080094e99.shtml*).

Przełączniki Nexus wydają się bardziej ukierunkowane na technologie open source, czego potwierdzeniem może być to, że ich domyślnym protokołem uwierzytelniania jest RADIUS, a mechanizm TACACS+ trzeba włączyć za pomocą polecenia `feature tacacs+`.

Aby skorzystać z serwera zabezpieczeń, musisz skonfigurować **grupę serwerów** (ang. *server group*). Grupy serwerów są logicznymi zbiorami jednostek serwerowych, które występują w konfiguracji pod jedną nazwą. Można wykorzystać domyślną grupę serwerów lub utworzyć własną.

Domyślne grupy serwerów RADIUS i TACACS+

Serwery TACACS+ są definiowane globalnie w systemie routera za pomocą polecenia `tacacs-`
`↪server`. Lokalizację serwera TACACS+ określa słowo kluczowe `host`:

```
tacacs-server host 10.100.100.100
```

Zamiast adresu IP można podać nazwę jednostki, ale konieczne jest wówczas wskazanie serwera DNS, z którego router będzie musiał skorzystać. Jeśli wymienionych zostanie więcej serwerów, odwołania do nich będą generowane w kolejności ich występowania na liście:

```
tacacs-server host 10.100.100.100
tacacs-server host 10.100.100.101
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Odwołania do kolejnych serwerów na liście są realizowane tylko wówczas, gdy serwer wcześniejszy zwróci informację o błędzie lub okaże się niedostępny. Nieudane logowanie nie jest uznawane za błąd.

Wiele implementacji mechanizmu wymaga zdefiniowania tajnego klucza, który jest przesyłany wraz z zapytaniem. Klucz ten pełni rolę hasła, które potwierdza tożsamość serwera (a nie uwierzytelnianego użytkownika). Ustawienie klucza sprowadza się do wykonania polecenia `tacacs-server` z dyrektywą `key`:

```
tacacs-server key Secret
```

Klucz zostanie zapisany w pliku konfiguracyjnym za pomocą zwykłego tekstu, chyba że wcześniej została włączona usługa szyfrowania haseł. Wiersz definicji klucza wygląda wówczas tak:

```
tacacs-server key 7 01200307490E12
```

W systemach NX-OS wszystkie wymienione instrukcje można zastąpić jednym poleceniem. Oto sposób opisania dwóch serwerów za pomocą pojedynczych wierszy:

```
NX-7K-1-Cozy(config)#tacacs-server host 10.100.100.100 key Secret
NX-7K-1-Cozy(config)#tacacs-server host 10.100.100.101 key Secret
```

Serwery RADIUS są konfigurowane w analogiczny sposób. Większość ich parametrów jest podawana w ramach jednego polecenia. Takie samo ustawienie (z zastosowaniem klucza) sprowadza się do wykonania instrukcji:

```
radius-server host 10.100.200.200 key Secret
```

W rezultacie w pliku konfiguracyjnym zapisywany jest wiersz:

```
radius-server host 10.100.200.200 auth-port 1645 acct-port 1646 key Secret
```

Port 1645 jest domyślnym portem usługi RADIUS. Router dodaje go więc automatycznie.

Podobnie jak w przypadku serwerów TACACS+ możliwe jest zdefiniowanie kilku adresów:

```
radius-server host 10.100.200.200 auth-port 1645 acct-port 1646 key Secret
radius-server host 10.100.200.201 auth-port 1645 acct-port 1646 key Secret2
```

Drugi z serwerów zostanie wykorzystany tylko w przypadku błędu w działaniu pierwszego. Warto jednak zwrócić uwagę na to, że w rozwiązaniu bazującym na usługach RADIUS można zdefiniować różne klucze w odwołaniach do poszczególnych jednostek. Konfiguracja mechanizmu TACACS+ uwzględnia globalne ustawienie klucza odnoszące się do wszystkich serwerów.

W systemach NX-OS instrukcje definiujące odwołania w protokole RADIUS są podobne, choć zmieniona jest nieznacznie kolejność opcji:

```
NX-5K-1(config)#radius-server host 10.100.200.200 key secret auth-port 1645 acct-port 1646
NX-5K-1(config)#radius-server host 10.100.200.201 key secret2 auth-port 1645 acct-port 1646
```

Niestandardowe grupy serwerów

Załóżmy, że dysponujemy dwoma grupami serwerów TACACS+, które pełnią różne zadania — dwa serwery służą do uwierzytelnienia logujących się użytkowników, a kolejne dwa wspomagają proces uwierzytelniania w protokole PPP.

Zarówno system IOS, jak i NX-OS umożliwiają zdefiniowanie niestandardowych grup serwerów RADIUS oraz TACACS+. Do ich utworzenia służy instrukcja `aaa group server`, którą trzeba uzupełnić o słowo kluczowe `tacacs+` lub `radius` oraz nazwę grupy. W obydwu systemach polecenia mają taką samą składnię:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
aaa group server tacacs+ Login-Servers
 server 10.100.1.100
 server 10.100.1.101

aaa group server radius PPP-Radius
 server 10.100.200.200 auth-port 1645 acct-port 1646
 server 10.100.200.201 auth-port 1645 acct-port 1646
```

Analogicznie jak w poprzednich przykładach router automatycznie dodaje numery portów usługi RADIUS. Uzyskanie powyższego wyniku jest możliwe po wprowadzeniu następujących poleceń:

```
R1(config)#aaa group server radius PPP-Radius
R1(config-sg-radius)#server 10.100.200.200
R1(config-sg-radius)#server 10.100.200.201
```

Jeśli jeden z serwerów grupy wymaga zdefiniowania odrębnego klucza uwierzytelniania, zamiast instrukcji `server` można wykorzystać polecenie `server-private` (tylko w systemie IOS):

```
aaa group server tacacs+ Login-Servers
 server-private 10.100.1.72 key Secret
```

Tworzenie list metod uwierzytelniania

Lista metod wyznacza określoną kolejność użycia poszczególnych mechanizmów uwierzytelniania. Na przykład można skonfigurować router w taki sposób, aby w pierwszej kolejności odwoływał się do serwerów TACACS+, ale w razie niepowodzenia operacji wykorzystał lokalne konto użytkownika. Po utworzeniu samej listy należy zdefiniować interfejs, który będzie jej używał.

Router może uwierzytelniać użytkowników za pomocą kilku mechanizmów. Oto ich lista:

Login
: Uwierzytelnianie typu Login jest stosowane podczas weryfikowania praw dostępu użytkownika do interfejsu wiersza poleceń routera.

PPP
: Uwierzytelnianie PPP znajduje zastosowanie w operacjach ustanawiania połączeń punkt-punkt z użyciem interfejsów szeregowych lub interfejsów o charakterystyce zbliżonej do połączeń modemowych.

ARAP
: Protokół zdalnego dostępu w sieciach AppleTalk (ARAP — ang. *AppleTalk Remote Access Protocol*) jest przeznaczony do weryfikowania tożsamości użytkowników w sieciach AppleTalk.

NASI
: Interfejs asynchronicznych usług NetWare (NASI — ang. *NetWare Asynchronous Services Interface*) jest protokołem zdalnego dostępu przeznaczonym dla użytkowników sieci Novell Netware.

W praktyce stosowane są jedynie rozwiązania Login i PPP. Wraz z rozwojem szerokopasmowych łączy dostępowych przeznaczonych dla odbiorców indywidualnych oraz coraz powszechniejszym stosowaniem połączeń VPN w sieciach firmowych rozwiązania bazujące na modemach stają się historią. Od lat nie widziałem użytkowego rozwiązania ARAP lub NASI. Z tego względu w dalszej części rozdziału omówione zostaną jedynie mechanizmy uwierzytelniania typu Login i PPP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Uwierzytelnianie typu Login w systemie IOS

Tożsamość użytkownika logującego się do urządzenia pracującego pod kontrolą systemu IOS może zostać zweryfikowana na wiele sposobów. Domyślne rozwiązanie polega na zażądaniu wprowadzenia hasła i porównaniu go z wartością przypisaną interfejsowi lub linii dostępowej. Na przykład poniższe polecenia zabezpieczają dostęp do konsoli za pomocą hasła:

```
line con 0
 password Secret1
```

W przypadku wykorzystania mechanizmu AAA taki sposób działania nazywa się **uwierzytelnianiem linii** dostępowej (ang. *line authentication*). Oto lista metod uwierzytelniania dostępnych podczas logowania:

`enable`
: Wykorzystanie hasła zdefiniowanego za pomocą polecenia `enable` jako hasła do uwierzytelnienia.

`krb5`
: Pobranie informacji uwierzytelniających z serwera Kerberos 5.

`krb5-telnet`
: Wykorzystanie protokołu Kerberos 5 Telnet do uwierzytelniania połączeń w protokole telnet. Jeśli metoda ta jest używana, musi zostać wymieniona jako pierwsza.

`line`
: Wykorzystanie hasła zdefiniowanego za pomocą polecenia `line` jako hasła do uwierzytelnienia.

`local`
: Wykorzystanie lokalnie zdefiniowanych nazw użytkowników i haseł (zapisanych w konfiguracji urządzenia).

`local-case`
: Działania analogiczne do metody `local`, ale z uwzględnieniem wielkości liter.

`none`
: Wyłączenie uwierzytelniania. Ustawienie wartości `none` jako jedynej powoduje, że użytkownik uzyskuje dostęp do urządzenia bez konieczności potwierdzania swojej tożsamości.

`group radius`
: Pobranie informacji uwierzytelniających z serwerów RADIUS.

`group tacacs+`
: Pobranie informacji uwierzytelniających z serwerów TACACS+.

`Własna` (`group` *`nazwa-grupy`*)
: Odwołanie do grupy zdefiniowanej w lokalnej konfiguracji urządzenia.

Lista metod może się składać z jednej lub z wielu pozycji. Musi mieć jednak określoną nazwę. Następna instrukcja jest przykładem utworzenia listy o nazwie `GAD-Method`, która wykonuje zadania właściwe dla mechanizmu `login` za pomocą metody `local`.

```
aaa authentication login GAD-Method local
```

Jeśli lista składa się z kilku pozycji, poszczególne metody będą stosowane w kolejności ich występowania. Druga metoda zostanie więc wykorzystana jedynie w przypadku błędu w działaniu pierwszego rozwiązania. Błędem nie jest odmowa uwierzytelnienia (a na przykład niemożność nawiązania połączenia z serwerem).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Oto przykład definicji listy metod o nazwie `GAD-Method`, która wymusza odwołanie w pierwszej kolejności do serwera TACACS+, a następnie użycie lokalnych kont użytkowników:

```
aaa authentication login GAD-Method group tacacs+ local
```

Słowa kluczowe `tacacs+` i `radius` są w tym przypadku odwołaniami do globalnie dostępnych serwerów TACACS+ i RADIUS. Jeśli chcesz wskazać własną grupę serwerów, musisz wpisać nazwę tej grupy (użytą podczas wprowadzania polecenia `aaa group server`). W jednym z wcześniejszych przykładów utworzyliśmy grupę `Login-Servers`. Aby dodać ją do listy, wystarczy wpisać jej nazwę poprzedzoną słowem kluczowym `group`:

```
aaa authentication login default group Login-Servers
```

W zaprezentowanym przykładzie została wykorzystana metoda domyślna (`default`), która po zdefiniowaniu odnosi się automatycznie do wszystkich interfejsów.

Jeśli polegasz na zewnętrznych serwerach i wystąpi problem w komunikacji z nimi, może się okazać, że dostęp wszystkich użytkowników do routera zostanie zablokowany. Metoda `none` umożliwia zaakceptowanie wszystkich odwołań do routera bez konieczności potwierdzania tożsamości użytkownika. Jeśli zostanie wymieniona jako ostatnia, każdy będzie mógł uzyskać dostęp do routera, ale tylko wówczas, gdy zawiodą inne metody uwierzytelnienia.

```
aaa authentication login default group tacacs+ local none
```

Również w tym przypadku błąd oznacza nieudaną próbę komunikacji z wymienionym serwerem, a nie wprowadzenie niewłaściwego hasła. Oczywiście dodanie opcji `none` może być niebezpieczne. Ktoś mógłby bowiem przeprowadzić atak typu DoS na serwery uwierzytelniania i w ten sposób uzyskać dostęp do urządzeń (co jest sprzeczne z zamierzeniami większości administratorów sieci). Osobiście preferuję użycie metody `local-case` jako ostatniego mechanizmu:

```
aaa authentication login default group tacacs+ local-case
```

Zazwyczaj zapisuję na wszystkich routerach standardową nazwę użytkownika i hasło, które są wykorzystywane tylko w przypadku awarii serwera lub sieci. Wolę również ustawienie `local-case` niż `local`, ponieważ mogę używać małych i wielkich liter w nazwach kont i hasłach. Taki sposób postępowania jest jednak niemile widziany w środowiskach, w których realizowane są transakcje z użyciem kart kredytowych. Organizacje czuwające nad bezpieczeństwem tego typu operacji narzucają zasady przechowywania i pobierania danych użytkowników.

Tę samą listę metod można wykorzystać w uwierzytelnianiu typu Login i PPP. Musisz jednak pamiętać, że utworzenie listy metod `login` nie powoduje automatycznie utworzenia listy metod `ppp`. Jeśli chcesz wykorzystywać listy Login i PPP o takich samych nazwach, musisz utworzyć je niezależnie:

```
aaa authentication login GAD-Login group GAD-Servers none
aaa authentication ppp GAD-Login group GAD-Servers none
```

Uwierzytelnianie typu Login w systemie NX-OS

System NX-OS wymaga definiowania nazw użytkowników (ustawienia tego nie można zmienić). Wprowadzenie instrukcji `login` w konfiguracji interfejsów VTY lub CON nie ma więc sensu. Nie można również przypisać hasła do konsoli lub linii VTY (tak jak w systemie IOS). Można natomiast wymusić użycie w procesie uwierzytelniania serwerów RADIUS i TACACS+. Stosowna instrukcja nie odbiega od wykorzystywanej w systemie IOS:

```
NX-7K-1-Cozy(config)# aaa authentication login default group Login-Servers none
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Uwierzytelnianie PPP

Uwierzytelnianie PPP jest stosowane w chwili nawiązywania połączenia punkt-punkt z danym routerem. Znajduje zastosowanie zarówno w połączeniach modemowych, jak i w łączach o wysokiej przepustowości, takich jak T1.

Oto lista metod uwierzytelniania PPP w ramach systemu AAA:

`if-needed`
: Jeśli użytkownik został uwierzytelniony podczas ustanawiania połączenia terminalowego (TTY), operacja nie zostanie wykonana.

`krb5`
: Pobranie informacji uwierzytelniających z serwera Kerberos 5 (tylko na potrzeby protokołu PAP).

`local`
: Wykorzystanie lokalnie zdefiniowanych nazw użytkowników i haseł (zapisanych w konfiguracji urządzenia).

`local-case`
: Działania analogiczne do metody `local`, ale z uwzględnieniem wielkości liter.

`none`
: Wyłączenie uwierzytelniania. Ustawienie wartości `none` jako jedynej powoduje, że użytkownik nie musi potwierdzać swojej tożsamości.

`group radius`
: Pobranie informacji uwierzytelniających z serwerów RADIUS.

`group tacacs+`
: Pobranie informacji uwierzytelniających z serwerów TACACS+.

`Własna` (`group` *`nazwa-grupy`*)
: Odwołanie do grupy zdefiniowane w lokalnej konfiguracji urządzenia.

Wykorzystanie wymienionych metod na listach `ppp` nie odbiega w żaden sposób od użycia ich na listach `login`. Ciekawym uzupełnieniem jest tutaj metoda `if-needed`. Informuje router o tym, że uwierzytelnienie należy przeprowadzić tylko w przypadku, gdy użytkownik nie został wcześniej zweryfikowany (podczas korzystania z terminali VTY, konsoli lub linii AUX). Oto przykład listy metod `ppp`:

```
aaa authentication ppp default group tacacs+ local-case
```

Przypisywanie listy metod uwierzytelniania

Po utworzeniu listy metod uwierzytelniania trzeba ją powiązać z interfejsem lub linią dostępową (na której ma działać). W przypadku mechanizmu Login do przypisania listy służy polecenie `login`. Przedstawiona poniżej instrukcja zapewnia powiązanie utworzonej wcześniej listy metod o nazwie `GAD-Login` z liniami 0 – 4 terminali VTY. Wynikiem jej działania jest poprzedzenie sesji telnet uwierzytelnieniem z użyciem mechanizmów zdefiniowanych na liście `GAD-Login`:

```
line vty 0 4
 login authentication GAD-Login
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Aby przypisać listę metod PPP do wybranego interfejsu, trzeba najpierw ustawić na nim enkapsulację PPP. Samo uwierzytelnianie zostaje włączone po wykonaniu polecenia `ppp authentication`. Wraz z nazwą listy musisz określić rodzaje protokołu uwierzytelniania. Oto przykład użycia protokołu CHAP i listy metod `GAD-Login`:

```
interface Serial0/0/0:0
 no ip address
 encapsulation ppp
 ppp authentication chap GAD-Login
```

Jeśli lista metod PPP nie została wcześniej utworzona, router wygeneruje komunikat o błędzie, ale zaakceptuje polecenie:

```
R1(config-if)#ppp authentication pap GAD-Login
AAA: Warning, authentication list "GAD-Login" is not defined for PPP.
                                    //AAA: Uwaga, lista uwierzytelniania „GAD-Login" nie została zdefiniowana jako lista PPP
```

Wykorzystanie mechanizmu AAA w uwierzytelnianiu PAP lub CHAP jest rozwiązaniem znacznie bardziej skalowalnym niż użycie lokalnie definiowanych nazw użytkowników i haseł.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 27.

Podstawy działania zapór sieciowych

Zapora sieciowa (ang. *firewall*) jest urządzeniem, które zapobiega wprowadzaniu do sieci lub wyprowadzaniu z niej pewnego rodzaju ruchu. Zazwyczaj, choć nie zawsze, chroni nas przed zagrożeniami, jakimi są próby uzyskania dostępu do sieci wewnętrznej przez osoby pracujące w internecie. Zapory sieciowe są często instalowane w miejscach łączenia sieci firmowych z systemami organizacji, którym nie można ufać (na przykład z systemami partnerów handlowych).

Funkcję zapory sieciowej może pełnić niezależne urządzenie, oprogramowanie uruchomione w routerze lub moduł zainstalowany w innym komponencie sieciowym (na przykład w przełączniku Cisco 6500). Często również są implementowane w powszechnie wykorzystywanych routerach bezprzewodowych (czyli urządzeniach realizujących zadania rotuera, zapory sieciowej i punktu dostępu bezprzewodowego).

Nowoczesne firewalle realizują wiele zadań, nawet jeśli nie wchodzą w skład takich urządzeń jak routery bezprzewodowe. Odpowiadają za obsługę połączeń VPN. Uruchomione jako aplikacje programowe mogą współdzielić serwery z innymi usługami, takimi jak DNS lub usługa poczty elektronicznej (choć ich działanie powinno się ograniczać do zabezpieczenia komunikacji). Sprzętowa zapora sieciowa firmy Cisco — Cisco Adaptive Security Appliance (ASA) — jest wyposażana w dodatkowe funkcje związane z bezpieczeństwem, takie jak usługi VPN oraz systemy IDS i IPS.

Sprawdzone rozwiązania

Swoim klientom powtarzam często:

> Bezpieczeństwo jest stanem równowagi między komfortem pracy a paranoją.

Wszyscy chcemy czuć się bezpiecznie. Jeśli powiem, że zagwarantuję bezpieczeństwo Twojej rodzinie, czy nie będziesz chciał skorzystać z mojej oferty? A jeśli dodam, że osiągnięcie tego celu wymaga wstawienia metalowych krat w oknach, zamurowania drzwi garażowych i zastąpienia dotychczasowych drzwi wejściowych drzwiami pancernymi? Zaczniesz się zastanawiać, czy takie rozwiązanie zapewni odpowiedni komfort życia, czyż nie? Firmy również chcą wysokiego poziomu bezpieczeństwa komunikacji sieciowej, ale nie są skore do rezygnowania z wielu udogodnień, z których trzeba zrezygnować, aby osiągnąć ten cel.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Niedawno pracowałem jako konsultant dużej firmy mieszczącej się na Manhattanie, która miała problem z zabezpieczeniem systemu informatycznego. Przedstawiłem więc listę rozwiązań wdrożonych z powodzeniem w innych organizacjach. Oto odpowiedzi na zgłoszone propozycje:

Tokeny generujące hasła jednorazowe
„Nie chcemy tego. Tokeny są nieporęczne, a samo logowanie trwa zbyt długo".

VPN
„Podoba nam się koncepcja VPN, ale czy można je skonfigurować w taki sposób, aby nie trzeba było podawać hasła?".

Przeniesienie serwera poczty za zaporę sieciową
„Czy będziemy musieli wprowadzać więcej niż jedno hasło? Jeśli tak, to zapomnij o tym rozwiązaniu".

Rotacja haseł
„Nie ma mowy. Nie chcemy obowiązku zmiany haseł!".

Nie muszę wyjaśniać, że spotkanie było dla mnie pewnym wyzwaniem. Klient chciał wdrożenia zabezpieczeń i był bardzo podekscytowany wizją zastosowania takich samych mechanizmów, jakie zaimplementowaliśmy w jednym z dużych banków. Jednak w chwili gdy uświadomił sobie, jakie będą konsekwencje takiej operacji, zrezygnował z projektu — wycofał się z uwagi na konieczność ograniczenia swobody działania.

Prędzej czy później firmy uświadamiają sobie, że potrzebują pewnych mechanizmów zabezpieczeń, nawet jeśli ich wdrożenie wiąże się ze zmniejszeniem komfortu pracy. Niestety bardzo często to przekonanie jest wynikiem złamania dotychczasowych zabezpieczeń sieciowych. Niekiedy powodem jest również konieczność spełnienia norm narzucanych przez określone organizacje nadzorcze.

Projektując system zabezpieczeń, zawsze stosuj się do kilku reguł:

Proste rozwiązania są najlepsze
Zasada ta odnosi się do wszystkich działań związanych z konfiguracją sieci, ale jest szczególnie ważna w systemach zabezpieczeń. Definiując reguły działania zapory sieciowej, zadbaj o ich prostotę. Treść poszczególnych wpisów musi być łatwa do zinterpretowania. Jeśli to możliwe, stosuj nazwy zamiast wartości liczbowych. Jeśli firewall składa się z 60 000 reguł, z pewnością są w nich pewne luki. Tak, naprawdę istnieją zapory sieciowe z dziesiątkami tysięcy reguł filtrowania ruchu.

Monitoruj dzienniki pracy systemu
Generowane przez zaporę sieciową informacje statusowe muszą być zapisywane na serwerze. Należy je również okresowo analizować. Osoby, które nie przeglądają komunikatów generowanych przez zaporę, mają fałszywe poczucie bezpieczeństwa sieci. W rzeczywistości ich systemy mogą być atakowane w każdej chwili, a powiadomienia o tego typu problemach nie są wykorzystywane. Diagnozowałem kiedyś sieć firmy, która ciągle zwiększała szerokość pasma w łączu dostępowym, zadziwiona tym, jakiej przepustowości potrzebuje. Gdy przeanalizowałem dzienniki zapory sieciowej, okazało się, że większość pasma zajmuje ruch związany z repozytoriami danych utworzonymi przez hakerów na wewnętrznych serwerach FTP. Nikt nigdy nie przeglądał plików dziennika, więc problem nie został wcześniej wykryty.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Blokuj wszystko, akceptuj wybrane rodzaje ruchu

To bardzo prosta zasada, która zadziwiająco często jest ignorowana. Spośród wszystkich zaleceń przestrzeganie tego przynosi najwięcej korzyści.

W praktyce blokowanie ruchu w obydwu kierunkach często wydaje się zbyt kłopotliwe. Dlatego wspomniana zasada powinna być stosowana przynajmniej w odniesieniu do strumieni wejściowych. Żaden pakiet nie powinien zostać przepuszczony przez zaporę sieciową, jeśli nie jest kierowany do udokumentowanej i zatwierdzonej usługi firmowej. Ograniczenie ruchu wyjściowego jest oczywiście pożądanym działaniem, ale definiowanie jego akceptowalnych form może się okazać administracyjnym koszmarem. Jest to więc przykład wyższości komfortu pracy nad bezpieczeństwem rozwiązania. Do zalet opisanego sposobu postępowania należy zaliczyć to, że można w ten sposób wyeliminować współdzielenie plików w sieciach P2P oraz ułatwić sobie poszukiwanie przyczyn takich problemów jak niepoprawne działanie komunikatorów internetowych w komputerach użytkowników sieci. Wadą jest konieczność poświęcenia znacznej ilości czasu na wyjaśnienie pracownikom firmy, jakie rodzaje komunikacji są dozwolone, a jakie nie (chyba że zostało to opisane w odpowiednim dokumencie polityki bezpieczeństwa).

Domyślny sposób działania wielu zapór sieciowych, w tym urządzeń Cisco PIX i ASA, polega na przenoszeniu ruchu wychodzącego bez żadnych ograniczeń. Jego ograniczenie może się okazać konieczne w danym środowisku lub korporacji. Jednak z doświadczenia wiem, że większość małych i średnich firm nie chce zwiększać złożoności rozwiązania. Ponadto duża część z nich nie ma określonej polityki korzystania z internetu, więc definiowanie zasad przekazywania ruchu wychodzącego jest niezwykłym wyzwaniem.

Wszystkie systemy nienależące do firmy są elementami spoza sieci chronionej

Jest to kolejna zasada, o której nie pamiętają początkujący administratorzy sieci. Wszystkie systemy firm trzecich, które są przyłączone do Twojej sieci, powinny być kontrolowane przez zaporę sieciową. Połączenia z innymi organizacjami, włącznie z usługami weryfikacji kart kredytowych, muszą być weryfikowane przez firewall.

Z tej samej reguły wynika również, że wszystkie rozwiązania firmowe powinny pracować po stronie wewnętrznego interfejsu zapory sieciowej (lub w opisanej dalej strefie DMZ). Jedyne urządzenia instalowane w takim obszarze sieci, którego nie monitoruje zapora sieciowa, to koncentratory VPN. Są one często uruchamiane równolegle do firewalli. Wszelkie pozostałe systemy powinny być uruchamiane po jednej lub drugiej stronie granicy wyznaczanej przez zaporę sieciową. Podział ułatwia wyróżnienie strefy DMZ lub większej liczby takich stref.

Firewalle są obwiniane za wszelkie problemy w komunikacji. Jest to niemal reguła obowiązująca w każdej korporacji. Sądzę, że są dwa powody takiego stanu rzeczy. Po pierwsze, zawsze zrzucamy winę na rozwiązania, których nie rozumiemy. Po drugie, zapory sieciowe są przeznaczone do blokowania strumieni danych. Zatem gdy dane nie są przekazywane, podejrzewanie, że jest to wina firewalla, jest uzasadnione.

Strefa DMZ

W konfiguracji zapory sieciowej można zazwyczaj zdefiniować **strefę DMZ**. Akronim DMZ pochodzi od angielskich słów *demilitarized zone* oznaczających strefę zdemilitaryzowaną, która pierwotnie nie miała nic wspólnego z systemami komputerowymi. Jest to określenie wojskowe

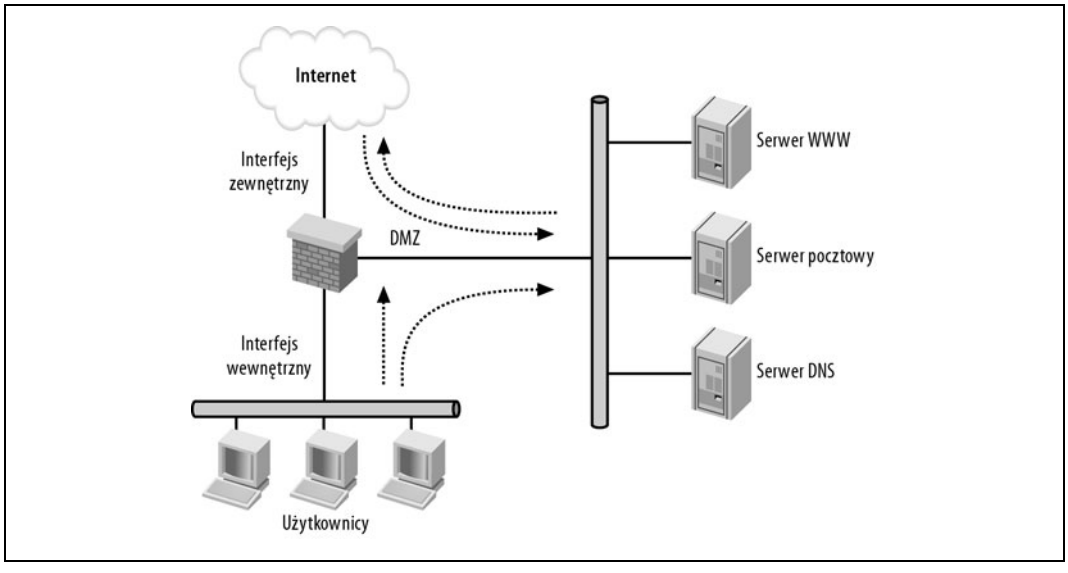
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

odnoszące się do obszaru rozdzielającego dwie walczące armie, w którym nie wolno prowadzić działań zbrojnych. Przykładem strefy zdemilitaryzowanej jest pas graniczny między Koreą Północną i Koreą Południową.

Wykorzystanie terminologii wojskowej w informatyce jest bardzo częste (od stref zdemilitaryzowanych po konie trojańskie). Najwyraźniej lubimy kojarzyć wszystkie nasze działania z operacjami militarnymi, choćby na poziomie nazewnictwa.

Z perspektywy bezpieczeństwa sieciowego strefa DMZ jest obszarem, który nie znajduje się ani po wewnętrznej, ani po zewnętrznej stronie zapory sieciowej. Powinny mieć do niej dostęp urządzenia sieci wewnętrznej (i zazwyczaj również zewnętrznej), ale mechanizmy zabezpieczeń firewalla powinny blokować wszelkie próby nawiązywania przez urządzenia DMZ połączeń z jednostkami w sieci wewnętrznej. Praca w sieci DMZ jest mniej bezpieczna od działania w sieci wewnętrznej, ale bezpieczniejsza niż praca w sieci zewnętrznej.

Typowy sposób konfiguracji strefy DMZ został przedstawiony na rysunku 27.1. Internet znajduje się po stronie zewnętrznego interfejsu zapory sieciowej, komputery użytkowników pracują po stronie wewnętrznej, a wszelkie serwery, które powinny być dostępne zarówno z internetu, jak i z sieci lokalnej, są umiejscowione w strefie DMZ.



Rysunek 27.1. Strefa DMZ

Konfiguracja widocznej na rysunku zapory sieciowej powinna być zgodna z następującymi wytycznymi:

Sieć wewnętrzna

Komputery sieci wewnętrznej muszą mieć możliwość nawiązywania połączeń z jednostkami innych sieci, ale połączenia inicjowane przez stacje spoza sieci wewnętrznej nie powinny być przepuszczane.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Sieć zewnętrzna
Komputery pracujące w sieci zewnętrznej nie mogą ustanawiać połączeń z jednostkami sieci wewnętrznej. Mogą natomiast komunikować się ze stacjami strefy DMZ.

Strefa DMZ
Jednostki strefy DMZ mogą nawiązywać połączenia z komputerami z sieci zewnętrznej, ale nie mogą się odwoływać do stacji pracujących w sieci wewnętrznej. Połączenia do strefy DMZ można ustanawiać z dowolnej sieci.

Największą zaletą przedstawionej topologii jest izolacja. Jeśli serwer pocztowy stanie się obiektem ataku i jego system zostanie przejęty, osoba atakująca nie będzie miała dostępu do komputerów użytkowników, które znajdują się w sieci wewnętrznej. Choć oczywiście włamywacz **będzie miał** dostęp do innych serwerów strefy DMZ, ponieważ pracują one w tej samej sieci (serwery można izolować jeszcze dokładniej dzięki specjalnym funkcjom przełącznika Cisco, takim jak prywatne sieci VLAN (ang. *private VLAN*), listy ACL przypisane do portów oraz mapy VLAN-ów; więcej informacji na ten temat znajduje się w rozdziale 25.).

Serwery strefy DMZ powinny zostać odizolowane od innych sieci w taki sam sposób, jaki zostałby wykorzystany, gdyby pracowały w internecie. W konfiguracji zapory sieciowej należy więc zapewnić dostęp do usług, ale tylko w potrzebnym zakresie. Oto przykład:

Serwer pocztowy
Niezbędne jest zaakceptowanie ruchu w protokołach POP, IMAP oraz SMTP (czyli ruchu TCP kierowanego do portów 110, 143 i 25). Dostęp z internetu do wszystkich pozostałych portów musi zostać zablokowany.

Serwer WWW
Dozwolony powinien być ruch w protokołach HTTP i HTTPS (czyli ruch TCP kierowany do portów 80 i 443). Wszystkie pozostałe porty powinny być niedostępne z internetu.

Serwer DNS
Akceptowane są jedynie zapytania DNS (w protokole UDP i kierowane do portu 53 oraz opcjonalnie w protokole TCP z portem 53). Wszystkie pozostałe porty powinny być niedostępne z internetu.

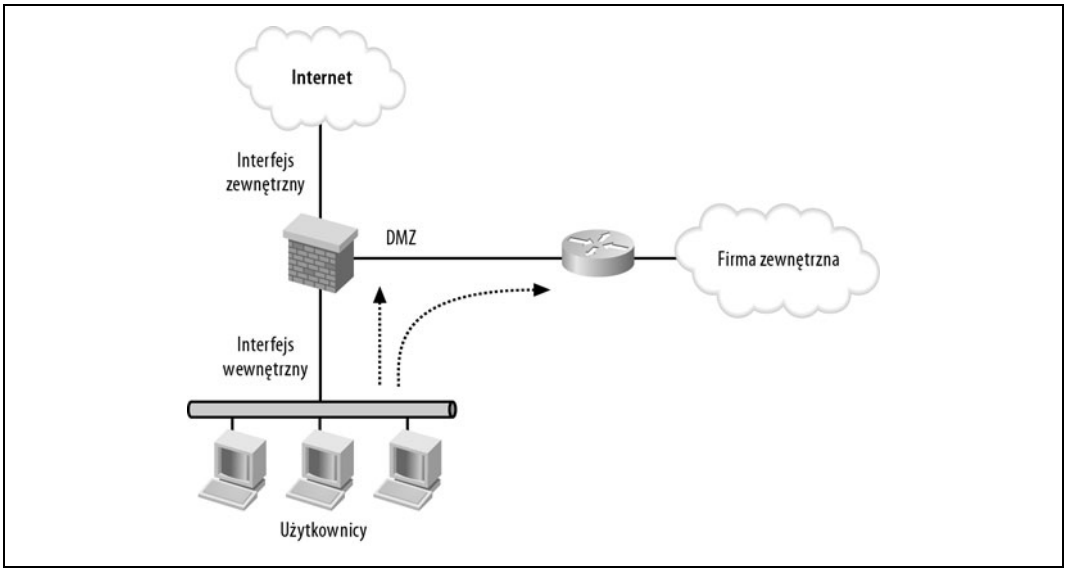
W idealnym rozwiązaniu dodatkowo między siecią wewnętrzną i strefą DMZ dozwolona mogłaby być komunikacja związana z zarządzaniem serwerami.

Inny przykład strefy DMZ

Innym częstym powodem wydzielania strefy DMZ jest konieczność utrzymywania połączenia z firmą zewnętrzną (na przykład z jednym z partnerów biznesowych). Na rysunku 27.2 przedstawiono topologię sieci, w której zewnętrzny kontrahent jest połączony za pomocą linii T1 z routerem strefy DMZ. Firmami zewnętrznymi mogą być na przykład jednostki zajmujące się obsługą transmisji z użyciem kart kredytowych lub dostawca półproduktów, który pozwala na korzystanie ze swojej bazy danych. Wiele przedsiębiorstw zleca również firmom zewnętrznym utrzymanie systemu pocztowego. Wówczas przedstawione łącze mogłoby zostać wykorzystane do komunikacji z serwerem pocztowym.

W sieci o budowie zbliżonej do pokazanej zapora sieciowa powinna działać w następujący sposób:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 27.2. Strefa DMZ z łączem do firmy zewnętrznej

Sieć wewnętrzna
Komputery sieci wewnętrznej muszą mieć możliwość nawiązywania połączeń z jednostkami innych sieci, ale połączenia inicjowane przez stacje spoza sieci wewnętrznej nie powinny być przepuszczane.

Sieć zewnętrzna
Komputery pracujące w sieci zewnętrznej nie mogą ustanawiać połączeń z jednostkami sieci wewnętrznej lub jednostkami strefy DMZ. Komputery sieci wewnętrznej mogą nawiązywać połączenia ze stacjami w sieci zewnętrznej, ale jednostki strefy DMZ takich połączeń nie mogą ustanawiać.

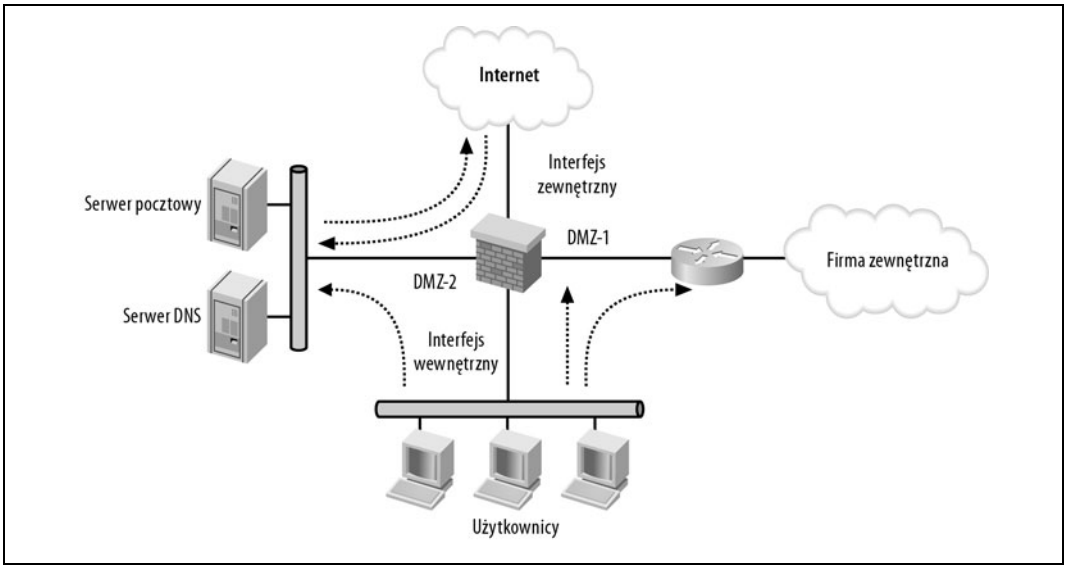
Strefa DMZ
Jednostki strefy DMZ nie mogą inicjować połączeń do żadnych sieci. Jedynie komputery sieci wewnętrznej mogą ustanawiać połączenia ze stacjami strefy DMZ.

Przykład wielu stref DMZ

Rzeczywiste funkcje sieci nie zawsze są tak precyzyjnie określone, jak to wynika z przedstawionych rysunków. Mimo że omówione rozwiązania są poprawne, wiele dużych organizacji wykorzystuje znacznie bardziej złożone topologie sieciowe. Często pojedyncza strefa DMZ nie jest wystarczająca.

Na rysunku 27.3 została przedstawiona sieć z wieloma strefami DMZ. Schemat jest połączeniem dwóch omówionych wcześniej rozwiązań. Po stronie zewnętrznej działają urządzenia internetowe, natomiast po stronie wewnętrznej pracują komputery użytkowników sieci. Obszar DMZ-1 obejmuje połączenie z firmą zewnętrzną. Z kolei w obszarze DMZ-2 uruchomione są serwery internetowe. Reguły przenoszenia ruchu są w zasadzie identyczne z opisanymi w poprzednim podrozdziale, ale trzeba się dodatkowo zastanowić, czy komputery strefy DMZ-1 powinny mieć prawo nawiązywania połączeń z jednostkami strefy DMZ-2 i odwrotnie. W tym przypadku odpowiedź jest przecząca.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 27.3. Wiele stref DMZ

Oto zalecany sposób konfiguracji zapory sieciowej:

Sieć wewnętrzna
: Komputery sieci wewnętrznej muszą mieć możliwość nawiązywania połączeń z jednostkami innych sieci, ale połączenia inicjowane przez stacje spoza sieci wewnętrznej nie powinny być przepuszczane.

Sieć zewnętrzna
: Komputery pracujące w sieci zewnętrznej nie mogą ustanawiać połączeń z jednostkami sieci wewnętrznej lub jednostkami strefy DMZ-1. Mogą jednak nawiązywać połączenia z komputerami pracującymi w sieci DMZ-2.

Strefa DMZ-1
: Jednostki strefy DMZ-1 nie mogą inicjować połączeń do żadnych sieci. Jedynie komputery sieci wewnętrznej mogą ustanawiać połączenia ze stacjami strefy DMZ-1.

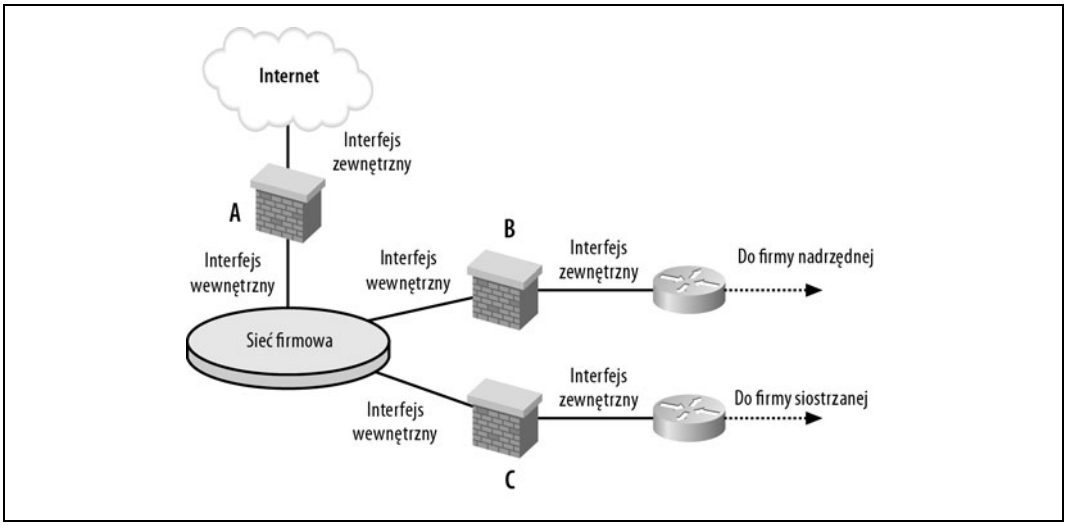
Strefa DMZ-2
: Komputery sieci DMZ-2 mogą inicjować jedynie połączenia kierowane do sieci zewnętrznej. Jednostki sieci wewnętrznej i zewnętrznej powinny mieć możliwość nawiązywania połączeń ze stacjami strefy DMZ-2.

Rozwiązania alternatywne

Zewnętrzny interfejs zapory sieciowej nie zawsze jest przyłączony do internetu. Wiele firm utrzymuje połączenia z innymi organizacjami (z firmami nadrzędnymi, siostrzanymi lub partnerskimi). W takich przypadkach (nawet jeśli firmy są od siebie zależne) dobrym rozwiązaniem jest oddzielenie firmy głównej od pozostałych właśnie za pomocą zapory sieciowej.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Na rysunku 27.4 przedstawiono topologię, w której *sieć firmowa* jest połączona z trzema zewnętrznymi obszarami. Zapora sieciowa A chroni ją przed dostępem ze strony internetu. Firewall B realizuje analogiczne zadanie na połączeniu z *firmą nadrzędną*, a zapora C reguluje dostęp od strony *firmy siostrzanej*.



Rysunek 27.4. Przykład zastosowania kilku zapór sieciowych

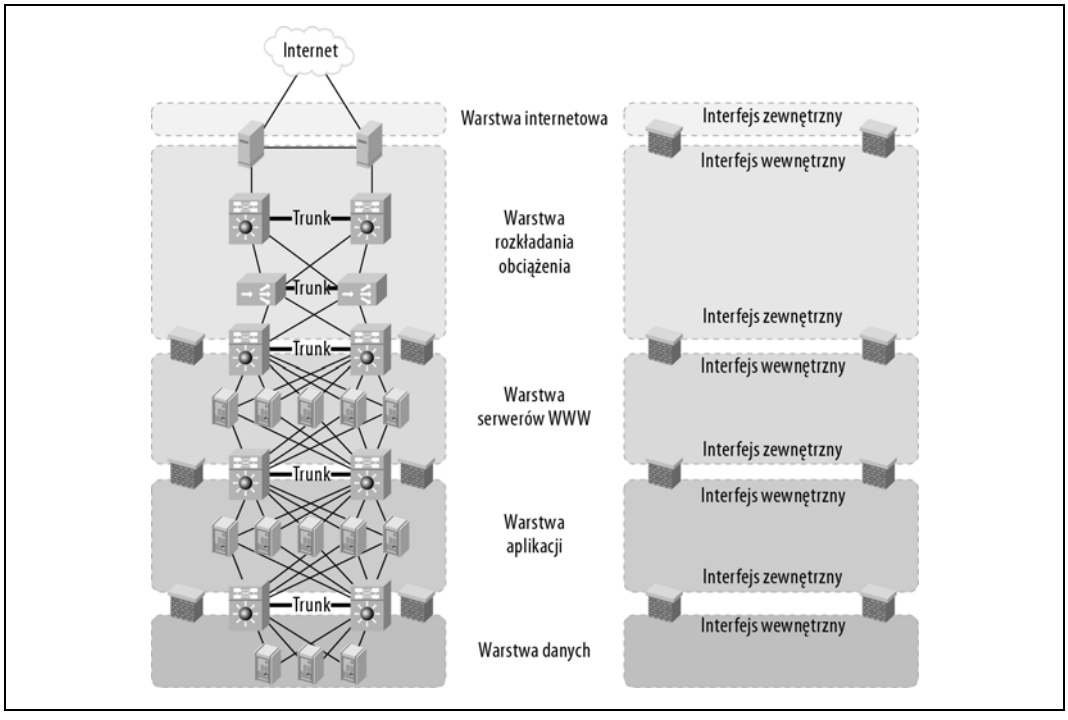
Każda z zapór sieciowych ma interfejs zewnętrzny i wewnętrzny. Wszystkie interfejsy wewnętrzne są przyłączone do tej samej sieci, ale każde z połączeń zewnętrznych prowadzi do innej sieci.

Firewalle są również często stosowane w systemach o architekturze wielowarstwowej (na przykład w usługach handlu elektronicznego). Wiele firm instaluje bowiem zapory sieciowe nie tylko w miejscu styku serwisu WWW z internetem, ale także pomiędzy warstwami. Rozwiązanie takie zostało pokazane na rysunku 27.5.

W systemach warstwowych (takich jak przedstawiony na rysunku) sieć wewnętrzna jednej zapory sieciowej jest jednocześnie siecią zewnętrzną kolejnej zapory. W analizowanym przykładzie do warstwy rozkładania obciążenia przyłączone zostały cztery firewalle. Dwa widoczne w górnej części rysunku (para zapewniająca połączenia zapasowe) łączą warstwę rozkładania obciążenia z warstwą internetową. Z ich perspektywy warstwa rozkładania obciążenia jest siecią wewnętrzną. Dolne zapory sieciowe (druga para z łączami zapasowymi) pośredniczy w komunikacji warstwy rozkładania obciążenia z warstwą serwerów WWW. W tym przypadku obszar rozkładania obciążenia jest siecią zewnętrzną.

Zapory sieciowe są kolejnymi elementami w zbiorze urządzeń sieciowych. Mimo że znane są ogólne zasady ich konfigurowania (wspomniane w tym rozdziale), o ostatecznym sposobie implementacji systemu zabezpieczeń decydują założenia biznesowe firmy.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 27.5. Usługa handlu elektronicznego

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 28.

Konfiguracja zapór sieciowych ASA

Informacje zamieszczone w tym rozdziale dotyczą zasad konfiguracji najczęściej wykorzystywanych funkcji zapory sieciowej ASA. Prezentowane przykłady odnoszą się do urządzenia ASA 5540, w którym dostępne są takie same polecenia, jakie obowiązują w innych komponentach z grupy ASA. Większość instrukcji można z powodzeniem zastosować również w zaporach sieciowych PIX, ale ponieważ urządzenia te nie są już sprzedawane (z wyjątkiem modułu usługi zapory sieciowej FWSM), prezentowane tutaj informacje odnoszą się do komponentów ASA.

Pomiędzy poszczególnymi modelami są pewne różnice. Na przykład podstawowy model urządzenia ASA 5505 nie może pracować w trybie przełączania awaryjnego (ang. *failover mode*) lub w trybie wielokontekstowym (ang. *multicontext mode*). Nieco inne jest także działanie modułu FWSM, który (jako moduł dodatkowy) nie ma konfigurowalnych interfejsów fizycznych.

Skrót ASA pochodzi od angielskich słów *Adaptive Security Appliance* i oznacza adaptacyjne urządzenie zabezpieczające. Z nazwy wynika więc, że urządzenia te mogą pełnić funkcje zapór sieciowych, koncentratorów VPN oraz elementów systemu IPS (lub IDS). Zagadnienia omówione w tym rozdziale odnoszą się jednak tylko do podstawowych usług zapory sieciowej.

Konfiguracja urządzania ASA może być nieco kłopotliwa dla osób, które są przyzwyczajone do pracy z systemami IOS. Mimo pewnych podobieństw sposób działania interpreterów poleceń istotnie się różni. Jednym z ciekawszych przykładów wspomnianych różnic jest możliwość wykonania w urządzeniu ASA (a nawet w zaporach sieciowych PIX) polecenia `show running-config` w trybie konfiguracji. Najnowsze wersje systemów IOS uwzględniają słowo kluczowe `do`, które pozwala na podobne operacje w trybie konfiguracji (na przykład `do show run`), ale instrukcje wykonywane w wierszu poleceń ASA wydają się bardziej naturalne. Ta sama funkcja jest zaimplementowana w interpreterze poleceń systemu NX-OS, czyli w przełączniku Cisco z grupy Nexus.

W tym samym czasie, gdy do sprzedaży trafiło urządzenie ASA, firma Cisco przebudowała oprogramowanie menedżera adaptacyjnych urządzeń zabezpieczających (ASDM — *Adaptive Security Device Manager*), które jest wykorzystywane do sterowania zaporami sieciowymi PIX i komponentami ASA. Po raz pierwszy oprogramowanie to stało się użytecznym narzędziem w pracy z firewallami Cisco. Mimo że wiele z omawianych w tym rozdziale rozwiązań można zaimplementować za pomocą menedżera ASDM, prezentując zasady ich wdrażania użyłem starego dobrego wiersza polecenia. Zasady posługiwania się oprogramowaniem ASDM wykraczają więc poza ramy tematyczne tego rozdziału. Nie zmienia to faktu, że ASDM jest doskonałym narzędziem i warto poświęcić trochę czasu na zapoznanie się z jego funkcjami.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Konteksty

Wiele urządzeń ASA obsługuje podział na konteksty. Wykorzystanie kontekstów jest podobne do działania systemu VMware, który umożliwia różnym serwerom i stacjom roboczym pracę w ramach jednego komputera fizycznego. Wielokontekstowy tryb pracy urządzenia ASA pozwala na definiowanie wielu wirtualnych zapór sieciowych w jednym urządzeniu fizycznym. Nie zapewnia jednak aż takiej uniwersalności jak rozwiązanie VMware, które umożliwia instalowanie dowolnych systemów operacyjnych. Tryb wielokontekstowy pozwala na definiowanie wielu logicznych jednostek ASA (i tylko ASA) w ramach pojedynczego urządzenia. Każde z urządzeń logicznych działa tak, jakby było niezależnym modułem. Istnieją jednak dość istotne ograniczenia w stosowaniu tego rozwiązania. W trybie wielokontekstowym nie są dostępne następujące funkcje:

- OSPF, RIP, EIGRP (w trybie wielokontekstowym można korzystać jedynie z routingu statycznego);
- VPN;
- routing w multiemisji (obsługa mostowania w multiemisji jest zachowana);
- detekcja zagrożeń;
- QoS;
- proxy telefonii.

Jak nietrudno się domyślić, wiele z tych mechanizmów jest niezbędnych w praktycznych instalacjach. Niemożność wykorzystania funkcji VPN często jest czynnikiem wykluczającym zastosowanie trybu wielokontekstowego. Jeśli jednak potrzeby firmy ograniczają się do wdrożenia klasycznej zapory sieciowej bez którejkolwiek z wymienionych funkcji, tryb wielokontekstowy może się okazać efektywnym sposobem na zaoszczędzenie pieniędzy. Więcej informacji na temat kontekstów pracy znajduje się w dalszej części rozdziału. Na tym etapie musisz jedynie wiedzieć, że domyślny ciąg zgłoszenia w wielokontekstowym urządzeniu ASA ma składnię `nazwa_urządzenia/nazwa_kontekstu#`. System wykorzystany w prezentowanych przykładach został nazwany `ASA-5540`. Do realizacji poleceń w ramach kontekstów został z kolei użyty kontekst `GAD`. Zatem zgłoszenie trybu konfiguracji w kontekście systemowym (domyślnym) ma postać:

```
ASA-5540#
```

Natomiast w kontekście `GAD` wiersz polecenia uwzględnia zgłoszenie:

```
ASA-5540/GAD#
```

Interfejsy i poziomy bezpieczeństwa

Każdy interfejs zapory sieciowej ASA musi mieć przypisane: nazwę fizyczną, nazwę logiczną, poziom bezpieczeństwa oraz adres IP. Oczywiście administrator może również zmieniać inne parametry interfejsów, na przykład szybkość transmisji czy tryb dupleksu.

W urządzeniu ASA 5540 standardowe fizyczne interfejsy są oznaczone symbolami od G0/0 do G0/3. Interfejsy dodane za pomocą karty rozszerzenia są oznaczane kolejnymi numerami począwszy od G1/0. Każdy z interfejsów musi mieć przypisaną nazwę logiczną. Nie wolno jednak stosować słów `inside` i `outside`, które mają specjalne znaczenie.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Wszystkie modele urządzeń ASA mogą pracować w trybie **transparentnym** (ang. *transparent*). Pełnią wówczas funkcję mostu, stanowiąc **„wybój na łączu"** (ang. *bump on the wire*), co moim zadaniem jest najgłupszym określeniem technicznym, jakie wymyślono. Praca w trybie transparentnym umożliwia włączenie zapory sieciowej do sieci bez konieczności zmiany adresacji w żadnym z jej obszarów. Konieczne jest jedynie zdefiniowanie adresu IP potrzebnego do zarządzania urządzeniem. Możliwość użycia transparentnego firewalla początkowo wydaje się doskonałym rozwiązaniem.

Jestem pewien, że jesteś zaintrygowany. Zalecam jednak zrezygnowanie z takiego sposobu konfiguracji zapory sieciowej. Wszystkie wdrożenia transparentnego firewalla, które widziałem, kończyły się niepowodzeniem. Dlaczego? Ponieważ instalację trzeba poprzedzić opracowaniem nowego projektu sieci, a nikt nie chce poświęcać na to czasu i pieniędzy. Transparentne firewalle są przeznaczone do stosowania w sieciach, w których nie przewidziano instalacji zapory sieciowej. Stanowczo zalecam jednak przebudowę sieci i zainstalowanie tradycyjnego firewalla warstwy trzeciej zamiast włączania pracy w trybie transparentnym.

Każdemu z interfejsów trzeba również przypisać poziom bezpieczeństwa. Zgodnie z domyślnym ustawieniem interfejsy o wyższym poziomie bezpieczeństwa mogą przekazywać pakiety do interfejsów o niższej wartości tego parametru. Transmisja w odwrotnym kierunku jest zabroniona. Sam poziom bezpieczeństwa jest definiowany za pomocą wartości liczbowych z przedziału od 0 do 100. Domyślna wartość parametru dla interfejsu wewnętrznego (`inside`) wynosi 100.

Z kolei interfejs zewnętrzny (`outside`) ma standardowo przypisany zerowy poziom bezpieczeństwa. Jeśli administrator przypisze wybranemu interfejsowi nazwę `inside` lub `outside`, automatycznie zdefiniowany zostanie również odpowiedni poziom bezpieczeństwa:

```
ASA-5540/GAD(config-if)#nameif outside
INFO: Security level for "outside" set to 0 by default.
```

Dodanie trzeciego interfejsu wiąże się zazwyczaj z wyznaczaniem pewnej pośredniej wartości tego parametru. Od wersji 7.x oprogramowania urządzeń ASA (PIX) istnieje możliwość definiowania wielu interfejsów o jednakowym poziomie bezpieczeństwa. Przykład konfiguracji trzech interfejsów zapory sieciowej (zewnętrznego, wewnętrznego oraz prowadzącego do strefy DMZ) został przedstawiony na rysunku 28.1.

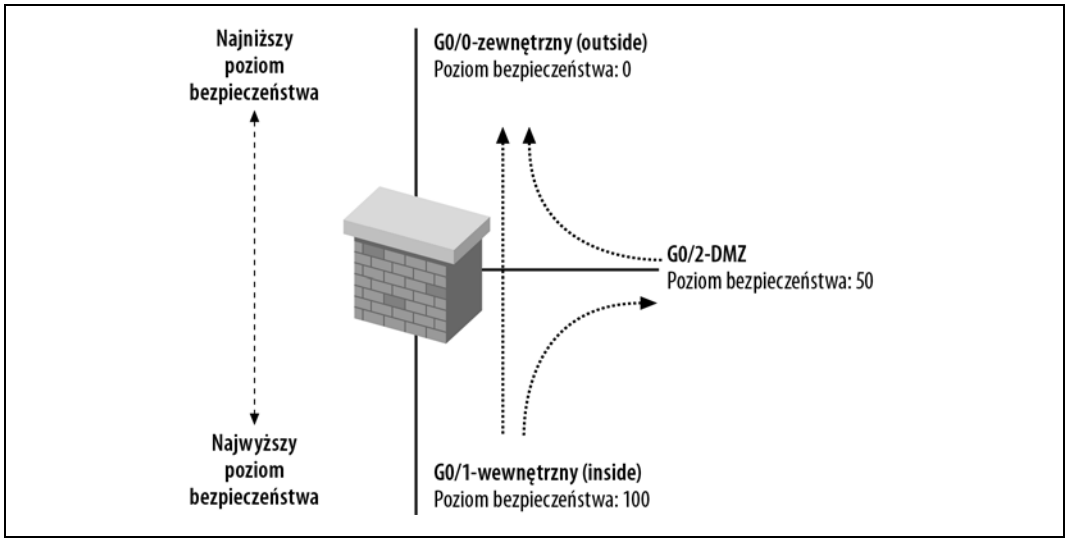
Zgodnie z domyślnymi ustawieniami ruch może być przenoszony w następujących kierunkach:

- z sieci wewnętrznej do sieci zewnętrznej;
- z sieci wewnętrznej do strefy DMZ;
- ze strefy DMZ do sieci zewnętrznej, ale nie do sieci wewnętrznej;
- ruch dostarczony do interfejsu zewnętrznego nie może zostać przekazany do żadnego innego interfejsu.

Ograniczenia wynikające z poziomu bezpieczeństwa interfejsu można nadpisać, definiując listę dostępu. Opisane reguły nie są więc niezmienne.

Aby ręcznie zdefiniować priorytet interfejsu, wystarczy wprowadzić polecenie `security-level` (w trybie konfiguracji interfejsu). Oto przykład ustawienia interfejsu wewnętrznego:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 28.1. Poziomy bezpieczeństwa interfejsów urządzenia ASA

```
ASA-5540/GAD(config)#int g0/0
ASA-5540/GAD(config-if)#nameif inside
ASA-5540/GAD(config-if)#security-level 100
```

oraz zewnętrznego:

```
ASA-5540/GAD(config)#int g0/1
ASA-5540/GAD(config-if)#nameif outside
ASA-5540/GAD(config-if)#security-level 0
```

Użyte wartości odpowiadają domyślnym parametrom urządzenia ASA. Aby dodać trzeci interfejs (zgodnie z rysunkiem 28.1), należy wpisać następujące instrukcje:

```
ASA-5540/GAD(config)#int g0/2
ASA-5540/GAD(config-if)#nameif DMZ
ASA-5540/GAD(config-if)#security-level 50
```

Za ustawienie szybkości transmisji i trybu dupleksu odpowiada polecenie `interface`:

```
ASA-5540(config)#int g0/0
ASA-5540(config-if)#speed 100
ASA-5540(config-if)#duplex full
```

W przypadku trybu wielokontekstowego parametry interfejsów można zmienić, jedynie pracując w kontekście systemowym. Więcej informacji na ten temat znajduje się w podrozdziale „Zarządzanie kontekstami” na stronie 465.

Aby sprawdzić stan interfejsów, należy wprowadzić polecenie `show interface`. Zestawienie wynikowe jest zbliżone do generowanego przez analogiczne instrukcje systemu IOS, choć jego treść jest zależna w pewnym stopniu od kontekstu, w którym zostało wpisane polecenie. Więcej informacji jest udostępnianych w kontekście systemowym, ponieważ umożliwia on modyfikowanie fizycznych parametrów interfejsów. Poniżej znajduje się przykład listingu uzyskanego w kontekście systemowym.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
ASA-5540#sho int g0/0
Interface GigabitEthernet0/0 "", is administratively down, line protocol is down
  Hardware is i82546GB rev03, BW 1000 Mbps, DLY 10 usec
        Full-Duplex, 100 Mbps
        Available for allocation to a context
        MAC address 0026.0b31.36d8, MTU not set
        IP address unassigned
        0 packets input, 0 bytes, 0 no buffer
        Received 0 broadcasts, 0 runts, 0 giants
        0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored, 0 abort
        0 L2 decode drops
        0 packets output, 0 bytes, 0 underruns
        0 output errors, 0 collisions, 0 interface resets
        0 late collisions, 0 deferred
        0 input reset drops, 0 output reset drops
        input queue (curr/max packets): hardware (0/0) software (0/0)
        output queue (curr/max packets): hardware (0/0) software (0/0)
```

A oto wynik wykonania tego samego polecenia w innym kontekście:

```
ASA-5540/GAD#sho int g0/0
Interface GigabitEthernet0/0 "inside", is down, line protocol is down
        MAC address 1200.0000.0700, MTU 1500
        IP address 10.10.10.1, subnet mask 255.255.255.0
  Traffic Statistics for "inside":
        0 packets input, 0 bytes
        0 packets output, 0 bytes
        0 packets dropped
```

Nazwy

Jedną z funkcji ułatwiających pracę z systemami ASA i PIX jest możliwość prezentowania nazw zamiast adresów IP. Aby włączyć tę funkcję, wystarczy wpisać w trybie konfiguracji ogólnej polecenie `names`.

```
ASA-5540/GAD(config)#names
```

Po włączeniu mechanizmu możesz przypisać każdemu adresowi IP dowolnie wybraną nazwę. Zasada jego działania jest zbliżona do pracy systemu DNS, z tą jednak różnicą, że nazwy są zapisywane lokalnie w konfigurowanym urządzeniu ASA lub PIX. Załóżmy, że adres IP `10.10.10.10` należy do komputera o nazwie `SerwerPlikow`. Użycie polecenia `name` pozwala na przypisanie jednostce `SerwerPlikow` podanego adresu:

```
ASA-5540/GAD(config)#name 10.10.10.10 SerwerPlikow
```

Dzięki tej operacji konfiguracja listy dostępu może zostać przeprowadzona w następujący sposób:

```
ASA-5540/GAD(config)#access-list Lista permit tcp any host 10.10.10.10 eq www
```

Listy dostępu (wraz z opcjami charakterystycznymi dla urządzeń ASA) zostały szczegółowo omówione w rozdziale 25. Nie uwzględniono w nim jednak skutków użycia nazw (takich jak `Lista`). Możesz ich więc używać, ale na własne ryzyko.

W czasie wyświetlania konfiguracji adres IP komputera jest zastępowany jego nazwą:

```
ASA-5540/GAD#sho run | include Lista
access-list Lista extended permit tcp any host SerwerPlikow eq www
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli wolisz analizować adresy IP, możesz wyłączyć funkcję prezentowania nazw, negując instrukcję `name`. Uzyskanie tego samego wyniku, ale z adresem IP, sprowadza się do wykonania poniższych poleceń:

```
ASA-5540/GAD(config)#no names
ASA-5540/GAD(config)#sho run | include Lista
access-list Dingle extended permit tcp any host 10.10.10.10 eq www
```

Listing wynikowy polecenia `show interface` zawsze zawiera adresy IP, nawet jeśli funkcja zamiany ich na nazwy została włączona.

Aby wyświetlić zestawienie wszystkich nazw skonfigurowanych w danym urządzeniu ASA, należy wykonać instrukcję `show names`:

```
ASA-5540/GAD #sho names
name 10.10.10.1 ASA-Outside
name 10.10.10.10 SerwerPlikow
name 192.168.1.1 ASA-Inside
```

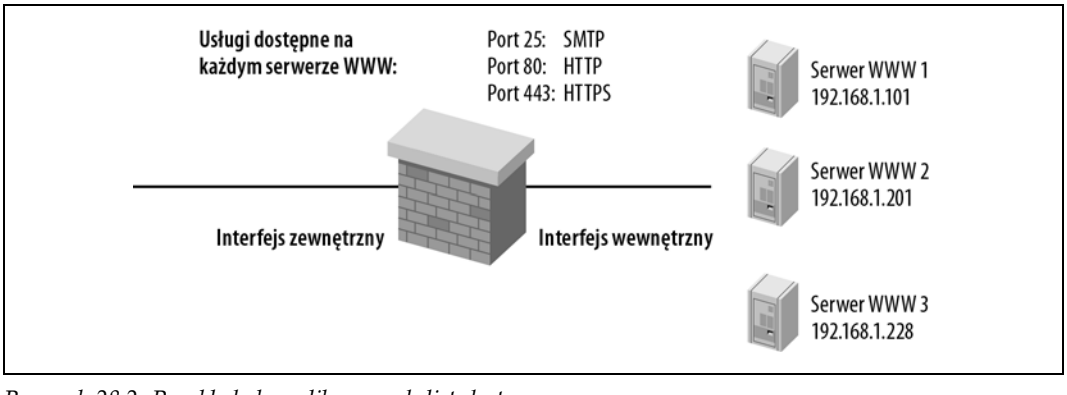
Nowa funkcja systemu operacyjnego jest również dostępna w poleceniu `show run names`:

```
ASA-5540/GAD #sho run names
name 10.10.10.1 ASA-Outside
name 10.10.10.10 SerwerPlikow
name 192.168.1.1 ASA-Inside
```

Włączenie mechanizmu translacji adresów na nazwy znacznie ułatwia analizowanie konfiguracji urządzenia ASA. W rozbudowanych plikach konfiguracyjnych liczba adresów IP bywa ogromna, a zapamiętanie ich wszystkich jest praktycznie niemożliwe.

Grupy obiektów

Mechanizm grup obiektów pozwala na reprezentowanie zbiorów adresów sieciowych, adresów IP, protokołów lub usług za pomocą jednej nazwy. Jest on niezwykle użyteczny podczas definiowania złożonych list dostępu. Wyobraź sobie sytuację przedstawioną na rysunku 28.2. W sieci pracują trzy serwery WWW, z których każdy obsługuje trzy takie same protokoły: SMTP (port TCP 25), HTTP (port TCP 80) oraz HTTPS (port TCP 443).



Rysunek 28.2. Przykład skomplikowanych list dostępu

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W typowej sieci korporacyjnej serwery WWW nie powinny pracować w części wewnętrznej, lecz w strefie DMZ. Choć oczywiście określenie „typowa sieć" jest bardzo uznaniowe.

Ponieważ adresy IP poszczególnych serwerów nie pochodzą z jednego zakresu i nie można ich objąć jedną maską podsieci, każdemu serwerowi musi odpowiadać osobny wpis na liście dostępu. Ponadto konieczne jest zdefiniowanie osobnego wpisu dla każdego protokołu każdego serwera.

W rezultacie opisanie trzech protokołów na każdym z trzech serwerów zajmuje dziewięć wierszy pliku konfiguracyjnego:

```
access-list In permit tcp any host 192.168.1.101 eq smtp
access-list In permit tcp any host 192.168.1.101 eq www
access-list In permit tcp any host 192.168.1.101 eq https
access-list In permit tcp any host 192.168.1.201 eq smtp
access-list In permit tcp any host 192.168.1.201 eq www
access-list In permit tcp any host 192.168.1.201 eq https
access-list In permit tcp any host 192.168.1.228 eq smtp
access-list In permit tcp any host 192.168.1.228 eq www
access-list In permit tcp any host 192.168.1.228 eq https
```

Już ten przykład wydaje się problematyczny, a wyobraź sobie zaporę sieciową o sześciu interfejsach obsługującą 140 serwerów. Widziałem firewalle Cisco, których listy dostępu zajmowały 17 stron wydruku. Wymienianie wszystkich kombinacji protokołów i serwerów może być niezwykle nużące. Złożoność list dostępu doprowadziła do tego, że wiele firm odrzucało propozycję zastosowania sprzętu Cisco w charakterze zapór sieciowych. Dzięki wprowadzeniu polecenia `object-group` administrator zyskał możliwość definiowania grup protokołów, sieci, typów komunikatów ICMP oraz usług i reprezentowania ich za pomocą wspólnej nazwy.

```
ASA-5540/GAD(config)# object-group ?

configure mode commands/options:
  icmp-type  Specifies a group of ICMP types, such as echo
                                                //Określa grupę typów komunikatów ICMP, na przykład echo
  network    Specifies a group of host or subnet IP addresses
                                                //Określa grupę adresów IP stacji lub podsieci
  protocol   Specifies a group of protocols, such as TCP, etc //Określa grupę protokołów, na przykład TCP itp.
  service    Specifies a group of TCP/UDP ports/services
                                                //Określa grupę portów (usług) protokołów TCP i UDP
```

Każdy z przedstawionych w przykładzie serwerów korzysta z tych samych portów TCP. Utworzenie grupy takich protokołów może znacznie skrócić definicję listy dostępu. Utwórzmy więc grupę obiektów o nazwie `Webserver-svcs`, która będzie grupą usług TCP. Do jej utworzenia trzeba wykorzystać instrukcje `port-object` uzupełniające polecenie `object-group`:

```
object-group service Webserver-svcs tcp
  description Obsluga serwerow WWW
  port-object eq smtp
  port-object eq www
  port-object eq https
```

Po tej operacji zamiast wymieniać wszystkie usługi powiązane z każdym serwerem, możemy określić grupę usług danego serwera. Służy do tego słowo kluczowe `object-group`, po którym następuje nazwa grupy:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
access-list In permit tcp any host 192.168.1.101 object-group Webserver-svcs
access-list In permit tcp any host 192.168.1.201 object-group Webserver-svcs
access-list In permit tcp any host 192.168.1.228 object-group Webserver-svcs
```

Efektem wprowadzonych zmian jest zmniejszenie liczby wpisów na liście dostępu z dziewięciu do trzech. Proces można jednak kontynuować. Wszystkie wymienione adresy IP odnoszą się do urządzeń tego samego typu — opisują serwery WWW. Utwórzmy więc kolejną grupę obiektów (o nazwie `Webservers`). Tym razem jako definicję typu grupy trzeba wykorzystać słowo kluczowe `network` oraz polecenia `network-object`:

```
object-group network Webservers
  description Serwery WWW
  network-object host 192.168.1.101
  network-object host 192.168.1.201
  network-object host 192.168.1.228
```

Definicję listy można więc uprościć jeszcze bardziej:

```
access-list In permit tcp any object-group Webservers object-group Webserver-svcs
```

Lista dostępu (złożona początkowo z dziewięciu wierszy) została skrócona do jednego wpisu. Niemniej po wprowadzeniu polecenia `show access-list` na ekranie są wyświetlane wszystkie elementy składowe grup obiektów:

```
ASA-5540/GAD(config)# show access-list Inbound
access-list Inboun; 9 elements
access-list In line 1 extended permit tcp any object-group Webservers object-group
Webserver-svcs 0x6484a1ba
  access-list In line 1 extended permit tcp any host 192.168.1.101 eq smtp (hitcnt=0) 0x9ba9ede4
  access-list In line 1 extended permit tcp any host 192.168.1.101 eq www (hitcnt=0) 0x34db9472
  access-list In line 1 extended permit tcp any host 192.168.1.101 eq https (hitcnt=0) 0x821a0cfe
  access-list In line 1 extended permit tcp any host 192.168.1.201 eq smtp (hitcnt=0) 0xb52bad46
  access-list In line 1 extended permit tcp any host 192.168.1.201 eq www (hitcnt=0) 0x881944d5
  access-list In line 1 extended permit tcp any host 192.168.1.201 eq https (hitcnt=0) 0x4c466f6a
  access-list In line 1 extended permit tcp any host 192.168.1.228 eq smtp (hitcnt=0) 0x6c6c1ba9
  access-list In line 1 extended permit tcp any host 192.168.1.228 eq www (hitcnt=0) 0xfca8e834
  access-list In line 1 extended permit tcp any host 192.168.1.228 eq https (hitcnt=0) 0x7d7e2239
```

Zwróć uwagę na jednakową wartość numeru wiersza w każdym wpisie (`line 1`). Oznacza ona, że poszczególne wpisy są rozwinięciem pierwszego wiersza listy dostępu (a zarazem jedynego wiersza tej listy).

Grupa obiektów nie może zostać usunięta, jeśli jest w danej chwili wykorzystywana. Ewentualna próba usunięcia kończy się zgłoszeniem błędu:

```
ASA-5540/GAD(config-service)#no object-group service Webserver-svcs
Removing object-group (Webserver-svcs) not allowed;
                                //Usunięcie grupy obiektów (Webserver-svcs) jest niedozwolone;
object-group (Webserver-svcs), being used in access-list, would become empty
                                //Grupa obiektów (Webserver-svc) wykorzystywana na liście dostępu stałaby się pusta
```

Trzeba więc najpierw usunąć odwołania do danej grupy z listy dostępu albo usunąć również listę dostępu. Po wykonaniu tego zadania usunięcie grupy nie stanowi problemu:

```
ASA-5540/GAD(config)#clear configure access-list In
ASA-5540/GAD(config)#no object-group service Webserver-svcs
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Inspekcje

W zaporach sieciowych PIX do weryfikacji protokołów aplikacji służyła funkcja o nazwie **fixup**. W nowszych rozwiązaniach została ona zastąpiona mechanizmem inspekcji aplikacji, nazywanym krótko **inspekcją**. Inspekcje są znacznie trudniejsze w konfigurowaniu, ponieważ ich definiowanie przypomina implementację podobnych rozwiązań w systemie IOS. Są one bowiem elementami globalnej polityki usług. Początkowo mogą się więc wydawać dość trudne, choć wcale nie muszą takie być.

Inspekcje zapewniają przekazywanie danych bardziej skomplikowanych protokołów, takich jak FTP, który wymaga przesyłania kilku strumieni. Gwarantują również większe bezpieczeństwo przenoszonych informacji. Na przykład inspekcja protokołu SMTP pozwala na zdefiniowanie określonych instrukcji, które będą akceptowane przez moduł ASA podczas sesji SMTP.

Aby zademonstrować jedną z typowych inspekcji aplikacji, połączyłem się z serwerem poczty, używając klienta telnetu. Ruch był przekazywany przez urządzenie ASA, w którym nie działała inspekcja SMTP (uruchamiana domyślnie). Gdy wprowadziłem polecenie SMTP `EHLO` *`nazwa_serwera`*, otrzymałem informacje na temat implementacji serwera pocztowego:

```
[GAD@someserver GAD]$telnet email.serwer.net 25
Trying 10.10.10.10...
Connected to email.serwer.net.
Escape character is '^]'.
220 email.serwer.net ESMTP Postfix
EHLO serwer
250-email.serwer.net
250-PIPELINING
250-SIZE 10240000
250-ETRN
250 8BITMIME
```

Informacje te nie są potrzebne do przesyłania wiadomości e-mail, a mogą się okazać użyteczne dla włamywacza. Na przykład haker mógłby próbować pobrać wiadomości z serwera za pomocą polecenia `ETRN deque`. Inspekcja SMTP przechwytuje jednak instrukcje `ETRN` i eliminuje ryzyko.

Funkcja `ETRN` jest bardzo użytecznym mechanizmem protokołu SMTP, ponieważ umożliwia dostawcy usług internetowych kolejkowanie poczty w czasie, gdy serwer klienta jest niedostępny. Jeśli chcesz korzystać z instrukcji `ETRN`, musisz wyłączyć inspekcję SMTP w firewallu.

Urządzenia ASA są dostarczane z domyślnie włączoną większością często wykorzystywanych inspekcji. Kolejny przykład przedstawia sposób uruchomienia inspekcji SMTP, co jednak najczęściej nie jest konieczne.

```
ASA-5540/GAD(config)#policy-map global_policy
ASA-5540/GAD(config-pmap)#class inspection_default
ASA-5540/GAD(config-pmap-c)#inspect esmtp
```

Po tej operacji zapora ASA przechwytuje i przetwarza wszystkie żądania SMTP:

```
[GAD@someserver GAD]$telnet mail.myserver.net 25
Trying 10.10.10.10...
Connected to email.serwer.net.
Escape character is '^]'.
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
220 *************************
EHLO serwer
502 Error: command not implemented
```

Porównaj zaznaczone fragmenty tego i poprzedniego listingu. Przy wyłączonej inspekcji SMTP do klienta telnetu została dostarczona odpowiedź serwera uwzględniająca nazwę, obsługiwaną wersję protokołu SMTP oraz rodzaj wykorzystywanego oprogramowania transferu poczty (MTA — *Mail Transfer Agent*).

```
220 email.serwer.net ESMTP Postfix.
```

Po włączeniu inspekcji SMTP firewall przechwycił odpowiedź i zmienił ją w bezużyteczny ciąg znaków:

```
220 *************************
```

Tym samym dostarczył ewentualnym hakerom znacznie mniej informacji. W ten sam sposób inspekcja uniemożliwia wykonanie instrukcji `EHLO` *`nazwa_serwera`*.

W różnych wersjach systemów ASA różne rodzaje inspekcji są włączane domyślnie. W wersji 8.0(4) standardowa konfiguracja to:

```
policy-map global_policy
 class inspection_default
  inspect dns preset_dns_map
  inspect ftp
  inspect h323 h225
  inspect h323 ras
  inspect netbios
  inspect rsh
  inspect rtsp
  inspect skinny
  inspect esmtp
  inspect sqlnet
  inspect sunrpc
  inspect tftp
  inspect sip
  inspect xdmcp
  inspect http
```

Część z wymienionych inspekcji bywa niepotrzebna i można je wyłączyć, choć włączone nie stanowią dodatkowego obciążenia urządzenia. Aby sprawdzić, które z nich są uaktywnione w danym systemie ASA, wystarczy wykonać instrukcję `show service-policy global`:

```
ASA-5540/GAD#sho service-policy global

Global policy:
  Service-policy: global_policy
    Class-map: inspection_default
    Inspect: dns preset_dns_map, packet 0, drop 0, reset-drop 0
    Inspect: ftp, packet 0, drop 0, reset-drop 0
    Inspect: h323 h225 _default_h323_map, packet 0, drop 0, reset-drop 0
    Inspect: h323 ras _default_h323_map, packet 0, drop 0, reset-drop 0
    Inspect: netbios, packet 0, drop 0, reset-drop 0
    Inspect: rsh, packet 0, drop 0, reset-drop 0
    Inspect: rtsp, packet 0, drop 0, reset-drop 0
    Inspect: skinny , packet 0, drop 0, reset-drop 0
    Inspect: esmtp _default_esmtp_map, packet 0, drop 0, reset-drop 0
    Inspect: sqlnet, packet 0, drop 0, reset-drop 0
    Inspect: sunrpc, packet 0, drop 0, reset-drop 0
    Inspect: tftp, packet 0, drop 0, reset-drop 0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Inspect: sip , packet 0, drop 0, reset-drop 0
Inspect: xdmcp, packet 0, drop 0, reset-drop 0
Inspect: http, packet 0, drop 0, reset-drop 0
```

Każda inspekcja jest powiązana z określonym protokołem. Szczegółowe informacje na temat ich działania są dostępne w dokumentacji opracowanej przez firmę Cisco.

Zarządzanie kontekstami

Tryb wielokontekstowy pozwala na utrzymywanie wielu wirtualnych zapór sieciowych w ramach pojedynczego firewalla ASA lub pary takich firewalli. Możliwa jest bowiem praca w tym trybie również wówczas, gdy urządzenia są przygotowane do przełączenia awaryjnego (ang. *failover*), także w konfiguracji przełączenia awaryjnego typu aktywny-aktywny. Przełączanie awaryjne aktywny-aktywny (ang. *active/active failover*) zostało opisane w podrozdziale „Przełączanie awaryjne" na stronie 475. Na rysunku 28.3 przedstawiono graficznie sposób definiowania wielu kontekstów w ramach jednego firewalla fizycznego.

Rysunek 28.3. Wiele kontekstów w fizycznej zaporze sieciowej

Praca w każdym kontekście jest zbliżona do działania niezależnego urządzenia, choć wymaga uwzględnienia szczególnego sposobu współdzielenia interfejsów. Każdy z kontekstów ma własny schemat adresowania IP i nic nie stoi na przeszkodzie, aby między kolejnymi kontekstami były powielane te same sieci.

W ramach każdego kontekstu dostępna jest większość standardowych funkcji firewalli ASA. Są jednak pewne ograniczenia. Wiadomo, że urządzenia ASA realizują wiele zadań, w tym VPN, IDS itp. Niestety większość z dodatkowych mechanizmów nie może być wykorzystywana w trybie wielokontekstowym. Przypominam, że następujące funkcje nie są w tym trybie dostępne:

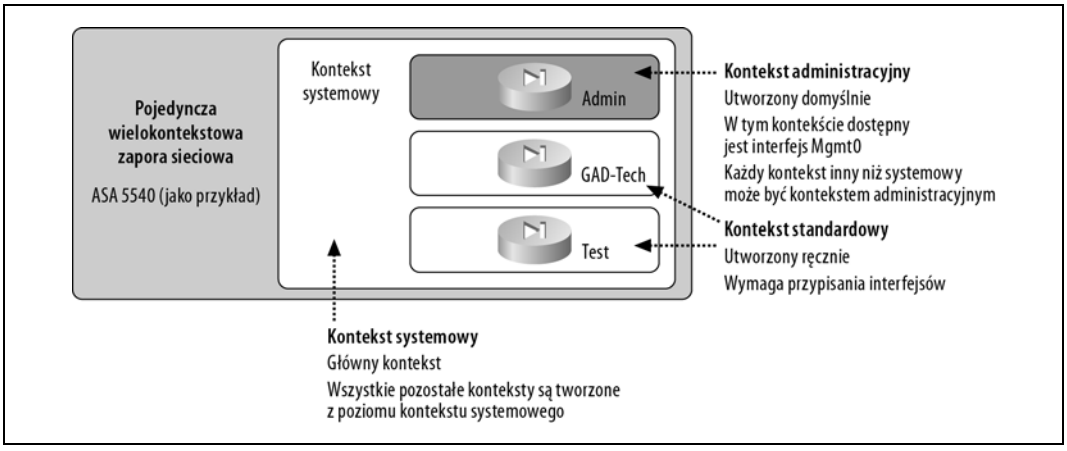
- OSPF, RIP, EIGRP (w trybie wielokontekstowym można korzystać jedynie z routingu statycznego);
- VPN;
- routing w multiemisji (obsługa mostowania w multiemisji jest zachowana);
- detekcja zagrożeń;
- QoS;
- proxy telefonii.

Lista jest dość długa i obejmuje istotne elementy! Dla wielu firm niemożność wykorzystania połączeń VPN czyni konteksty bezużytecznymi. Wyłączenie mechanizmu QoS i pośredniczenia w ruchu telefonicznym również bywa postrzegane jako znaczące ograniczenie. Tryb kontekstowy sprawdza się więc przede wszystkim wówczas, gdy zastępuje wiele tradycyjnych zapór sieciowych PIX. Korzystałem z niego w sieciach z serwisami WWW oraz w centrach danych, w których każdy kontekst obsługiwał ruch jednego klienta. W takich rozwiązaniach mechanizm ten sprawdza się doskonale.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Rodzaje kontekstów

W ramach wielokontekstowych systemów ASA dostępne są trzy rodzaje kontekstów: kontekst systemowy, kontekst administratora i kontekst, który nazywam **standardowym** lub po prostu **kontekstem**. Każdy z nich ma pewne funkcje specjalne i pewne ograniczenia opisane na rysunku 28.4.



Rysunek 28.4. Rodzaje kontekstów urządzenia ASA

Kontekst systemowy

Kontekst systemowy jest tym, w którego ramach definiowane i obsługiwane są inne konteksty. Można go postrzegać jako kontekst główny. Nie jest on wirtualnym firewallem (jak inne konteksty), ponieważ nie można w nim używać interfejsów i reguł zabezpieczeń. Wyjątkiem są interfejsy wykorzystywane do przełączania awaryjnego.

W urządzeniu ASA pracującym w trybie wielokontekstowym szybkość i tryb dupleksu interfejsów można definiować **tylko** w kontekście systemowym. Rozwiązanie to może doprowadzić do furii osobę, która po raz pierwszy korzysta z kontekstów. Osobiście przeklinałem Cisco za wyprodukowanie urządzenia, które najwyraźniej nie obsługuje szybkości 10 Mb/s i 100 Mb/s, nieświadomy tego, że szybkość **można** zmieniać, ale w innym kontekście.

Zmiana parametrów interfejsów jest wykonywana w ramach kontekstu systemowego, ponieważ interfejsy fizyczne są współdzielone przez wszystkie inne konteksty. Klient A nie może mieć przecież możliwości zmiany trybu dupleksu na interfejsie współdzielonym z klientem B.

Kontekst administracyjny

Kontekst administracyjny pełni funkcję podobną do innych kontekstów, ale pozwala również na przełączenie się do kontekstu systemowego lub innego kontekstu. System umożliwia wyznaczenie tylko jednego kontekstu jako administracyjnego, choć może nim być dowolny kontekst. Wykonuje się w nim zadania specjalne, takie jak uaktualnianie kodu.

Kontekst

Zgodnie z wcześniejszymi informacjami kontekst jest wirtualną zaporą sieciową. Jego działanie nie odbiega szczególnie od innych rodzajów kontekstów. Pracując w jego ramach, nie zauważysz niczego, co by sugerowało, że korzystasz z wirtualnego urządzenia. Wyjąt-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

kiem jest brak niektórych poleceń związanych z konfiguracją przełączania awaryjnego oraz szybkości i trybu dupleksu interfejsów. Przejście do innych kontekstów może się okazać niemożliwe, jeśli nie zostało poprzedzone przełączeniem do kontekstu systemowego lub administracyjnego. W ramach kontekstu (ale nie między kontekstami) można również wyzwalać awaryjne przełączanie, choć dostępność tej funkcji zależy od globalnej konfiguracji mechanizmu przełączania awaryjnego.

Klasyfikator

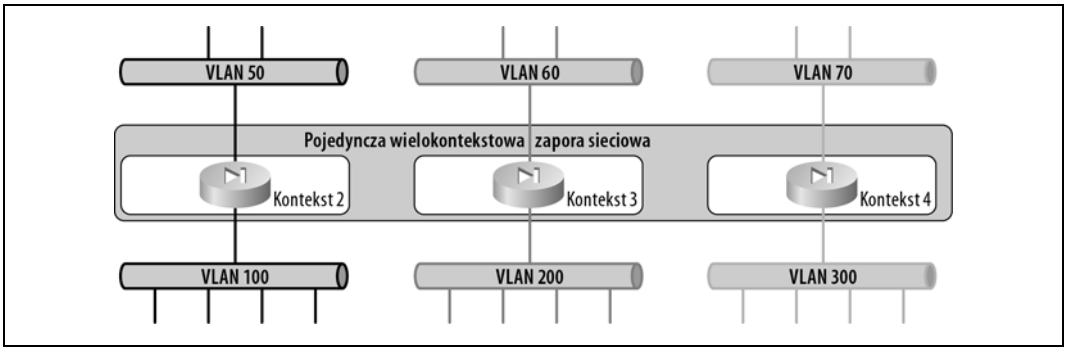
Skoro wiele kontekstów współdzieli jedno łącze fizyczne, musi istnieć jakiś mechanizm, który decyduje o dostarczaniu poszczególnych pakietów do określonych kontekstów. Zadanie to należy do procesu nazywanego **klasyfikatorem** (ang. *classifier*).

Klasyfikator jest istotnym elementem wielokontekstowego urządzenia ASA. Każdy administrator musi znać zasady jego działania oraz ograniczenia w dostarczaniu pakietów. Nieznajomość zasad pracy klasyfikatora niesie ryzyko zawału serca, depresji, a co najważniejsze, niewłaściwego działania urządzenia.

Różne sposoby tworzenia kontekstów (różne ich ograniczenia i wymagania konfiguracyjne) wynikają właśnie z mnogości trybów pracy klasyfikatora.

Brak współdzielonych interfejsów

Najmniej skomplikowana konfiguracja urządzenia ASA została przedstawiona na rysunku 28.5. Widać na nim trzy konteksty. Każdy dysponuje oddzielnym interfejsem wewnętrznym i zewnętrznym. Każdy interfejs jest przypisany tylko do jednego kontekstu. W prezentowanym rozwiązaniu każdy z interfejsów może być podinterfejsem opisanym za pomocą znacznika sieci VLAN (nie należy do współdzielonego interfejsu). Gdy każdy kontekst dysponuje własnym interfejsem, niemożliwa jest interakcja pomiędzy kontekstami. Rozwiązanie to w pełni odpowiada konfiguracji, w której każdy kontekst jest niezależną zaporą sieciową.

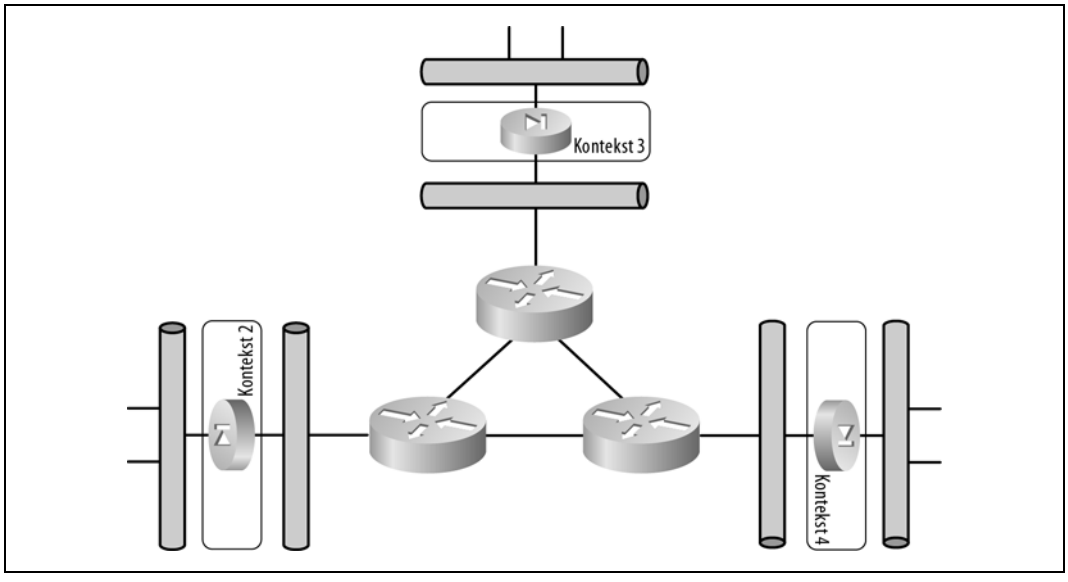


Rysunek 28.5. Wiele kontekstów bez współdzielonych interfejsów

Jest to najprostszy sposób konfiguracji firewalla wielokontekstowego, w którym nie trzeba podejmować żadnych specjalnych działań związanych z wykorzystaniem interfejsów.

Inna technika została przedstawiona na rysunku 28.6. Ponownie nie ma tutaj współdzielenia interfejsów lub sieci pomiędzy kontekstami.

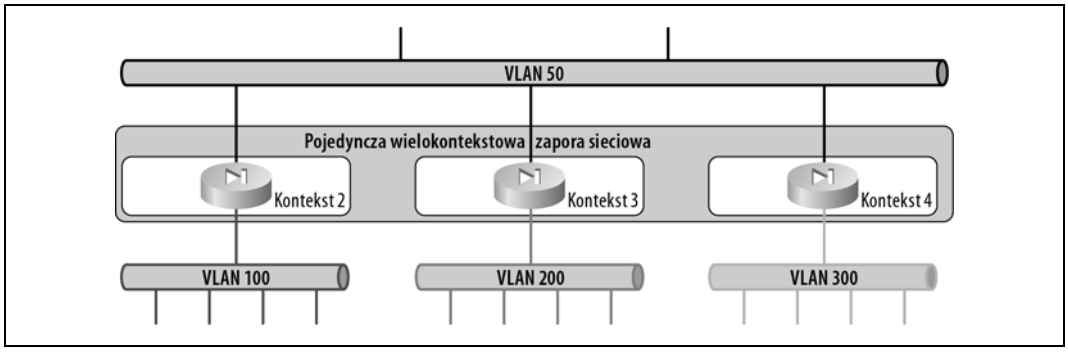
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 28.6. Kolejny układ wielokontekstowy bez współdzielonego interfejsu

Współdzielenie interfejsu zewnętrznego

W urządzeniach wielokontekstowych możliwe jest współdzielenie interfejsów przez większą liczbę kontekstów. Sytuację taką przedstawia rysunek 28.7. Widać na nim, że interfejs G0/0 jest przypisany do wszystkich trzech kontekstów, z których każdy ma adres IP pochodzący z tej samej sieci i korzysta z tej samej bramy domyślnej.



Rysunek 28.7. Wiele kontekstów współdzielących zewnętrzny interfejs

Przedstawione rozwiązanie jest znaczenie bardziej rozpowszechnione wśród firm ISP, które uruchamiają wiele zapór sieciowych przyłączonych do wspólnej zewnętrznej sieci IP (powiązanej z jednym interfejsem VLAN). Zasada działania klasyfikatora nie jest w tym przypadku szczególnie skomplikowana. Brama domyślna jest bowiem zazwyczaj osiągalna przez interfejs zewnętrzny. Pakiety o nieznanych adresach docelowych są więc przekazywane do bramy domyślnej, która jest jednakowa we wszystkich kontekstach.

Pakiety przychodzące są dostarczane przez bramę domyślną do kolejnego routera na trasie. W tym przypadku wszystkie routery mają ten sam adres MAC (domyślnie). Klasyfikator

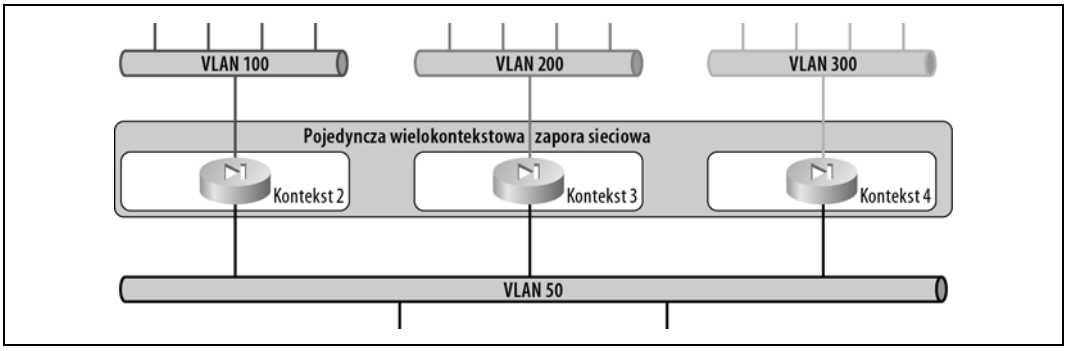
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

wybiera więc odpowiedni kontekst na podstawie docelowego adresu IP. Docelowy adres pakietu będzie odpowiadał adresowi mechanizmu NAT zdefiniowanemu w ramach kontekstu lub będzie zapisany (pośrednio) w tablicy routingu danego kontekstu.

Korzystniejszym rozwiązaniem w przedstawionym przykładzie jest przypisanie każdemu kontekstowi oddzielnego adresu MAC. Dzięki temu zewnętrzny router (brama domyślna) może dostarczać pakiety bezpośrednio do odpowiedniego kontekstu, bez angażowania klasyfikatora.

Współdzielenie interfejsu wewnętrznego

Inny sposób współdzielenia interfejsów został przedstawiony na rysunku 28.8. Jest to rozwiązanie odwrotne do przedstawionego w poprzednim przykładzie (na rysunku 28.7). W tym przypadku interfejsy zewnętrzne są niezależne, natomiast kilka interfejsów wewnętrznych współdzieli jeden interfejs fizyczny.



Rysunek 28.8. Wiele kontekstów współdzielących jeden interfejs wewnętrzny

Mimo że z technicznego punktu widzenia taka konfiguracja jest możliwa, nie należy jej implementować, szczególnie jeśli zapora sieciowa jest przyłączona do internetu. Przyczyną jest zasada działania klasyfikatora i sposób dostarczania pakietów do poszczególnych kontekstów. Ponieważ brama domyślna firewalla pracuje zazwyczaj po stronie interfejsu zewnętrznego, docelowy adres IP pakietu może być dowolnym adresem spoza wykorzystywanych po stronie wewnętrznej. Klasyfikator nie może ustalić, która brama domyślna powinna zostać wybrana, ponieważ w żadnym kontekście nie istnieje wpis w tablicy routingu odpowiadający docelowemu adresowi IP. Nie ma również żadnych wpisów w tablicy NAT. W takim przypadku klasyfikator może odrzucić pakiet. Z tego powodu współdzielenie interfejsów wewnętrznych nie jest zalecane, poza nielicznymi przypadkami, w których zewnętrzne adresy podlegają translacji NAT.

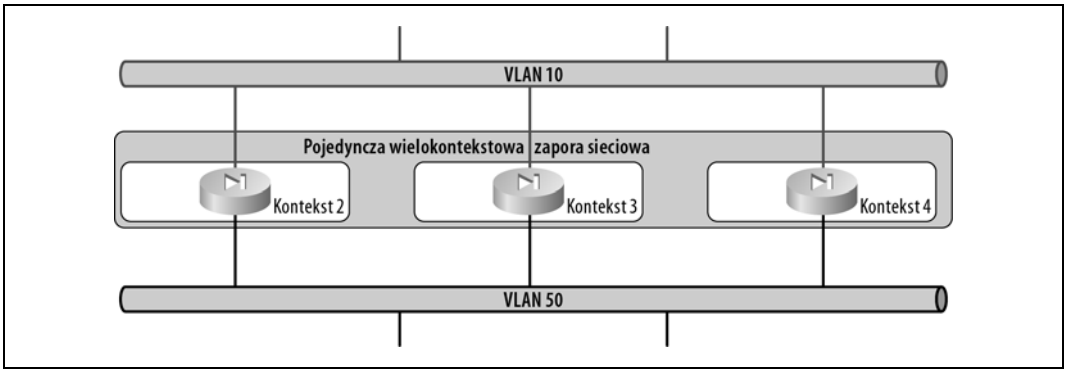
Współdzielenie zewnętrznego i wewnętrznego interfejsu

Teoretycznie współdzielenie interfejsu wewnętrznego i zewnętrznego jest możliwe (zgodnie z rysunkiem 28.9). W praktyce jednak nie jest zalecane, ponieważ doprowadza do szału zarówno klasyfikator, jak i administratora.

Konteksty kaskadowe

Konteksty można łączyć **kaskadowo**. Kaskadowe konfigurowanie kontekstów polega na tym, że jeden z kontekstów jest umiejscawiany przed lub za innym. W poprzednich przykładach przedstawione zostały techniki współdzielenia interfejsów. Jednak we wszystkich prezentowanych

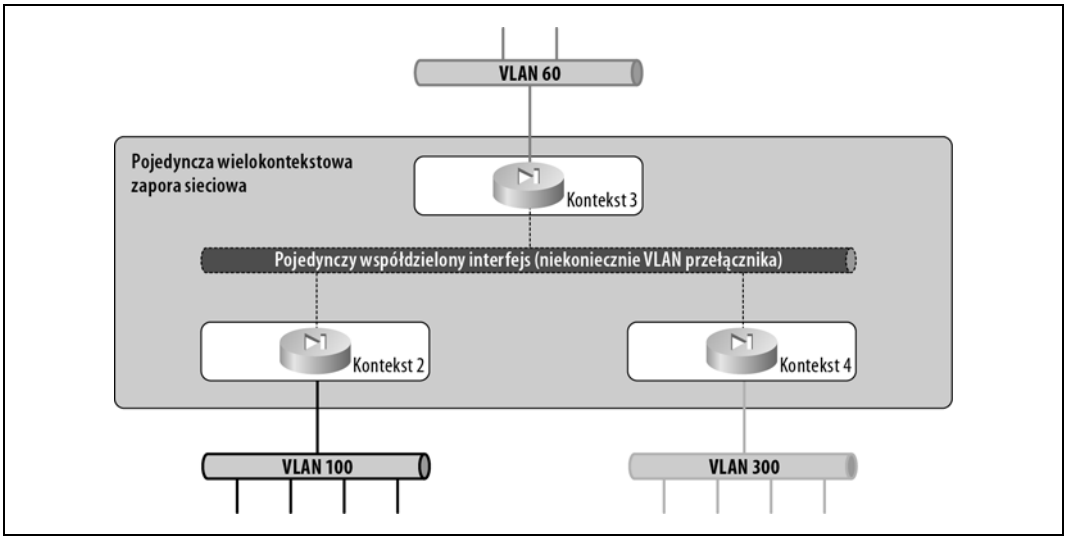
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 28.9. Wiele kontekstów współdzielących interfejs wewnętrzny i interfejs zewnętrzny

rozwiązaniach interfejsy znajdowały się po tej samej stronie wszystkich kontekstów. Przykładem takiej konfiguracji było współdzielenie interfejsu zewnętrznego w sieci ISP połączonej z publiczną siecią IP.

Dzięki łączeniu kaskadowemu możesz powiązać wewnętrzny interfejs jednego kontekstu z interfejsem zewnętrznym kolejnego kontekstu (zgodnie z rysunkiem 28.10).



Rysunek 28.10. Konteksty kaskadowe

Takie rozwiązanie bywa użyteczne w przypadku, w którym pojedynczy firewall zarządzany przez dostawcę usług internetowych musi chronić sieć pomimo tego, że poszczególni klienci zarządzają własnymi zaporami sieciowymi. Konfigurując taką sieć, pamiętaj, że prawdopodobnie konieczne będzie powielenie odpowiednich reguł filtrowania ruchu na każdym poziomie hierarchii.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Konfiguracja kontekstów

Zapory sieciowe obsługujące konteksty mogą pracować w dwóch trybach: pojedynczego kontekstu (`single`) i wielu kontekstów (`multi`). Tryb pojedynczego kontekstu udostępnia administratorowi wszystkie funkcje, których się spodziewamy w urządzeniu ASA. Tryb wielokontekstowy umożliwia definiowanie niezależnych firewalli wirtualnych kosztem wszystkich najciekawszych funkcji zapory sieciowej (takich jak routing i połączenia VPN). Aby sprawdzić, w jakim trybie urządzenie pracuje w danej chwili, należy wpisać polecenie `show mode`. Oto przykład wykonania tej instrukcji w zaporze sieciowej ASA 5540 pracującej w trybie pojedynczego kontekstu:

```
ASA-5540#sho mode
Security context mode: single  //Tryb kontekstu zabezpieczeń: pojedynczy
```

Rozpoczynamy pracę od włączenia trybu wielokontekstowego, do czego wykorzystamy polecenie `mode`. Konsekwencje jego użycia są dość poważne, ponieważ zmienia ono istotnie sposób działania urządzenia. System ASA wyświetla nawet stosowne ostrzeżenie:

```
ASA-5540(config)#mode multiple
WARNING: This command will change the behavior of the device  //Uwaga: To polecenie zmieni sposób pracy
                                                              //urządzenia
WARNING: This command will initiate a Reboot  //Uwaga: To polecenie spowoduje ponowne uruchomienie systemu
Proceed with change mode? [confirm]  //Czy kontynuować zmianę trybu pracy? [potwierdzam]
```

Po potwierdzeniu zostaniesz zapytany, czy chcesz dokonać konwersji konfiguracji systemu. Zaakceptowanie tej propozycji powoduje ponowne uruchomienie systemu i zainicjowanie procesu.

```
Convert the system configuration? [confirm]  //Skonwertować konfigurację systemu? [potwierdzam]
!
The old running configuration file will be written to flash
                                 //Poprzedni plik konfiguracji bieżącej zostanie zapisany w pamięci flash
The admin context configuration will be written to flash
                                 //Konfiguracja kontekstu administracyjnego zostanie zapisana w pamięci flash
The new running configuration file was written to flash
                                 //Nowy plik konfiguracji bieżącej został zapisany w pamięci flash
Security context mode: multiple  //Tryb kontekstu zabezpieczeń: wielokrotny
***
*** --- SHUTDOWN NOW ---         //Zamykanie systemu
***
*** Message to all terminals:    //Komunikat do wszystkich terminali:
***
*** change mode                  //zmiana trybu
Rebooting....                    //Ponowne uruchamianie…
```

Po ponownym uruchomieniu systemu ASA na ekranie może zostać wyświetlonych wiele komunikatów o błędach. Jeśli wcześniej urządzenie miało zdefiniowaną pewną konfigurację, wiele z jej instrukcji prawdopodobnie nie zostanie wykonanych poprawnie (szczególnie jeśli odnosiły się do nieobsługiwanych funkcji, takich jak VPN, QoS lub routing). Nie uda się również wykonać poleceń odnoszących się do interfejsów, ponieważ interfejsy te nie są jeszcze przypisane do poszczególnych kontekstów. Niemniej wpisanie polecenia `show mode` pozwala upewnić się, że tryb pracy został zmieniony:

```
ASA-5540#sho mode
Security context mode: multiple
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Teraz pracujemy w kontekście systemowym. Możemy więc zarządzać innymi kontekstami. W tym trybie możliwe jest również definiowanie mechanizmu przełączania awaryjnego (opisanego w kolejnym podrozdziale). Po ponownym uruchomieniu urządzenia udostępniany jest jeden wstępnie skonfigurowany kontekst:

```
ASA-5540#sho context
Context Name     Class      Interfaces          URL
*admin           default    Management0/0       disk0:/admin.cfg
Total active Security Contexts:
```

Z kontekstem administracyjnym powiązana jest następująca konfiguracja:

```
admin-context admin
context admin
  allocate-interface Management0/0
  config-url disk0:/admin.cfg
```

Utworzenie nowego kontekstu jest bardzo łatwe i sprowadza się do wykonania polecenia `context`. Nie oznacza to jednak, że nowy kontekst będzie od razu dostępny:

```
ASA-5540(config)#context GAD-Tech
Creating context 'GAD-Tech'... Done. (4)   //Tworzenie kontekstu 'GAT-Tech'… Zrobione. (4)
```

Zauważ, że zmienił się tryb pracy na tryb konfiguracji kontekstu:

```
ASA-5540(config-ctx)#
```

Następna czynność polega na przypisaniu kontekstowi wybranych interfejsów:

```
ASA-5540(config-ctx)#allocate-interface g0/0
ASA-5540(config-ctx)#allocate-interface g0/1
```

Operacja ta nie jest ograniczona do interfejsów fizycznych. W ten sam sposób można przypisać do kontekstów określone podinterfejsy. Wystarczy określić identyfikator podinterfejsu:

```
ASA-5540(config-ctx)#allocate-interface g0/1.100
```

Nie możemy jeszcze zmienić kontekstu, o czym za każdym razem przypomina mi następujący komunikat o błędzie:

```
ASA-5540#changeto context GAD-Tech
ERROR: Context hasn't been initialized with 'config-url' //BŁĄD: W kontekście nie został zdefiniowany
parametr 'config-url'
```

Musimy osobiście określić plik, w którym będzie przechowywana konfiguracja danego kontekstu:

```
ASA-5540(config)#context GAD-Tech
ASA-5540(config-ctx)#config-url disk0:/gad-tech.cfg
INFO: Context GAD-Tech was created with URL disk0:/gad-tech.cfg
                                   //INFO: Kontekst GAT-Tech został utworzony pod adresem URL disk0:/gad-tech.cfg
```

Aby wyświetlić informacje o stanie wszystkich kontekstów, trzeba w kontekście systemowym wpisać polecenie `show context`. W kolejnym przykładzie widać nowo utworzony kontekst oraz domyślny kontekst administracyjny:

```
ASA-5540#show context
Context Name     Class      Interfaces          URL
*admin           default    Management0/0       disk0:/admin.cfg
GAD-Tech         default    GigabitEthernet0/0, disk0:/gad-tech.cfg
                            GigabitEthernet0/1
Test             default    GigabitEthernet0/0, disk0:/test.cfg
                            GigabitEthernet0/2
Total active Security Contexts: 3
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Po dodaniu kontekstu `GAD-Tech` zdefiniowałem również kontekst `Test`, aby wykazać, że obydwa konteksty mogą współdzielić jeden interfejs (zgodnie z wcześniej omówionym schematem). W tym przypadku jest to interfejs G0/0.

Przełączmy się więc do kontekstu `GAD-Tech`. Służy do tego instrukcja `changeto`.

Niektóre wersje przełączników Nexus obsługują konteksty urządzeń wirtualnych (VDC — *Virtual Device Context*), które w działaniu są zbliżone do kontekstów urządzeń ASA. Aby zmylić użytkowników swoich produktów, firma Cisco postanowiła zaimplementować polecenie `changeto` do zmiany kontekstów w zaporach sieciowych ASA oraz instrukcję `switchto` do zmiany kontekstów VDC w przełącznikach Nexus. Jeśli pracujesz w centrum przetwarzania danych, które dysponuje obydwoma rodzajami urządzeń, z pewnością zauważysz, że ciągłe błędne wprowadzanie instrukcji na poszczególnych platformach istotnie wzbogaca zasób wulgaryzmów.

Zmieńmy więc kontekst pracy na `GAD-Tech`:

```
ASA-5540#changeto context GAD-Tech
```

Zauważ, że ciąg zgłoszenia zmienił się z *`nazwa_urządzenia#`* na *`nazwa_urządzenia/kontekst#`*. Treść zgłoszenia można konfigurować, ale osobiście uważam, że domyślny jest bardzo użyteczny. Następnie możemy sprawdzić, że wynik wykonania polecenia `show context` jest ograniczony do informacji właściwych dla określonego kontekstu:

```
ASA-5540/GAD-Tech#sho context
Context Name     Class      Interfaces          URL
GAD-Tech         default    GigabitEthernet0/0, disk0:/gad-tech.cfg
                            GigabitEthernet0/1
```

Aby przełączyć się z powrotem do kontekstu systemowego, wystarczy wpisać polecenie `changeto system`.

```
ASA-5540/GAD-Tech#changeto system
ASA-5540#
```

Jeśli korzystasz z mechanizmu VDC przełącznika Nexus, działanie systemu ASA może Ci się wydać irytujące. W systemie Nexus powrót do poprzedniego kontekstu następuje po wpisaniu instrukcji `switchback`. W urządzeniu ASA nie ma polecenia `changeback`.

Opuszczenie kontekstu bez użycia instrukcji `changeto system` powoduje przejście do kontekstu systemowego, ale poza tryb uprzywilejowany:

```
ASA-5540/GAD-Tech#exit

Logoff
Type help or '?' for a list of available commands.
ASA-5540>
```

Interfejsy i konteksty

Po przypisaniu interfejsu do kontekstu można go zobaczyć na zestawieniach danego kontekstu. Na przykład w prezentowanym urządzeniu ASA 5540 dostępne są cztery interfejsy: G0/0 – 3. Oto przykład konfiguracji, w której kontekstowi `GAT-Tech` zostały przypisane interfejsy G0/0, G0/1.100 oraz G0/2.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
ASA-5540#sho run context GAD-Tech
context GAD-Tech
  allocate-interface GigabitEthernet0/0
  allocate-interface GigabitEthernet0/1.100
  allocate-interface GigabitEthernet0/2
  config-url disk0:/gad-tech.cfg
```

Analizując rezultat wykonania polecenia w ramach kontekstu, można odnieść wrażenie, że w urządzeniu są zainstalowane tylko wymienione interfejsy:

```
ASA-5540/GAD-Tech#sho int
Interface GigabitEthernet0/0 "", is down, line protocol is down
        Available but not configured via nameif
Interface GigabitEthernet0/1.100 "", is down, line protocol is down
        Available but not configured via nameif
Interface GigabitEthernet0/2 "", is down, line protocol is down
        Available but not configured via nameif
```

Pracując w kontekście innym niż systemowy, nie można zmienić jakiegokolwiek parametru fizycznego. Oczywiście ciągle o tym zapominam i gapię się w ekran, jakbym właśnie zauważył oznaki działania Matriksa, ale później sobie przypominam, że zmiana szybkości transmisji lub trybu dupleksu wymaga przełączenia się do kontekstu systemowego. Nie chcielibyśmy przecież, aby jeden z klientów pracujących w swoim kontekście mógł zmieniać szybkość i tryb dupleksu interfejsu, przenosząc te ustawienia na wszystkie inne konteksty, nad którymi nie ma kontroli.

Gdy interfejsy są dostępne (w ramach kontekstu), można im przypisać adresy IP, nazwy, priorytety i wszystkie standardowe parametry.

Zachowanie polecenia write memory

Gdy wykonujesz polecenie `write memory`, pracując w danym kontekście, zapisujesz jedynie ustawienia związane z tym kontekstem.

Analogicznie wprowadzenie instrukcji `write mem` w kontekście systemowym powoduje zapisanie ustawień kontekstu systemowego. Nie towarzyszy temu żadna informacja o zakresie działania polecenia. Jest to jednak rozwiązanie spójne z działaniem instrukcji w innych kontekstach.

Polecenie `write mem` zapisuje jedynie ustawienia bieżącego kontekstu.

```
ASA-5540(config)#wri mem
Building configuration...
Cryptochecksum: 60a2296d 904d7bb1 f2e9284a 5576c1dd

1023 bytes copied in 3.320 secs (341 bytes/sec)
[OK]
```

Jedną z korzyści wynikających z pracy w ramach kontekstu systemowego jest to, że za pomocą polecenia `write mem all` możemy zapisać konfigurację wszystkich kontekstów.

```
ASA-5540(config)# wri mem all
Building configuration...
Saving context : system : (000/003 Contexts saved)
Cryptochecksum: 60a2296d 904d7bb1 f2e9284a 5576c1dd
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
1023 bytes copied in 3.320 secs (341 bytes/sec)
Saving context : admin : (001/003 Contexts saved)
Cryptochecksum: f69086ec aa68219f d3ef18ff 391fe121

1780 bytes copied in 0.260 secs
Saving context : GAD-Tech : (002/003 Contexts saved)
Cryptochecksum: 156bed5f 17788005 9fb9d182 231910ab

1503 bytes copied in 0.160 secs
Saving context : Test : (003/003 Contexts saved)
Cryptochecksum: f39b8a45 e1b47f79 22f4b4e6 95b0cc8d

1499 bytes copied in 0.220 secs
[OK]
```

Przełączanie awaryjne

Większość urządzeń ASA można połączyć w pary o wysokiej dostępności, o ile obydwa urządzenia z pary mają identyczną konfigurację sprzętową. Istnieją dwa sposoby zapewnienia awaryjnego przełączania modułów ASA: rozwiązanie typu aktywny-oczekujący lub aktywny-aktywny. W konfiguracji aktywny-oczekujący awaria jednego urządzenia powoduje automatyczne przejście jego funkcji przez firewall zapasowy. W rozwiązaniu aktywny-aktywny obydwa urządzenia ASA przekazują ruch, jednak nie jest to ustawienie tak korzystne, jak by się mogło początkowo wydawać.

Aby skorzystać z funkcji przełączania, trzeba wykupić odpowiednią licencję. Wszystkie moduły z serii 5520 i wyższych są w nią wyposażane automatycznie. Sprawdzenie, czy dana zapora sieciowa ma zaimplementowany mechanizm przełączania awaryjnego, sprowadza się do wykonania polecenia `show version`:

```
ASA-5540#sho version | begin Licensed
Licensed features for this platform:
Maximum Physical Interfaces : Unlimited
490 | Chapter 28: ASA Firewall Configuration
Maximum VLANs : 200
Inside Hosts : Unlimited
Failover : Active/Active
VPN-DES : Enabled
VPN-3DES-AES : Enabled
Security Contexts : 10
GTP/GPRS : Disabled
VPN Peers : 5000
WebVPN Peers : 2
AnyConnect for Mobile : Disabled
AnyConnect for Linksys phone : Disabled
Advanced Endpoint Assessment : Disabled
UC Proxy Sessions : 2
```

Aby dwa urządzenia ASA mogły stanowić parę przełączania awaryjnego, muszą mieć zainstalowaną tę samą wersję oprogramowania. Tolerowane są ewentualne drobne odstępstwa, które są sygnalizowane za pomocą następującego komunikatu:

```
************WARNING****WARNING****WARNING*******************************
Mate version 8.0(3) is not identical with ours 8.0(4)  //Wersja partnera 8.0(3) nie jest identyczna z naszą 8.0(4)
************WARNING****WARNING****WARNING*******************************
Beginning configuration replication: Sending to mate.  //Rozpoczynanie konfiguracji replikacji: Wysyłanie
                                                       //do partnera.
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Każde urządzenie pary awaryjnego przełączania musi również pracować z tą samą konfiguracją. Nawet nazwy urządzeń muszą być jednakowe. Ewentualna próba zmiany ustawień firewalla zapasowego powoduje wygenerowanie komunikatu ostrzegawczego z informacją o tym, że wszelkie wprowadzone zmiany nie zostaną zsynchronizowane:

```
ASA-5540#conf t
**** WARNING ***
Configuration Replication is NOT performed from Standby unit to Active unit.
                                        //Replikacja konfiguracji NIE jest wykonywana z jednostki zapasowej do aktywnej.
Configurations are no longer synchronized.  //Konfiguracje nie są zsynchronizowane.
```

Wprowadzanie zmian nie jest jednak zabronione. Sam wpatrywałem się w ten komunikat wielokrotnie, szczególnie po osiemnastu godzinach ciągłej pracy.

Terminologia przełączania awaryjnego

Jednostki wchodzące w skład pary przełączania awaryjnego są określane za pomocą specjalnych nazw, zależnych od pełnionej roli.

Primary

Słowo `primary` służy do określenia jednostki podstawowej w danej parze przełączania awaryjnego. Do ustawienia urządzenia ASA jako podstawowego służy polecenie `failover lan unit primary`. Zazwyczaj po zainicjowaniu mechanizmu jest ono również urządzeniem **aktywnym** (`active`). Po wyznaczeniu rola jednostki się nie zmienia. Taki sposób określania funkcji jest nazywany wyznaczaniem fizycznym (ang. *physical designation*).

Secondary

Słowo `secondary` definiuje zapasową jednostkę danej pary przełączania awaryjnego. Samo urządzenie zazwyczaj nie jest bezpośrednio konfigurowane, chyba że podstawowa jednostka ASA ulegnie awarii. Po zainicjowaniu mechanizmu jednostka zapasowa pozostaje w stanie **oczekiwania** (`standby`), który się nie zmienia. Jest to wyznaczenie fizyczne.

Aktywność

Określenia jednostka aktywna lub aktywny kontekst opisują urządzenie lub kontekst, które analizują pakiety, wykorzystując do tego celu systemowe adresy IP przypisane do poszczególnych interfejsów. Jest to wyznaczenie logiczne. Zarówno podstawowa jednostka ASA, jak i zapasowa mogą być urządzeniami aktywnymi lub zawierać aktywny kontekst.

Oczekiwanie

Określenia jednostka oczekująca lub kontekst oczekujący odnoszą się do urządzenia lub kontekstu, które nie przetwarzają pakietów, a na interfejsach wykorzystują adresy IP mechanizmu przełączania awaryjnego. Gdy aktywna zapora sieciowa lub jej kontekst przestaną działać, system zapasowy lub kontekst zapasowy przejmują zadanie i stają się elementami aktywnymi. Taki sposób wyznaczania roli jest nazywany wyznaczaniem logicznym. Jednostką oczekującą może być urządzenie podstawowe lub zapasowe. Analogicznie kontekst oczekujący może być zdefiniowany w jednostce podstawowej lub zapasowej.

W trybie aktywny-aktywny obydwa firewalle (podstawowy i zapasowy) mogą obejmować oczekujące konteksty.

Przełączanie awaryjne z zachowaniem stanu

Gdy aktywne urządzenia ASA lub kontekst ulegają awarii i ich rolę przejmują jednostki lub konteksty oczekujące, wszystkie połączenia (zgodnie z domyślnym ustawieniem) są

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

przerywane. Aby zapobiec utracie połączeń, między aktywnym i oczekującym urządzeniem ASA można zestawić specjalne połączenie ethernetowe, za którego pomocą będą przekazywane informacje o stanie każdego z realizowanych połączeń. Jednostki oczekujące lub konteksty oczekujące, które zostały skonfigurowane z myślą o zachowaniu stanu podczas przełączania awaryjnego (ang. *statefull failover*), są na bieżąco aktualizowane, dzięki czemu w razie awarii żadne połączenie nie zostanie przerwane.

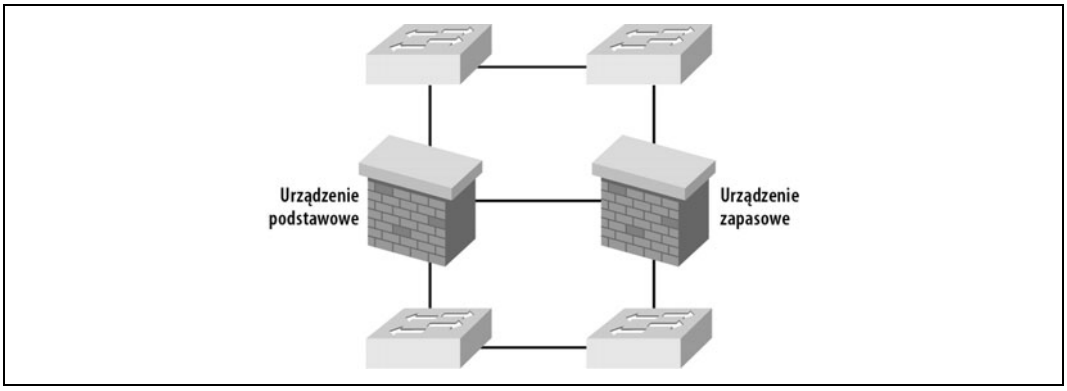
Podstawy przełączania awaryjnego

Podstawowe i zapasowe urządzenia ASA komunikują się za pośrednictwem specjalnie skonfigurowanego interfejsu. Jednostki ASA nie umożliwiają wykorzystania do tego celu kabla szeregowego, tak jak miało to miejsce w przypadku zapór sieciowych PIX.

Każde z urządzeń ASA monitoruje status mechanizmu przełączania awaryjnego, zasilania oraz interfejsu drugiej jednostki. Każdy z modułów za pomocą łącza awaryjnego oraz wszystkich aktywnych interfejsów w regularnych odstępach czasu wysyła pakiet **hello**. Jeśli jeden z dwóch modułów ASA nie odbierze pakietu hello w czasie dwóch kolejnych interwałów, dany interfejs jest przełączany w tryb **testowania**. Jeżeli przez kolejne dwa okresy ponawiania pakiet hello nie zostanie odebrany na interfejsie awaryjnym urządzenia oczekującego, jednostka ta inicjuje przełączenie awaryjne.

Konfiguracja platformy ASA jest bardzo elastyczna, więc nie będę opisywał wszystkich możliwych scenariuszy przełączania awaryjnego. Podstawowy mechanizm jest zawsze taki sam. Jeśli jedno urządzenie wykryje niedostępność drugiego, uznaje je za uszkodzone. W trybie wielokontekstowym obowiązują te same zasady, ale odnoszące się do pojedynczych kontekstów (z pewnymi ograniczeniami, które zostały opisane w dalszej części rozdziału).

Aby przełączanie awaryjne działało poprawnie, każde urządzenie ASA musi mieć możliwość komunikacji z drugą jednostką za pośrednictwem wszystkich interfejsów objętych działaniem tego mechanizmu. Zazwyczaj zapory są łączone z użyciem dwóch przełączników, z których każdy jest przyłączony do jednego urządzenia. Pomiędzy przełącznikami instalowane jest dodatkowe łącze (zazwyczaj trunkowe). Przykład takiej konfiguracji został przedstawiony na rysunku 28.11. Bezpośrednie połączenie między jednostkami ASA jest łączem przełączania awaryjnego (umożliwiającym przekazywanie informacji o stanie połączeń), które nie powinno być realizowane z użyciem przełącznika.



Rysunek 28.11. Typowa konfiguracja mechanizmu przełączania awaryjnego w urządzeniach ASA

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Firma Cisco zaleca, aby porty przełączników wykorzystane do połączenia zapór sieciowych ASA były ustawiane jako `spanning-tree portfast`. Standardowy algorytm drzewa rozpinającego uniemożliwia bowiem wymienianie pakietów hello w czasie krótszym niż trzy interwały ponawiania tych pakietów (w czasie inicjowania drzewa rozpinającego). To z kolei może doprowadzić do uruchomienia procedury przełączania awaryjnego w jednym z urządzeń ASA.

Jednostki ASA można uruchamiać w **trybie transparentnym** (ang. *transparent mode*) (standardowe działanie jest określane jako praca w **trybie routingu** (ang. *routed mode*)). Działanie firewalli odpowiada wówczas funkcjom mostów, które mają nieco inny wpływ na sposób budowania drzewa rozpinającego. Szczegółowych informacji na temat wykorzystania urządzeń ASA w trybie transparentnym musisz poszukać w dokumentacji firmy Cisco, a najlepiej w ogóle nie korzystaj z takiego trybu pracy.

W czasie przełączania awaryjnego oczekujące urządzenie ASA przejmuje adresy IP i MAC wszystkich interfejsów jednostki aktywnej, które zostały objęte tym mechanizmem. Cała operacja pozostaje niezauważona dla innych stacji sieciowych. Na poniższym listingu widać różnice w wartościach adresów MAC łącza G0/0 sprzed i po operacji awaryjnego przełączania:

```
ASA-5540/GAD#sho int g0/0 | inc MAC
        MAC address 0200.0000.1300, MTU 1500
ASA-5540/GAD#Failover LAN Failed
ASA-5540/GAD#sho int g0/0 | inc MAC
        MAC address 1200.0000.0c00, MTU 1500
```

Przełączanie awaryjne w urządzeniach ASA działa tak dobrze, że uszkodzenia jednostek nie są zauważane przez stacje sieciowe. Dlatego konieczne jest nadzorowanie pracy jednostek ASA i rozwiązywanie zaistniałych problemów. Sam widziałem przypadki, w których podstawowe urządzenie ASA uległo awarii i przez wiele miesięcy było zastępowane przez jednostkę zapasową, aż do czasu, gdy ona też zawiodła. Wyłączenie zapasowego urządzenia (niezabezpieczonego kolejnym urządzeniem) powoduje całkowity brak dostępu do sieci. Poza tym bez monitorowania stanu firewalli sieć nie jest tak bezpieczna, jak by się mogło wydawać. Użycie protokołu SMTP i oprogramowania zarządczego, takiego jak CiscoWorks lub OpenView, pozwala na bieżące reagowanie na zdarzenia przełączania awaryjnego między jednostkami ASA.

Konfiguracja przełączania awaryjnego — wariant aktywny-oczekujący

Obydwa urządzenia pary przełączania awaryjnego muszą korzystać z tej samej wersji systemu operacyjnego (w przeciwnym przypadku nie zsynchronizują swoich konfiguracji). W razie jakichkolwiek odstępstw skutki operacji mogą być nieprzewidywalne. Jednakowa musi być również konfiguracja sprzętowa. Co prawda mechanizm przełączania może działać poprawnie także w przypadkach pewnych nieznacznych różnic między jednostkami, ale zaleca się zachowanie pełnej zgodności urządzeń. Sam system generuje komunikaty ostrzegawcze po wykryciu różnic pomiędzy jednostkami:

```
************WARNING****WARNING****WARNING*******************************
Mate version 8.0(4) is not identical with ours 8.0(3)  //Wersja partnera 8.0(4) nie jest identyczna z naszą 8.0(3)
************WARNING****WARNING****WARNING*******************************
```

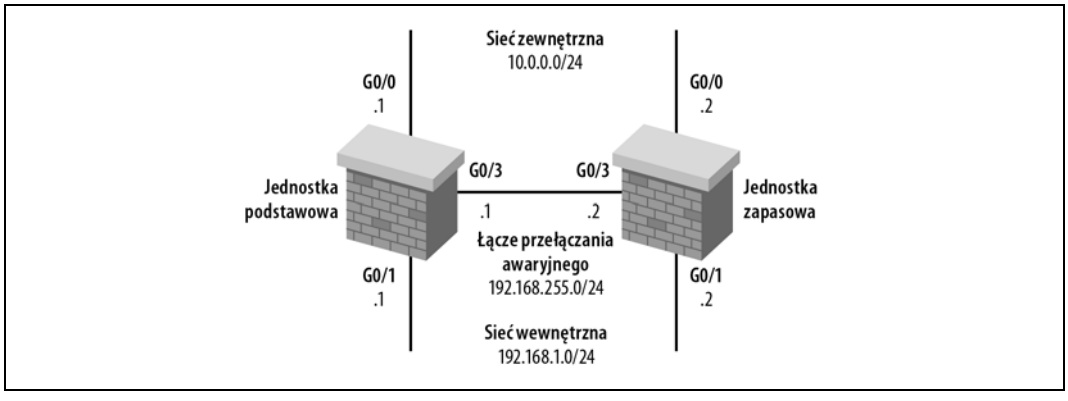

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Pierwsza czynność konfiguracyjna polega na włączeniu mechanizmu za pomocą polecenia `failover`:

```
ASA-5540(config)#failover
```

Każdy interfejs, który powinien zostać uwzględniony w działaniu mechanizmu, musi mieć odrębny adres IP przeznaczony na potrzeby awaryjnego przełączania. Z tego względu warto pamiętać o przydzielaniu par adresów zaporom sieciowym podczas projektowania sieci, nawet jeśli instalowane jest tylko jedno urządzenie. Na przykład jeśli wyznaczasz jako wewnętrzny adres firewalla adres 192.168.1.1, zarezerwuj również adres 192.168.1.2. Dzięki temu rozbudowa sieci o dodatkowe urządzenie ASA będzie mniej kłopotliwa. Adres mechanizmu przełączania awaryjnego będzie wstępnie zarezerwowany.

Analizę konfiguracji mechanizmu przełączania awaryjnego przeprowadzimy na parze urządzeń przedstawionych na rysunku 28.12. Jedyną różnicą jest to, że połączenie między urządzeniami bazuje na przełącznikach, a nie na kablu z przeplotem.



Rysunek 28.12. Przykładowa konfiguracja mechanizmu przełączania awaryjnego

Konfiguracja podstawowej jednostki ASA jest wówczas następująca:

```
interface GigabitEthernet0/0
 nameif outside
 security-level 0
 ip address 10.0.0.1 255.255.255.0 standby 10.0.0.2
!
interface GigabitEthernet0/1
 nameif inside
 security-level 100
 ip address 192.168.1.1 255.255.255.0 standby 192.168.1.2
!
interface GigabitEthernet0/3
 description LAN/STATE Failover Interface!    //Interfejs przełączania awaryjnego z podtrzymaniem stanu!
```

Wykorzystane w tym przykładzie zapory sieciowe to urządzenia ASA 5540 z czterema interfejsami Gigabit Ethernet (w każdej). Interfejsy G0/0 i G0/1 pełnią rolę interfejsów zewnętrznych (`outside`) i wewnętrznych (`inside`). Z kolei interfejs G0/3 wyznacza łącze przełączania awaryjnego z podtrzymaniem stanu. Opis interfejsu jest dodawany automatycznie przez system.

Każdy interfejs musi otrzymać adres IP przełączania awaryjnego odpowiadający urządzeniu zapasowemu. Przedstawione polecenia odnoszą się do jednostki podstawowej. Adresy IP zapisane

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

za słowem kluczowym `standby` wyznaczają więc adresy IP przełączania awaryjnego w każdym z interfejsów urządzenia podstawowego. Do interfejsu przełączania awaryjnego nie został przypisany żaden adres IP.

Konfiguracja interfejsu przełączania awaryjnego (z zachowaniem stanu) w podstawowym urządzeniu ASA sprowadza się do wykonania następujących poleceń:

```
failover lan unit primary
failover lan interface Failover GigabitEthernet0/3
failover link Failover GigabitEthernet0/3
failover interface ip Failover 192.168.255.1 255.255.255.0 standby 192.168.255.2
```

Instrukcja `failover lan interface` przekształca interfejs G0/3 w interfejs przełączania awaryjnego. Polecenie `failover link` sprawia, że staje się on interfejsem przeznaczonym do wymiany informacji o stanie połączeń. Oznacza to, że obydwie funkcje będą realizowane przez ten sam interfejs, co jest rozwiązaniem dozwolonym. Można je również rozdzielić, przekazując je nawet do odpowiednich sieci VLAN. Zazwyczaj jednak staram się wydzielić do tego celu osobne łącze fizyczne.

Aby ustanowić połączenie w sieci złożonej z przełączników, interfejs przełączania awaryjnego musi zostać skonfigurowany również po drugiej stronie (w urządzeniu zapasowym). Oto niezbędne polecenia:

```
failover lan unit secondary
failover lan interface Failover GigabitEthernet0/3
failover interface ip Failover 192.168.255.1 255.255.255.0 standby 192.168.255.2
```

Po zakończeniu konfiguracji wystarczy włączyć mechanizm za pomocą polecenia `failover`. Jeśli operacja przebiegła poprawnie, na ekranie powinien zostać wyświetlony następujący komunikat:

```
ASA-5540#failover
ASA-5540#
Beginning configuration replication: Sending to mate.
                                        //Rozpoczynanie konfiguracji replikacji: Wysyłanie do partnera.
End Configuration Replication to mate   //Zakończenie konfiguracji replikacji
```

Nieużywane interfejsy jednostek objętych przełączaniem awaryjnym powinny pozostawać w stanie `shutdown`. Aktywne, ale niepodłączone interfejsy mogą powodować przełączanie, ponieważ firewall będzie bezskutecznie próbował nawiązać kontakt z adresem przełączania awaryjnego (jeśli takowy został wcześniej skonfigurowany).

Monitorowanie przełączania awaryjnego

Podstawowym narzędziem do sprawdzania bieżącego stanu mechanizmu przełączania awaryjnego jest polecenie `show failover`:

```
ASA-5540#sho failover
Failover On
Failover unit Primary
Failover LAN Interface: Failover GigabitEthernet0/3 (up)
Unit Poll frequency 1 seconds, holdtime 15 seconds
Interface Poll frequency 5 seconds, holdtime 25 seconds
Interface Policy 1
Monitored Interfaces 2 of 250 maximum
Version: Ours 8.0(4), Mate 8.0(4)
Last Failover at: 14:07:24 UTC Aug 17 2010
        This host: Primary - Active
                Active time: 1313 (sec)
                slot 0: ASA5540 hw/sw rev (2.0/8.0(4)) status (Up Sys)
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
                Interface outside (10.0.0.1): Link Down (Waiting)
                Interface inside (192.168.1.1): Link Down (Waiting)
              slot 1: empty
        Other host: Secondary - Standby Ready
              Active time: 457 (sec)
              slot 0: ASA5540 hw/sw rev (2.0/8.0(4)) status (Up Sys)
                Interface outside (10.0.0.2): Link Down (Waiting)
                Interface inside (192.168.1.2): Link Down (Waiting)
              slot 1: empty

Stateful Failover Logical Update Statistics
        Link : Failover GigabitEthernet0/3 (up)
        Stateful Obj    xmit       xerr       rcv        rerr
        General         33         0          33         0
        sys cmd         33         0          33         0
        up time         0          0          0          0
        RPC services    0          0          0          0
        TCP conn        0          0          0          0
        UDP conn        0          0          0          0
        ARP tbl         0          0          0          0
        Xlate_Timeout   0          0          0          0
        VPN IKE upd     0          0          0          0
        VPN IPSEC upd   0          0          0          0
        VPN CTCP upd    0          0          0          0
        VPN SDI upd     0          0          0          0
        VPN DHCP upd    0          0          0          0
        SIP Session     0          0          0          0
Logical Update Queue Information
                Cur     Max     Total
Recv Q:         0       17      33
Xmit Q:         0       1024    2780
```

Zestawienie wynikowe polecenia zawiera informacje na temat obydwu zapór sieciowych pary oraz dane statystyczne o łączu przełączania awaryjnego. W przypadku zwiększających się wartości liczników błędów należy się liczyć z ryzykiem utracenia połączeń w czasie awaryjnego przełączania urządzeń.

Pamiętaj, że włączenie funkcji zachowania stanu podczas awaryjnego przełączania może sprawić, że cała operacja pozostanie niezauważona. Polecenie to przydaje się więc do sprawdzenia, czy urządzenie podstawowe nadal działa, czy uległo awarii. Jeśli funkcję jednostki podstawowej pełni firewall wyznaczony pierwotnie do roli urządzenia zapasowego, jednostka podstawowa musiała zostać uszkodzona:

```
ASA-5540#sho failover | inc host
        This host: Primary - Standby Ready  //Ta jednostka: Podstawowa — oczekująca w gotowości
        Other host: Secondary - Active      //Inna jednostka: Zapasowa — aktywna
```

Gdy aktywna jednostka ASA zawodzi, jej rolę przejmuje urządzenie zapasowe. Jednak przywrócenie sprawności uszkodzonej wcześniej jednostki nie powoduje automatycznej zmiany ról. Zgodnie z dokumentacją Cisco „nie ma powodu do zamiany ról aktywnego i oczekującego urządzenia" w takich przypadkach. Osobiście uważam, że przydałaby się opcja `preempt` (taka, jaka jest dostępna w rozwiązaniu HSRP), ale firma Cisco nie poprosiła mnie o napisanie kodu mechanizmu przełączania awaryjnego.

Aby wymusić przejęcie funkcji aktywnej zapory sieciowej przez jednostkę oczekującą, trzeba wykonać instrukcję `no failover active` w urządzeniu aktywnym lub `failover active` w firewallu oczekującym:

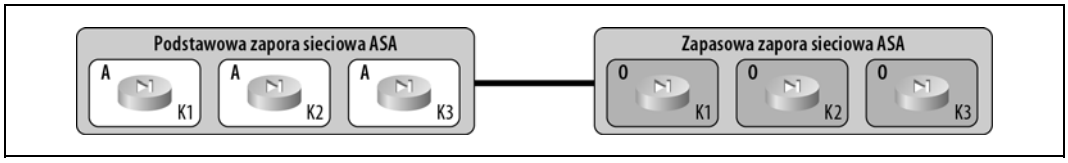
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
ASA-5540#failover active
Switching to Active  //Przełączenie w tryb aktywny
```

Jeśli operacja się powiedzie, podstawowa jednostka ASA stanie się jednocześnie urządzeniem aktywnym.

Konfiguracja przełączania awaryjnego — wariant aktywny-aktywny

W przypadku zdefiniowania wielu kontekstów i przełączania awaryjnego typu aktywny-oczekujący aktywne są wszystkie konteksty aktywnej jednostki ASA. Wynikowa konfiguracja jest analogiczna do przedstawionej na rysunku 28.13.

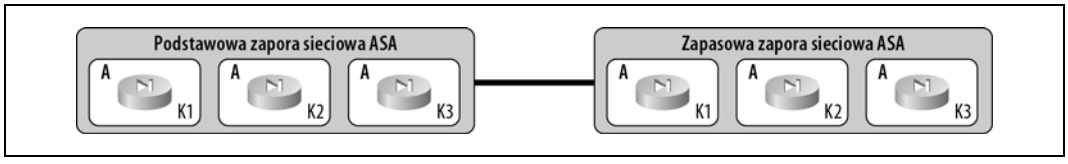


Rysunek 28.13. Przełączanie awaryjne typu aktywny-oczekujący z wieloma kontekstami

Gdy po raz pierwszy usłyszałem o wariancie aktywny-aktywny, ucieszyłem się, że firma Cisco w końcu opracowała rozwiązanie zbliżone do klastra urządzeń Checkpoint, w którym wiele fizycznych zapór sieciowych może pracować jako pojedynczy komponent logiczny z mechanizmem rozkładania obciążenia. Informacja o tym, że jest inaczej, bardzo mnie więc rozczarowała.

Gdy kończyłem pisanie tego rozdziału, przedstawiciel firmy Cisco powiedział mi, że w kolejnych wersjach systemów operacyjnych zawarte będą mechanizmy typu aktywny-aktywny charakterystyczne dla rozwiązań Checkpoint. Niestety nie umożliwiono mi przetestowania tej funkcji przed wydaniem książki.

Określenie aktywny-aktywny, mimo że technicznie poprawne, jest raczej terminem marketingowym. Tryb aktywny-aktywny w wielokontekstowych zaporach sieciowych ASA nie działa tak samo jak klaster firewalli Checkpoint. Jednostki ASA nie stanowią klastra i nie mogą jednocześnie przekazywać pakietów wszystkich kontekstów. Gdy słyszę o rozwiązaniu typu aktywny-aktywny, spodziewam się, że wszystkie konteksty będą aktywne w obydwu zaporach sieciowych, tak jak na rysunku 28.14.



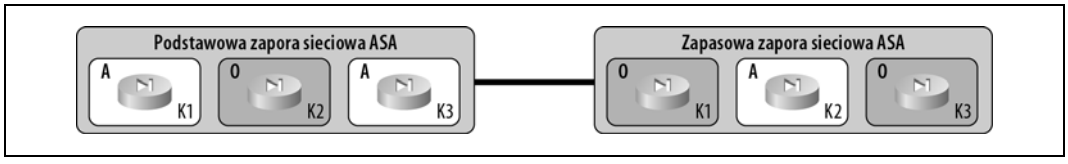
Rysunek 28.14. Idealny mechanizm przełączania awaryjnego typu aktywny-aktywny

Działanie urządzeń ASA w rozwiązaniu aktywny-aktywny jest **inne**. Po pierwsze, obydwie jednostki muszą pracować w trybie wielokontekstowym. Wynika to z tego, że część kontekstów jest aktywna po jednej stronie, a część po drugiej stronie wspólnego łącza.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Nie można skonfigurować wielokontekstowych zapór sieciowych ASA w taki sposób, aby wszystkie konteksty były aktywne po obydwu stronach łącza. Tak naprawdę **żaden** z kontekstów nie może być aktywny po obydwu stronach.

Wariant aktywny-aktywny oznacza, że część kontekstów jest aktywna w podstawowej jednostce ASA, a część w jednostce oczekującej. Ostateczne rozwiązanie jest więc zbliżone do przedstawionego na rysunku 28.15.



Rysunek 28.15. Rzeczywista konfiguracja przełączania awaryjnego w wariancie aktywny-aktywny

Stosując tę technikę, trzeba jednak zachować dużą ostrożność, ponieważ wiążą się z nią pewne pułapki. Na przykład wpisanie polecenia `write standby` w kontekście systemowym podstawowego urządzenia ASA (przy założeniu, że kontekst administracyjny po stronie jednostki podstawowej jest aktywny) powoduje zerwanie wszystkich połączeń obsługiwanych przez aktywne konteksty zapasowego urządzenia ASA. Oto fragment dokumentacji Cisco odnoszący się do tego ograniczenia:

> Wpisanie polecenia `write standby` w przestrzeni systemowej powoduje, że konfiguracja systemowa oraz konfiguracje wszystkich kontekstów jednostki ASA są zapisywane w urządzeniu zdalnym. Operacja ta obejmuje również dane na temat kontekstów pozostających w trybie oczekiwania. Polecenie trzeba wykonywać w przestrzeni systemowej tego urządzenia, którego grupa 1 przełączania awaryjnego pracuje w trybie aktywnym.
>
> **Uwaga**. Jeśli w urządzeniu zdalnym jakiekolwiek konteksty są aktywne, instrukcja `write standby` spowoduje przerwanie wszystkich obsługiwanych przez nie połączeń. Przed wpisaniem polecenia `write standby` wykonaj instrukcję `failover active` w systemie, w którym zmieniasz konfigurację. Dzięki temu wszystkie aktywne konteksty będą utrzymywane w jednym urządzeniu.
>
> Jeśli wpiszesz polecenie `write standby` w ramach danego kontekstu, do jednostki zdalnej zostanie przekazana jedynie konfiguracja tego kontekstu. Należy je więc wprowadzać w systemie tej jednostki, która utrzymuje aktywny kontekst.

Nie używaj zbyt często (jeśli w ogóle musisz) polecenia `write standby`. Jest ono potrzebne tylko w przypadkach, w których wykonanie pewnych instrukcji po stronie urządzenia zapasowego spowodowało rozsynchronizowanie konfiguracji. W normalnych warunkach urządzenia same zsynchronizują swoje ustawienia.

Jeśli po przeczytaniu tego opisu zastanawiasz się „po co jest ta funkcja?”, wiedz, że nie jesteś sam. Poza omówionymi ograniczeniami wariantu aktywny-aktywny, trzeba pamiętać, że sumaryczna przepustowość obydwu firewalli nie może przekroczyć pojemności jednego urządzenia. Jeśli jedna z zapór sieciowych ulegnie uszkodzeniu, cały ruch zostanie przekierowany do drugiej. Jeżeli więc każda z jednostek zostanie obciążona w stopniu przekraczającym połowę jej pojemności, po przełączeniu awaryjnym jednostka zapasowa zostanie przeciążona.

Biorąc pod uwagę wszystkie opisane czynniki, nie zalecam stosowania rozwiązań typu aktywny-aktywny, ponieważ rozkładanie ruchu tylko po to, by zająć urządzenia, nie ma sensu. Nie zyskujemy w ten sposób dodatkowej pojemności (przy założeniu poprawnej konfiguracji). Nie wzrasta więc przepustowość systemu. Niemniej można sobie wyobrazić sytuację, w której

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

trzeba skierować cały wrażliwy na opóźnienie ruch do jednego urządzenia oraz ruch o dużym natężeniu (w którym opóźnienia nie są istotne) do drugiej jednostki. W takich przypadkach dodatkowa opcja okazuje się przydatna.

Jeżeli nadal chcesz skorzystać z wariantu aktywny-aktywny, oto przepis na jego konfigurację. Tryb aktywny-aktywny jest uruchamiany po przypisaniu kontekstów do jednej z dwóch grup, do czego służy polecenie `failover group`:

```
ASA-5540(config)#failover group ?

configure mode commands/options: //polecenia (opcje) trybu konfiguracji:
  <1-2> group number             //numer grupy
```

Możliwe jest zdefiniowanie tylko dwóch grup — po jednej na każdą jednostkę ASA pary przełączania awaryjnego:

```
ASA-5540(config)#failover group 1
ASA-5540(config-fover-group)# ?
Failover User Group configuration mode:  //tryb konfiguracji własnych grup przełączania awaryjnego
  help              Help for user Failover Group configuration commands
                              //Pomoc w korzystaniu z poleceń konfiguracji własnych grup przełączania awaryjnego
  interface-policy  Set the policy for failover due to interface failures
                              //Ustawienie polityki przełączania awaryjnego wynikającego z uszkodzenia interfejsu
  mac               Specify the virtual mac address for a physical interface
                              //Określenie wirtualnego adresu MAC interfejsu fizycznego
  no                Remove user failover group configuration
                              //Usunięcie własnej grupy przełączania awaryjnego
  polltime          Configure failover interface polling interval
                              //Konfiguracja interwału odpytywania interfejsu przełączania awaryjnego
  preempt           Allow preemption of lower priority active unit
                              //Umożliwienie wywłaszczania aktywnych jednostek o niższym priorytecie
  primary           Primary unit has higher priority
                              //Przypisanie jednostce podstawowej wyższego priorytetu
  replication       Configure the replication option   //Konfiguracja opcji replikacji
  secondary         Secondary unit has higher priority //Przypisanie jednostce zapasowej wyższego priorytetu
```

Zgodnie z domyślnym ustawieniem konfigurowana grupa jest grupą podstawową. Zatem wprowadzenie instrukcji `primary` nie zostanie odzwierciedlone w pliku konfiguracyjnym. Myśl, co chcesz, ale ja lubię, gdy grupa podstawowa jest grupą 1, a grupa zapasowa jest grupą 2:

```
ASA-5540(config)#failover group 1
ASA-5540(config-fover-group)#primary
ASA-5540(config-fover-group)#preempt
ASA-5540(config-fover-group)#exit

ASA-5540(config)#failover group 2
ASA-5540(config-fover-group)#secondary
ASA-5540(config-fover-group)#exit
```

Po utworzeniu grup można rozpocząć przypisywanie kontekstów do poszczególnych spośród zdefiniowanych grup. Służy do tego polecenie `join-failover-group` *`grupa`*. Każdy kontekst powiązany z grupą 1 zostanie uaktywniony w podstawowej jednostce ASA. Natomiast konteksty przypisane do grupy 2 staną się aktywne w zapasowym urządzeniu ASA:

```
context GAD-Tech
  join-failover-group 1

context Test
  join-failover-group 2
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Wyświetlenie informacji na temat stanu mechanizmu przełączania awaryjnego w każdym z kontekstów sprowadza się do wprowadzenia polecenia `show failover` w ramach kontekstu systemowego. Jego wynik bywa niekiedy mylący, więc analizuj go starannie. Dane na temat grup zostały wyróżnione:

```
ASA-5540#sho failover
Failover On
Failover unit Primary
Failover LAN Interface: Failover GigabitEthernet0/3 (up)
Unit Poll frequency 1 seconds, holdtime 15 seconds
Interface Poll frequency 5 seconds, holdtime 25 seconds
Interface Policy 1
Monitored Interfaces 5 of 250 maximum
Version: Ours 8.0(4), Mate 8.0(4)
Group 1 last failover at: 07:33:09 UTC Jun 29 2010
Group 2 last failover at: 07:32:57 UTC Jun 29 2010

  This host:    Primary
  Group 1       State:          Active
                Active time:    92 (sec)
  Group 2       State:          Standby Ready
                Active time:    0 (sec)

                slot 0: ASA5540 hw/sw rev (2.0/8.0(4)) status (Up Sys)
                  admin Interface management (192.168.1.1): No Link (Waiting)
                  GAD-Tech Interface outside (10.10.10.1): Link Down (Waiting)
                  GAD-Tech Interface inside (0.0.0.0): Link Down (Waiting)
                  Test Interface outside (10.10.10.21): Link Down (Waiting)
                  Test Interface inside (0.0.0.0): Link Down (Waiting)
                slot 1: empty

Other host:     Secondary
  Group 1       State:          Standby Ready
                Active time:    2008 (sec)
  Group 2       State:          Active
                Active time:    2101 (sec)

                slot 0: ASA5540 hw/sw rev (2.0/8.0(4)) status (Up Sys)
                  admin Interface management (0.0.0.0): No Link (Waiting)
                  GAD-Tech Interface outside (10.10.10.2): Link Down (Waiting)
                  GAD-Tech Interface inside (0.0.0.0): Link Down (Waiting)
                  Test Interface outside (10.10.10.20): Link Down (Waiting)
                  Test Interface inside (0.0.0.0): Link Down (Waiting)
                slot 1: empty
Stateful Failover Logical Update Statistics
        Link : Unconfigured.
```

Zauważ, że mimo przypisania interfejsów do poszczególnych kontekstów status tych kontekstów nie jest prezentowany w bloku danych na temat grupy (w wyniku wykonania polecenia `show failover`). Nie należy więc szukać informacji o kontekstach w sekcji odnoszącej się do interfejsów. Po chwili zastanowienia dojdziesz do wniosku, że to ma sens. W omawianym rozwiązaniu interfejs G0/0 jest przypisany zarówno do kontekstu `GAD-Tech`, jak i `Test`. Kontekst `GAD-Tech` (należący do grupy 1) jest aktywny w urządzeniu podstawowym. Natomiast kontekst `Test` (należący do grupy 2) pracuje w urządzeniu zapasowym. Interfejsy nie są zależne od grupy, a konteksty tak.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

NAT

Translacja adresów sieciowych (NAT — *Network Address Translation*) jest nazwą, którą firma Cisco stosuje do określenia operacji tłumaczenia jednego adresu IP na inny. Większość instalacji, włącznie z sieciami domowymi, dokonuje translacji wielu adresów IP na pojedynczy adres. Taka technika jest nazywana **translacją adresów i portów** (PAT — *Port Address Translation*). W systemach IOS mechanizm PAT jest również określany jako **przeciążony NAT** (ang. *NAT overload*).

W systemach ASA rzecz jest jeszcze bardziej skomplikowana, ponieważ mechanizm NAT jest wykorzystywany na wiele sposobów, nie zawsze oczywistych. Na przykład można użyć polecenia `nat` do przekazania pakietów z jednego interfejsu do drugiego, mimo że na obydwu interfejsach wykorzystywane są publiczne adresy IP, które nie wymagają translacji.

Polecenia NAT

W większości praktycznych zastosowań mechanizmu NAT wykorzystywanych jest zaledwie kilka poleceń, choć niektóre z nich, takie jak `nat`, mają wiele opcji (nie wszystkie zostaną tutaj wymienione). Sama translacja adresów sieciowych w urządzeniach ASA jest zagadnieniem na osobną książkę. Moim celem jest natomiast zwięzłe przedstawienie problemu. Jeśli więc szukasz szczegółowych informacji na ten temat, skorzystaj z dokumentacji poleceń urządzenia (przynajmniej na początek). Do najczęściej wykorzystywanych instrukcji należą:

`nat`
: Polecenie `nat` jest stosowane do tłumaczenia adresów podczas przenoszenia ruchu z interfejsu o wyższym poziomie bezpieczeństwa do interfejsu o niższym poziomie bezpieczeństwa. Instrukcji `nat` należałoby użyć na przykład w przypadku tłumaczenia adresów sieci wewnętrznej na publiczny adres IP lub większą liczbę takich adresów. Z technicznego punktu widzenia wspomniane adresy nie muszą pochodzić z zakresów prywatnych lub publicznych, tak jak to zostało opisane w zaleceniu RFC 1918. W dokumentacji urządzenia ASA wykorzystano terminy „globalny" i „lokalny" do odróżnienia adresów sieci wewnętrznej od adresów zewnętrznych.

`static`
: Polecenie `static` odpowiada za translację adresów podczas przenoszenia ruchu z interfejsu o niższym poziomie bezpieczeństwa do interfejsu o wyższym poziomie bezpieczeństwa. Na przykład jeśli musisz przekazać pakiety do serwera działającego po wewnętrznej stronie jednostki ASA, możesz za pomocą instrukcji `static` powiązać prywatny i publiczny adres IP serwera.

`global`
: Polecenie `global` jest wykorzystywane w konfiguracji mechanizmu PAT, w którym wiele adresów IP jest tłumaczonych na jeden adres IP. Znajduje również zastosowanie w definiowaniu puli adresów mechanizmu NAT i występuje najczęściej w połączeniu z instrukcją `nat`.

Przykłady konfiguracji NAT

Istnieje wiele zastosowań mechanizmu NAT, z których część jest bardzo skomplikowana. Przedstawię więc tylko kilka najczęściej implementowanych rozwiązań.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

We wszystkich przykładach wykorzystałem adres 10.0.0.0 do zasymulowania publicznej sieci IP oraz adresu 192.168.1.0 jako adresu sieci prywatnej.

Prosta konfiguracja PAT z użyciem interfejsu zewnętrznego

Jednym z najczęstszych zastosowań zapór sieciowych jest ochrona sieci biurowej przed zagrożeniami internetowymi. Jeśli wszystkie jednostki sieci wewnętrznej korzystają z internetu, ale nie są do nich kierowane żadne żądania z zewnątrz, wówczas najkorzystniejsze wydaje się zastosowanie nieskomplikowanego mechanizmu PAT.

W opisywanym tutaj rozwiązaniu skonfigurowałem zewnętrzny interfejs tak, aby jego adres stał się globalnym adresem mechanizmu PAT. Innymi słowy, adresy wszystkich pakietów wychodzących z sieci wewnętrznej będą tłumaczone na jeden adres, odpowiadający zewnętrznemu interfejsowi urządzenia ASA:

```
global (outside) 1 interface
nat (inside) 1 0.0.0.0 0.0.0.0 0 0
```

Tłumaczeniem zostaną objęte wszystkie wewnętrzne adresy IP, ponieważ instrukcja `nat` zawiera definicję zakresu o wartości `0.0.0.0`, która odpowiada wszystkim adresom.

Prosta konfiguracja PAT z użyciem dedykowanego adresu IP

Aby wprost zdefiniować globalny adres mechanizmu PAT (bez wskazywania go pośrednio za pomocą interfejsu), należy zastosować polecenie `global`, jak w poprzednim przypadku, ale z adresem IP zamiast słowa kluczowego `interface`.

```
global (outside) 1 10.0.0.5
nat (inside) 1 0.0.0.0 0.0.0.0 0 0
```

Jeśli testujesz kolejne rozwiązania, z pewnością zauważyłeś, że przed wpisaniem kolejnych instrukcji konieczne jest usunięcie wcześniejszej konfiguracji. W przeciwnym przypadku otrzymasz powiadomienie o błędzie *Duplicate NAT entry*, informujące, że został powielony wpis w konfiguracji mechanizmu NAT.

Prosta konfiguracja PAT z publicznymi serwerami po stronie wewnętrznej

W niewielkich sieciach zdarza się, że serwery są uruchamiane po wewnętrznej stronie zapory sieciowej (a nie w strefie DMZ) i muszą być dostępne z internetu. Choć nie jest to najlepsze rozwiązanie, czasami trzeba je wdrożyć. Są one stosowane w niewielkich firmach (lub nawet w sieciach domowych), w których wyznaczenie stref DMZ jest niepraktyczne lub niemożliwe.

Zaproponowana konfiguracja obejmuje globalny mechanizm PAT z tłumaczeniem wszystkich adresów wewnętrznych na adres interfejsu zewnętrznego oraz dwa wpisy statyczne. Pierwszy z tych wpisów powoduje, że wszystkie pakiety dostarczane na adres `10.0.0.10` są przekazywane do jednostki o adresie prywatnym `192.168.1.10`. Drugi wpis odpowiada z kolei za translację adresu `10.0.0.11` na adres prywatny `192.168.1.11`:

```
global (outside) 1 interface
nat (inside) 1 0.0.0.0 0.0.0.0 0 0
static (inside,outside) 10.0.0.10 192.168.1.10 netmask 255.255.255.255 0 0
static (inside,outside) 10.0.0.11 192.168.1.11 netmask 255.255.255.255 0 0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Polecenie `static` nadpisuje ustawienia bardziej ogólnej instrukcji `nat`, więc można je stosować łącznie bez obaw o konflikt. Sama definicja `static` bywa jednak problematyczna z uwagi na różną kolejność adresów sieci i nazw interfejsów. Jeśli uważniej przyjrzysz się wcześniejszemu listingowi, zauważysz, że poszczególne odwołania zostały zapisane w następującej kolejności:

```
(int_wewnętrzny,int_zewnętrzny) sieć_zewnętrzna sieć_wewnętrzna
```

Pamiętaj, że przedstawione instrukcje `static` umożliwiają nawiązywanie połączeń zewnętrznych z dwoma jednostkami sieci wewnętrznej (o odpowiednio określonych adresach IP). Może się to wydawać niebezpieczne, ale dzięki temu, że interfejsowi zewnętrznemu odpowiada poziom bezpieczeństwa o wartości 0, a interfejs wewnętrzny ma przypisany poziom 100, ruch nie może być przenoszony z zewnątrz do wewnątrz bez uprzedniego zaakceptowania go za pomocą listy dostępu. Oznacza to, że musisz utworzyć listę dostępu, która zapewni przekazywanie określonego ruchu.

Przekierowanie portów

Przekierowanie portów różni się nieco od mechanizmu PAT, który zamienia grupę adresów na pojedyncze adresy, modyfikując numery portów w przesyłanych pakietach. Przekierowanie portów służy do czegoś zupełnie innego.

Umożliwia ono skonfigurowanie statycznego mechanizmu NAT, w którym wyznaczony zostanie jeden publiczny adres IP oraz wiele prywatnych adresów IP powiązanych z odpowiednimi portami. Rozwiązanie PAT nie pozwala na realizację połączeń przychodzących w przeciwieństwie do przekierowania portów.

Załóżmy, że mamy tylko osiem publicznych adresów IP, które zostały wykorzystane w następujący sposób:

- `.0` — adres sieci;
- `.1` — router ISP VIP (HSRP);
- `.2` — router ISP 1;
- `.3` — router ISP 2;
- `.4` — podstawowa jednostka ASA;
- `.5` — zapasowa jednostka ASA;
- `.6` — publiczny adres serwera WWW;
- `.7` — adres rozgłoszeniowy.

Choć instalowanie tak wielu urządzeń nadmiarowych w tak niewielkiej sieci może się wydawać nierealne, w praktyce nie takie rozwiązania można zobaczyć. Sieci niewielkich firm często są ograniczone do ośmiu adresów, ponieważ najczęściej więcej nie jest potrzebnych.

W omawianym przykładzie skonfigurowana została tylko jedna statyczna translacja NAT — do serwera WWW. Oto stosowne polecenia:

```
global (outside) 1 interface
nat (inside) 1 0.0.0.0 0.0.0.0 0 0
static (inside,outside) 10.0.0.6 192.168.1.6 netmask 255.255.255.255 0 0
```

Ich działanie jest zgodne z oczekiwaniami. Załóżmy jednak, że konieczne okazało się udostępnienie kolejnego serwera WWW — bezpiecznego serwera, który wykorzystuje protokół HTTPS

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

i nasłuchuje na porcie TCP 443. Zadanie okazuje się problematyczne, ponieważ nie ma wolnego publicznego adresu IP. Wiedząc, że pierwszy z serwerów pracuje na porcie TCP 80, możemy rozwiązać problem przez zastosowanie mechanizmu przekierowania portów.

Dzięki rozwiązaniom wprowadzonym w wersji 6.x oprogramowania możemy zapisać regułę, zgodnie z którą ruch przychodzący kierowany na adres `10.0.0.6` i na port 80 zostanie przekazany do serwera o jednym z wewnętrznych adresów IP. Natomiast pakiety dostarczane na ten sam zewnętrzny adres IP, ale na port TCP 443, będą przesyłane do jednostki o innym adresie IP:

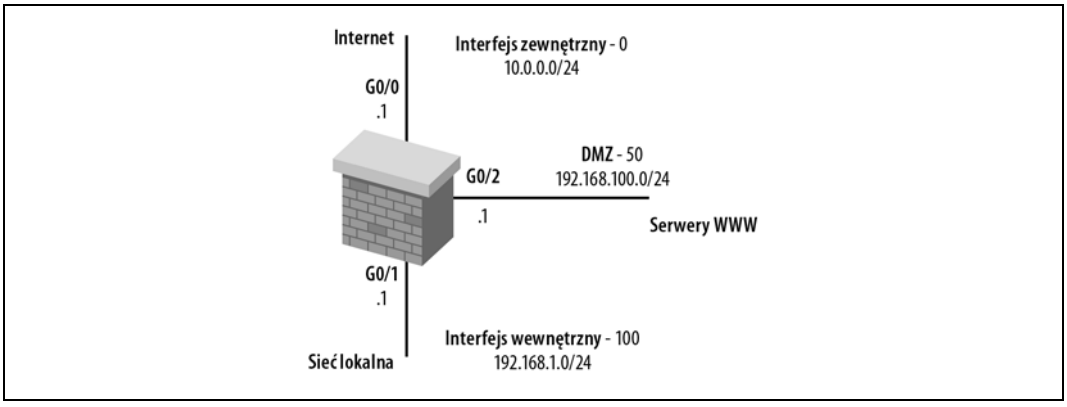
```
static (inside,outside) tcp 10.0.0.6 80 192.168.1.6 80 netmask 255.255.255.255
static (inside,outside) tcp 10.0.0.6 443 192.168.1.7 443 netmask 255.255.255.255
```

Standardowa wersja instrukcji `static` obejmuje jedynie adresy IP (zewnętrzny i wewnętrzny) i powoduje przekazywanie wszystkich pakietów między wskazanymi jednostkami. Uzupełnienie polecenia o numery portów sprawia jednak, że staje się ono bardziej selektywne.

W rezultacie pakiety kierowane na adres `10.0.0.6` będą otrzymywały różne wewnętrzne adresy IP w zależności od portów docelowych — adresy `10.0.0.6:80` będą zastępowane adresami `192.168.1.6:80`, natomiast adresy `10.0.0.6:443` zostaną przetłumaczone na `192.168.1.7:443`. Podczas tej operacji można również zmienić numery portów.

DMZ

Rozważmy pewną bardzo popularną konfigurację. Firma zainstalowała zaporę sieciową ASA z myślą o zabezpieczeniu sieci przed zagrożeniami internetowymi. Niektóre z jej serwerów muszą być jednak dostępne publicznie. Zostały więc umieszczone w strefie DMZ. Zewnętrzny interfejs zapory jest przyłączony do internetu, wewnętrzny do firmowej sieci LAN, a w strefie DMZ pracują serwery dostępne z internetu. Struktura takiej sieci została pokazana na rysunku 28.16.



Rysunek 28.16. Zapora sieciowa ze strefą DMZ

Konfigurując działanie mechanizmu NAT, musimy pamiętać o istnieniu poziomów bezpieczeństwa. Zewnętrznemu interfejsowi odpowiada poziom 0, wewnętrznemu interfejsowi poziom 100, a interfejsowi DMZ poziom 50.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W analizowanym przypadku serwery pracujące w strefie DMZ powinny być dostępne z sieci zewnętrznej. Ponadto możliwość komunikacji z nimi powinny mieć również jednostki sieci wewnętrznej. Nie chcemy jednak, aby serwery strefy DMZ nawiązywały połączenia z komputerami sieci LAN.

Po pierwsze, musimy wprowadzić instrukcje `nat` i `global` zapewniające dostęp z sieci wewnętrznej do internetu:

```
global (outside) 1 interface
nat (inside) 1 192.168.1.0 255.255.255.0 0 0
```

Zdefiniowanie określonej sieci zamiast użycia wartości `0.0.0.0` w poleceniu `nat` powoduje, że tylko jednostki należące do podanego zakresu adresów IP będą mogły korzystać z internetu. Jeśli po stronie wewnętrznej zostanie wydzielona kolejna sieć, która będzie musiała mieć zagwarantowany dostęp do internetu, wystarczy dodać kolejną instrukcję `nat (inside) 1`.

Następna czynność polega na przygotowaniu poleceń, które zapewnią dostęp do serwerów DMZ z internetu:

```
static (DMZ,outside) 10.0.0.11 192.168.100.11 netmask 255.255.255.255
static (DMZ,outside) 10.0.0.12 192.168.100.12 netmask 255.255.255.255
static (DMZ,outside) 10.0.0.13 192.168.100.13 netmask 255.255.255.255
```

Zgodnie z domyślnymi ustawieniami komputery pracujące w strefie DMZ nie mogą się komunikować z jednostkami sieci wewnętrznej, ponieważ interfejs DMZ ma niższy poziom bezpieczeństwa niż interfejs sieci wewnętrznej. Dlatego musimy użyć polecenia `static`, które umożliwi realizację połączeń. Nie chcemy jednak opisywać translacji adresów, a jedynie zezwolić na połączenia. Choć może się to wydawać dziwne, konieczne jest w tym przypadku zdefiniowanie statycznej operacji tłumaczenia adresów sieci wewnętrznej na te same adresy:

```
static (inside,DMZ) 192.168.1.0 192.168.1.0 netmask 255.255.255.0
```

W zaporze sieciowej ASA przekazywanie ruchu z interfejsu o niższym poziomie bezpieczeństwa do interfejsu o wyższym poziomie bezpieczeństwa wymaga zdefiniowania translacji adresów. Zadanie jest nieco mylące, ponieważ oznacza opisanie „translacji", która nie powoduje zamiany adresów. W tym przypadku jest to konieczne, aby komputery sieci wewnętrznej mogły się komunikować z jednostkami strefy DMZ. Adresy IP nie muszą być zmieniane, ale konieczne jest zdefiniowanie trasy ich przekazywania.

Po skonfigurowaniu mechanizmu NAT trzeba jeszcze przygotować listy dostępu, które umożliwią przenoszenie odpowiedniego ruchu ze strefy DMZ do sieci wewnętrznej.

Operacje uzupełniające

W tym podrozdziale opisałem kilka problemów, z którymi spotykam się w codziennej pracy.

Zdalny dostęp

Aby mieć możliwość ustanowienia połączenia telnet lub SSH z urządzeniem ASA, trzeba określić sieci, z których takie połączenia są dozwolone. Służą do tego instrukcje `telnet` i `ssh`:

```
ASA-5540(config)#telnet 192.168.1.0 255.255.255.0 inside
ASA-5540(config)#ssh 192.168.1.0 255.255.255.0 inside
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zapisywanie zmian w konfiguracji

Jeśli czytałeś pierwsze wydanie tej książki, pewnie pamiętasz, jak narzekałem na to, że zapory sieciowe PIX nie uznają ciągu `wri` jako skróconej wersji instrukcji `write memory`. Z radością informuję, że ciąg `wri` jest akceptowany w systemie ASA. Oczywiście przedstawiciele firmy Cisco twierdzą, że należy stosować polecenie `copy run start`, ale ja nie lubię zmian:

```
ASA-5540(config)#wri
Building configuration...
Cryptochecksum: 7401aa77 0584c65f 9ff42fa9 d5ed86ab
3191 bytes copied in 3.280 secs (1063 bytes/sec)
[OK]
```

Oczywiście instrukcja `copy running startup` również działa, ale jest mniej elegancka:

```
ASA-5540(config)#copy run startup-config

Source filename [running-config]?
Cryptochecksum: 7401aa77 0584c65f 9ff42fa9 d5ed86ab

3191 bytes copied in 3.280 secs (1063 bytes/sec)
```

W czasie konfigurowania jednostki ASA, która w parze przełączania awaryjnego jest modułem aktywnym, każde wpisywane polecenie powinno być automatycznie przekazywane do jednostki oczekującej. Z kolei zapisanie konfiguracji w urządzeniu aktywnym (za pomocą polecenia `write memory`) powinno spowodować zachowanie konfiguracji również w urządzeniu oczekującym. Aby wymusić zapis ustawień w jednostce zdalnej, należy użyć polecenia `write standby`:

```
ASA-5540#write standby
Building configuration...
[OK]
ASA-5540#Sync Started  //Synchronizacja rozpoczęta
.
Sync Completed         //Synchronizacja zakończona
```

Ciąg `Sync Started` nie jest poleceniem, ale komunikatem mechanizmu rejestracji zdarzeń (jeśli jest on włączony).

Polecenia `write standby` nie należy używać w konfiguracji aktywny-aktywny przełączania awaryjnego, ponieważ powoduje ono uaktywnienie wszystkich kontekstów w aktywnej zaporze sieciowej ASA. Więcej informacji na ten temat znajduje się w podrozdziale „Konfiguracja przełączania awaryjnego — wariant aktywny-aktywny” na stronie 482.

Rejestracja zdarzeń

Jeśli korzystasz z zapory sieciowej, powinieneś okresowo przeglądać generowane przez nią komunikaty. Włączenie rejestracji zdarzeń powoduje, że urządzenie generuje wiele użytecznych informacji, które mogą się okazać nieocenione nawet w niewielkiej sieci.

Choć zapisywanie informacji w buforze jednostki ASA mogłoby się wydawać dobrym rozwiązaniem, trzeba pamiętać, że dzienniki szybko się przepełnią, a tym samym staną się bezużyteczne. Jeśli będziesz wysyłał zbyt dużo szczegółowych komunikatów na konsolę, możesz doprowadzić do obniżenia wydajności zapory sieciowej. Nie najlepszym rozwiązaniem jest

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

również kierowanie ich do monitora (sesji telnet), ponieważ częstotliwość generowania powiadomień jest na tyle duża, że aby móc popracować, sam wyłączysz rejestrację zdarzeń.

Rozważ więc wysyłanie powiadomień z firewalla ASA do serwera syslog. Ja wykorzystuję do tego pewną wersję systemu Unix, ale oczywiście nie jest to jedyne rozwiązanie. Aby włączyć rejestrację zdarzeń, trzeba wykonać dwie czynności — uruchomić generowanie powiadomień za pomocą polecenia `logging on` oraz określić co najmniej jeden serwer docelowy. Przy definiowaniu parametrów serwera docelowego konieczne jest także wskazanie poziomu ważności komunikatów. Oto lista poziomów:

0 — System bezużyteczny

1 — Konieczna natychmiastowa reakcja

2 — Błąd krytyczny

3 — Komunikat o błędzie

4 — Ostrzeżenie

5 — Standardowa, ale istotna informacja

6 — Informacja

7 — Tryb debugowania

Najbardziej użyteczne informacje na temat ruchu przenoszonego przez firewall są zapisywane na poziomie 6. Poniżej został przedstawiony fragment dziennika z komunikatami poziomu 6.

```
302015: Built outbound UDP connection 3898824 for outside:11.1.1.1/123 (11.1.1.1/123) to
➥inside:192.168.1.5/123 (10.1.1.5/334)
302013: Built inbound TCP connection 3898825 for outside:12.2.2.2/2737 (12.2.2.2/2737) to
➥inside:192.168.1.21/80 (10.1.1.21/80)
302013: Built inbound TCP connection 3898826 for outside:13.3.3.3/49050 (13.3.3.3/49050) to
➥inside:192.168.1.21/80 (10.1.1.21/80)
304001: 15.5.5.5 Accessed URL 10.1.1.21:/lab/index.html/
```

W sieci użytkowej informacje są rejestrowane z taką szybkością, że wyświetlanie ich na ekranie nie ma sensu (ich przeczytanie nie jest możliwe). Niestety komunikaty z procesu debugowania mają wyższy poziom, co oznacza, że powiadomienia generowane w trybie debugowania będą przeplatane innymi wpisami.

Poniższe instrukcje odpowiadają za włączenie mechanizmu rejestracji zdarzeń w taki sposób, aby komunikaty poziomu 5 były wyświetlane na konsoli, a zdarzenia poziomu 7 rejestrowane w dziennikach pracy.

```
logging on
logging console notifications
```

Przedstawione polecenia odnoszą się do dzienników urządzenia ASA. Aby włączyć przekazywanie powiadomień do serwera syslog, trzeba ustawić poziom komunikatów, kategorię komunikatów oraz adres komputera:

```
logging trap debugging
logging facility 22
logging host inside 192.168.1.200
```

Oczywiście konieczne jest również skonfigurowanie serwera, który będzie odbierał komunikaty. Szczegółowy opis przygotowywania serwera syslog wykracza poza ramy tematyczne książki. Ograniczę zatem omówienie problemu do przedstawienia odpowiedniej zawartości pliku */etc/syslog.conf* z systemu Ubuntu:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
# Rejestracja zdarzeń zapory sieciowej ASA
local6.debug                            /var/log/ASA
```

Wiersz ten zapewnia przechwycenie powiadomień z urządzenia ASA i zapisanie ich w pliku */var/log/ASA*. Zauważ, że po stronie zapory sieciowej kategoria zdarzeń została ustawiona za pomocą instrukcji `facility 22`. Natomiast serwer zawiera w pliku konfiguracyjnym wpis `local6.debug`. Poszczególne kategorie odwzorowywane są w taki sposób, że wartość `16` odpowiada ustawieniu `local0`, natomiast wartość `23` jest odpowiednikiem ustawienia `local7`. Domyślnie wykorzystywaną kategorią zdarzeń w urządzeniu ASA jest `20` (`local4`).

Zapisywanie informacji o zdarzeniach w pliku pozwala na przeglądanie ich i analizowanie bez zajmowania okna konsoli. W systemach Unix doskonale sprawdza się w tym przypadku polecenie `tail -f /var/log/ASA`, które umożliwia bieżące monitorowanie pracy urządzeń. Oczywiście można również zastosować pewne filtry. Na przykład aby wyodrębnić jedynie te wpisy, które zawierają ciąg `/lab/index/html/`, wystarczy użyć polecenia `tail -f /var/log/ASA | grep '/lab/`
`↪index/html'`.

Więcej informacji na temat rejestracji zdarzeń można uzyskać, wpisując polecenie `help logging` w trybie konfiguracji urządzenia ASA. Z kolei opis serwera syslog zainstalowanego w systemie Unix jest dostępny po wprowadzeniu polecenia `man syslog` lub `man syslogd`.

Rozwiązywanie problemów

Zmiana listy dostępu, reguł translacji lub innych mechanizmów, które wpływają na przekazywanie pakietów przez zaporę sieciową, czasami jest niewidoczna do czasu wpisania polecenia `clear xlate`.

Słowo `xlate` oznacza translację. Definicja translacji jest tworzona przez urządzenie ASA na potrzeby każdego realizowanego połączenia. Aby sprawdzić, jakie translacje są wykorzystywane w danej chwili, należy wpisać polecenie `show xlate`:

```
ASA-5540#show xlate
4 in use, 4 most used
Global 10.0.0.10 Local 192.168.1.10
Global 10.0.0.11 Local 192.168.1.11
PAT Global 10.0.0.6(80) Local 192.168.1.6(80)
PAT Global 10.0.0.6(443) Local 192.168.1.7(443)
```

Wpisy `PAT Global` odpowiadają aktywnym połączeniom między moim komputerem a serwerami WWW. Generując to zestawienie, miałem uruchomioną przeglądarkę z kilkoma otwartymi stronami i pobierałem plik. Ostatni wpis jest wynikiem statycznej konfiguracji (zdefiniowanej we wcześniejszej części rozdziału).

Aby usunąć bieżące translacje, wystarczy wykonać polecenie `clear xlate`:

```
ASA-5540#clear xlate
```

Usunięcie translacji powoduje zerwanie wszystkich sesji zapory sieciowej i wymaga ponownego ustanowienia połączeń. Jeśli zadaniem urządzenia jest zabezpieczanie dostępu do komercyjnej witryny, przerwanie transakcji może spowodować niezadowolenie klientów. Nie należy więc używać tego polecenia, jeśli nie ma ku temu istotnych powodów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Wykonaniu polecenia nie towarzyszą żadne ostrzeżenia, mimo że powoduje ono przerwanie wszystkich połączeń. Skutkiem wyczyszczenia translacji było ponowne uruchomienie klienta IM w moim komputerze oraz przerwanie procesu pobierania pliku.

Innym poleceniem przydatnym podczas rozwiązywania problemów jest `show conn`. Umożliwia ono wyświetlenie wszystkich aktywnych połączeń.

```
ASA-5540#sho conn
8 in use, 199 most used
TCP out 10.233.161.147:80 in LAB-PC2:1151 idle 0:00:18 Bytes 6090 flags UIO
TCP out 10.46.109.49:1863 in LAB-SVR1:1736 idle 0:03:28 Bytes 7794 flags UIO
TCP out 10.188.8.176:5190 in LAB-PC2:4451 idle 0:00:52 Bytes 32827 flags UIO
TCP out 10.120.37.15:80 in LAB-PC:1789 idle 0:00:03 Bytes 19222477 flags UIO
TCP out 10.120.37.15:80 in LAB-PC:1802 idle 0:00:02 Bytes 20277173 flags UIO
TCP out 10.172.118.250:19093 in LAB-SVR2:80 idle 0:00:09 Bytes 11494 flags UIOB
TCP out 10.172.118.250:19075 in LAB-SVR2:80 idle 0:00:09 Bytes 219866 flags UIOB
UDP out 10.67.79.202:123 in RTR1:123 idle 0:00:32 flags -
```

Poszczególne wiersze listingu wynikowego zawierają informacje na temat protokołu, kierunku transmisji, adresów źródłowego i docelowego, a także czasu nieaktywności danego połączenia oraz liczby przesłanych bajtów. Użyteczne są również dołączane na końcu znaczniki. Te same dane (ale nieco inaczej sformatowane) wraz z listą znaczników są prezentowane po dodaniu słowa kluczowego `detail`. Oto przykładowe zestawienie wygenerowane kilka minut później niż poprzednie:

```
ASA-5540#sho conn detail
4 in use, 4 most used
Flags: A - awaiting inside ACK to SYN, a - awaiting outside ACK to SYN,
       B - initial SYN from outside, C - CTIQBE media, D - DNS, d - dump,
       E - outside back connection, F - outside FIN, f - inside FIN,
       G - group, g - MGCP, H - H.323, h - H.225.0, I - inbound data,
       i - incomplete, J - GTP, j - GTP data, K - GTP t3-response
       k - Skinny media, M - SMTP data, m - SIP media, n - GUP
       O - outbound data, P - inside back connection, p - Phone-proxy TFTP connection,
       q - SQL*Net data, R - outside acknowledged FIN,
       R - UDP SUNRPC, r - inside acknowledged FIN, S - awaiting inside SYN,
       s - awaiting outside SYN, T - SIP, t - SIP transient, U - up,
       V - VPN orphan, W - WAAS,
       X - inspected by service module
TCP outside:10.46.109.49/1863 inside:LAB-PC2/1736 flags UIO
TCP outside:10.188.8.176/5190 inside:LAB-PC/4451 flags UIO
TCP outside:10.241.244.1/48849 inside:LAB-SVR1/80 flags UIOB
UDP outside:10.30.70.56/161 inside:RTR1/1031 flags -
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 29.

Sieci bezprzewodowe

Praca w sieciach bezprzewodowych jest bardzo rozległym tematem. Obejmuje zagadnienia od podstaw komunikacji bezprzewodowej przez implementację mechanizmów roamingu w sieciach korporacyjnych po problematykę nieustannie zmieniających się zagrożeń sieciowych. Uwzględnienie ich wszystkich w książce poświęconej również wielu innym rozwiązaniom było po prostu niemożliwe. Poza tym wydano już wiele publikacji na ten temat. Szczególnie polecam książkę Matthew Gasta pt. *802.11 Wireless Networks: The Definitive Guide, Second Edition* (*http://oreilly.com/catalog/9780596001834*), w której przedstawiono szczegółowe omówienie tej fascynującej technologii.

Sieci bezprzewodowe są uruchamiane w dwóch trybach pracy: **infrastrukturalnym** (ang. *infrastructure*) i **ad hoc**. Infrastrukturalna bezprzewodowa sieć LAN (WLAN — *Wireless LAN*) składa się z co najmniej jednego punktu dostępu bezprzewodowego (WAP — *Wireless Access Point*), który zapewnia komputerom dostęp do sieci. W tym rozwiązaniu wszelkie formy komunikacji angażują punkt dostępowy, nawet jeśli połączenie jest realizowane między dwiema jednostkami sieci WLAN. Aby stacja kliencka mogła skorzystać z usług punktu dostępowego, musi się uwierzytelnić (jeśli opcja uwierzytelniania została włączona), a następnie stowarzyszyć z urządzeniem WAP. Klient może być stowarzyszony tylko z jednym punktem dostępowym w danej chwili. Jeśli znajdzie się poza zasięgiem tego punktu, może ustanowić podobną relację z kolejnym urządzeniem WAP, zachowując wcześniejszą konfigurację. Taka operacja nazywa się roamingiem.

Punkty dostępowe mogą być niezależnymi urządzeniami lub uproszczonymi komponentami sieciowymi nadzorowanymi przez kontroler bezprzewodowej sieci LAN. Kontrolery stosuje się przede wszystkim w sieciach bardzo rozległych, ponieważ pozwalają one na zarządzanie dużą liczbą modułów WAP z jednego miejsca.

Sieć ad hoc jest tworzona między dwoma stacjami końcowymi lub większą liczbą takich stacji. Do działania sieci nie jest potrzebny punkt dostępowy, gdyż każda stacja dostarcza dane bezpośrednio do odpowiedniego odbiorcy. Analogiczne rozwiązania od dziesięcioleci są stosowane w amatorskiej radiokomunikacji i nazywają się sieciami **packet radio**. Sieci ad hoc nie są wykorzystywane w systemach korporacyjnych.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Standardy komunikacji bezprzewodowej

Każdy, kto korzystał z sieci bezprzewodowych, wie, że obowiązuje w nich wiele różnych standardów i protokołów. Wystarczy jednak przeanalizować dokumenty IEEE, by się dowiedzieć, że jest ich więcej, niż można by przypuszczać. Są wśród nich specyfikacje takie jak 802.11k (rozszerzenia do pomiaru zasobów radiowych), 802.11y (sieci działające w Stanach Zjednoczonych na częstotliwościach 3650 – 3700 MHz) i wiele innych. Praktyczne zastosowanie ma tylko kilka z nich. Oto lista wykorzystywanych obecnie rozwiązań:

802.11a

Rozwiązanie to bazuje na standardzie 802.11-1997, który opisuje transmisję na częstotliwości 5 GHz z maksymalną przepływnością 54 Mb/s. Wartość przepływności odnosi się do strumienia uwzględniającego bity kontroli błędów oraz inne dane nadmiarowe. Użyteczna szybkość transmisji nie przekracza 20 Mb/s. Połączenia 802.11a mają relatywnie krótki zasięg (w porównaniu na przykład z systemem 802.11b).

802.11b

Standard 802.11b jest znany większości osób korzystających z sieci bezprzewodowych. Był bardzo popularny w czasie upowszechniania się tego rodzaju transmisji (około 2000 roku). Dzięki większemu zasięgowi był intensywnie promowany przez dostawców sprzętu sieciowego i stał się de facto standardem na wiele lat. Nawet dzisiaj wiele punktów dostępowych obsługuje transmisję zgodną ze specyfikacją 802.11b.

802.11g

Rozwiązania 802.11g również pracują na częstotliwości 2,4 GHz, ale wykorzystują znacznie efektywniejszą technikę modulacji, nazywaną **ortogonalnym zwielokrotnianiem w dziedzinie częstotliwości** (OFDM — *Orthogonal Frequency-Division Multiplexing*). Technika OFDM jest również stosowana w urządzeniach 802.11a, ale dzięki wykorzystaniu pasma 2,4 GHz moduły 802.11g mają większy zasięg. Maksymalna szybkość transmisji danych wynosi 54 Mb/s, ale po uwzględnieniu korekcji błędów i innego narzutu spada ona do około 22 Mb/s.

Ze stosowaniem urządzeń 802.11g wiąże się pewien problem. Standard 802.11g zakłada zachowanie zgodności z rozwiązaniami 802.11a/b. Dlatego w czasie komunikacji jednostek różnego typu (802.11b i 802.11g) szybkość transmisji może zostać obniżona do wartości wynikającej ze specyfikacji 802.11b.

802.11n

Standard 802.11n został zatwierdzony w październiku 2009 roku. Większość producentów sprzedawała jednak urządzenia 802.11n znacznie wcześniej. Po przyjęciu standardu po prostu nakleiła wielkie etykiety 802.11n na pudełkach. Dlaczego? Ponieważ producenci urządzeń bezprzewodowych zawiązali wcześniej konsorcjum o nazwie WiFi Alliance i oznaczali produkty jako zgodne z rozwiązaniem 802.11n, bazując na zgłoszonej w 2007 roku propozycji specyfikacji. Członkowie konsorcjum testują dostarczane komponenty, chcąc zapewnić zgodność między urządzeniami różnych dostawców. Dzięki temu odbiorcy nie są zmuszeni do kupowania modułów bezprzewodowych tylko jednej firmy.

Urządzenia 802.11n pracują zarówno w paśmie 2,4 GHz, jak i na częstotliwości 5 GHz. Korzystają z pasma o szerokości 40 MHz, czyli dwa razy większego niż w rozwiązaniach 802.11a/b/g. Dodatkowo obsługują multipleksowanie strumieni z wykorzystaniem wielu

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

anten (technika MIMO). Dzięki temu teoretyczna przepustowość połączenia 802.11n wynosi 600 Mb/s. W praktyce żadne z urządzeń nie może uzyskać wyniku nawet zbliżonego do tej wartości, głównie z powodu zakłóceń pochodzących od innych sygnałów radiowych.

Prywatnie korzystam z punktu dostępowego, który w mojej ocenie działał poprawnie. Moja żona skarżyła się jednak na niską przepływność danych i zrywanie połączeń. Po roku słuchania żali postanowiłem zobaczyć, w czym tkwi problem. No co, jestem dość zajętym człowiekiem. Wszystkie ustawienia wydawały mi się poprawne. Poza tym sam nie zauważałem żadnych problemów.

Kolejne testy przeprowadziłem po kupieniu nowego laptopa MacBook Pro, który również miał problemy z pracą sieciową. Nie mogłem dłużej winić za tę sytuację Windowsa ani taniego laptopa. Przeanalizowałem więc działanie sieci bezprzewodowej w domu i odkryłem, że na trasie między punktem dostępowym a miejscem pracy żony znajduje się kuchenka mikrofalowa oraz stacja bazowa telefonu bezprzewodowego. Gdy przeniosłem punkt dostępowy w inne miejsce, wszystkie problemy zniknęły.

Wynika z tego wniosek, że umiejscowienie punktów dostępowych ma znaczenie. Od lat używałem tego samego urządzenia, ustawiając je w wolnym miejscu na biurku (jeśli takie było). W związku z tym nie dość, że korzystałem ciągle ze starego standardu, to jeszcze przyczyniłem się do większości problemów w komunikacji sieciowej laptopa żony, które wynikały z interferencji pochodzących od innych źródeł sygnału. W sieci korporacyjnej taka sytuacja nie mogłaby się przydarzyć, ponieważ instalacja sieci jest poprzedzana pomiarami sygnałów radiowych w danym miejscu.

Bezpieczeństwo

Bezpieczeństwo od zawsze było największym zmartwieniem osób zajmujących się komunikacją bezprzewodową. Nawet dzisiaj, gdy technologia sieci bezprzewodowych jest powszechnie znana, mogę się przejść po osiedlu mieszkaniowym z laptopem i z pewnością znajdę niejedną niezabezpieczoną instalację. Znam osoby, które przez lata nie płaciły za internet, ponieważ ich sąsiedzi nie zabezpieczali odpowiednio sieci bezprzewodowych. Każdy mógł więc korzystać z ich łącza. Choć kradzież pasma dostępowego jest istotnym problemem moralnym, znacznie większym zmartwieniem powinien być nieograniczony dostęp osób postronnych do zasobów sieciowych. Dostęp do sieci i komputerów obcych osób jest przecież czymś zupełnie innym niż możliwość darmowego surfowania w internecie.

Zagwarantowanie odpowiedniego bezpieczeństwa sieci bezprzewodowej też nie jest wystarczającym rozwiązaniem. Dużo podróżuję i wiele czasu spędzam w hotelach. Gdy tylko się uwierzytelnię i uzyskam dostęp do sieci bezprzewodowej, sprawdzam, ile komputerów informuje o swoim istnieniu oraz jakie zasoby sieciowe udostępnia. Wiele osób nie chroni osobistych fotografii, dokumentacji medycznej i finansowej, a nawet list haseł. Wyświadcz więc sobie przysługę i wyłączaj udostępnianie plików na czas podróży. Wiadomo, że w sieci czai się wielu znudzonych maniaków komputerowych, którzy nie mają lepszego zajęcia niż przeszukiwanie udostępnionych udziałów sieciowych.

Istnieje kilka technik zabezpieczania sieci bezprzewodowych, które generalnie dzielą się na dwie kategorie — uwierzytelnianie i szyfrowanie. Oto niektóre z nich:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Filtrowanie adresów MAC

Filtrowanie adresów MAC jest użyteczną techniką uwierzytelniania, która może być stosowana w niewielkich sieciach albo w organizacjach zatrudniających osobę odpowiedzialną za aktualizowanie list adresów. Wdrożenie tego rozwiązania wymaga bowiem rejestrowania adresów MAC wszystkich bezprzewodowych kart sieciowych uprawnionych do korzystania z zasobów sieciowych. Wpis na liście jest sprawdzany w chwili, gdy stacja próbuje się stowarzyszyć z punktem dostępowym. Jeśli wpisu nie ma, użytkownik nie ma prawa dostępu do sieci. Jest to bardzo prosty mechanizm, który nie skaluje się odpowiednio i nie znajduje zastosowania w dużych sieciach korporacyjnych.

WEP

Skrót *WEP* pochodzi od słów *Wired Equivalent Privacy* oznaczających prywatność odpowiadającą połączeniom kablowym. Mechanizm WEP jest formą szyfrowania opracowaną specjalnie dla sieci bezprzewodowych. Niestety zabezpieczenia WEP są tak niedoskonałe, że powszechnie dostępne narzędzia łamią je w ciągu kilku minut. Technika ta jest najprostszą formą zabezpieczenia i wymaga od obydwu stron — punktu dostępowego i stacji klienckiej — przechowywania relatywnie krótkiego klucza (40- lub 128-bitowego). Nie należy stosować tego rozwiązania w żadnej nowo budowanej sieci, no, chyba że zabezpieczenia wcale nie są potrzebne, albo administrator jest leniwy, a sieć i tak pracuje głęboko w lesie. Skonfigurowanie połączenia WEP wymaga wstępnego zapisania hasła w punkcie dostępowym, a następnie w każdym komputerze klienckim. Niestety błędy projektowe algorytmu sprawiły, że jest on tylko nieznacznie lepszy od przesyłania kluczy w postaci jawnego tekstu. Zatem mimo że transmisja danych w sieciach zabezpieczonych protokołem WEP podlega szyfrowaniu, każdy może pobrać z internetu narzędzie, które złamie klucz w ciągu kilku minut.

Nieautoryzowany dostęp do mojej sieci domowej wymaga przejścia przez ogrodzenie. Za ogrodzeniem z kolei bezpieczeństwa sieci bezprzewodowej bronią dwa sześćdziesięciokilogramowe psy o bardzo nieprzyjaznym nastawieniu i małej wyrozumiałości dla intruzów. Oczywiście wcześniej trzeba ominąć wartowników, miny, pułapki oraz snajperów. No i mimów. Strzeż się mimów.

EAP

Rozszerzalny protokół uwierzytelniania (EAP — *Extensible Authentication Protocol*) jest środowiskiem uwierzytelniania, które zapewnia dynamiczne generowanie kluczy WEP i dostarczanie ich do urządzeń bezprzewodowych. Z uwagi na łatwość złamania stałego hasła dynamiczna zmiana kluczy gwarantuje wyższy poziom bezpieczeństwa. Mechanizm EAP występuje w wielu wersjach, w tym na przykład EAP-MD5, EAP-FAST i EAP-TLS. Do jego rozszerzeń należą także uproszczony rozszerzalny protokół uwierzytelniania (LEAP — *Lightweight Extensible Authentication Protocol*) oraz zabezpieczony rozszerzalny protokół uwierzytelniania (PEAP — *Protected Extensible Authentication Protocol*). Silniejsze algorytmy EAP (takie jak EAP-TLS) są wykorzystywane w rozwiązaniach WPA2-Enterprise.

TKIP

Protokół zachowania integralności kluczy tymczasowych (TKIP — *Temporal Key Integrity Protocol*) obejmuje kilka algorytmów, które zostały opracowane w celu wzmocnienia rozwiązania WEP. W standardowym mechanizmie WEP każdy pakiet jest szyfrowany z użyciem tego samego klucza. Technika TKIP gwarantuje zmianę klucza podczas przesyłania każdego pakietu, czyniąc algorytm trudniejszym do złamania. Oczywiście również w mecha-

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

nizmie TKIP wykryto luki, ale jest on znaczenie bezpieczniejszy niż WEP. Rozwiązanie TKIP warto stosować w tych urządzeniach, które nie obsługują sprzętowo algorytmu AES-CCMP. Jest rozwiązaniem obowiązkowym w systemie WPA.

AES-CCMP

Skrót AES-CCMP pochodzi od słów *Advanced Encryption Standard — Counter Mode with Cipher Block Chaining Message Authentication Code*. Tak, nazwa jest długa, a jej znajomość zbyteczna. Nie trzeba wiedzieć, co oznaczają poszczególne wyrazy, żeby sprawdzić, czy konkretne urządzenie może obsługiwać dany protokół. Nowoczesne urządzenia z grupy *N* muszą uwzględniać implementację tego protokołu, żeby uzyskać certyfikat WiFi Alliance gwarantujący zgodność modułu ze standardem WPA2.

WPA i WPA2

Skrót WPA oznacza bezpieczny dostęp do sieci WiFi (ang. *WiFi protected Access*). Sam protokół WPA bazuje na standardzie 802.11i opracowanym przez organizację IEEE w celu wyeliminowania niedoskonałości techniki WEP. Rozwiązanie WPA2 jest udoskonaloną wersją WPA. Uwzględnia bowiem algorytm AES-CCMP. Aby skorzystać z systemów WPA lub WPA2, trzeba zainstalować odpowiednie mechanizmy szyfrowania — TKIP w przypadku WPA oraz AES-CCMP w rozwiązaniach WPA2. Standard WPA2 należy do najlepszych sposobów zabezpieczania sieci bezprzewodowych. Protokół WPA należy stosować jedynie w starszych instalacjach, w których urządzenia nie obsługują mechanizmu AES-CCMP.

Algorytm WPA ma dwie wersje — **Personal** (osobistą), bazującą na wspólnych kluczach (PSK — *PreShared Key*) przeznaczonych do uwierzytelniania, oraz **Enterprise** (korporacyjną), która wymaga zastosowania protokołu 802.1x i zewnętrznego serwera uwierzytelniania, takiego jak RADIUS. Dlatego niekiedy rozwiązanie WPA2-Personal jest nazywane WPA2-PSK, a WPA2-Enterprise określa się jako WPA2-dot1x.

Zatem którego z rozwiązań należy używać? Po raz kolejny odsyłam do książek poświęconych bezpieczeństwu sieciowemu, projektowaniu sieci oraz implementowaniu poszczególnych rozwiązań, w których problem ten został omówiony znacznie bardziej szczegółowo. W przypadku systemów domowych algorytm WEP powinien wystarczyć, ale tylko wtedy, gdy masz psy i wyszkolonego mima. Większość małych sieci bazuje na rozwiązaniu WPA2-PSK, które wydaje się optymalnym połączeniem odpowiedniego poziomu bezpieczeństwa i małej złożoności. W przypadku większych przedsiębiorstw standardem jest instalowanie kontrolerów sieci bezprzewodowych.

Jeśli chcesz zapewnić całkowite bezpieczeństwo pracy sieciowej, skonfiguruj sieć tak, aby użytkownicy nie mieli dostępu do żadnych usług poza połączeniami VPN. Wymuś dwuetapowe uwierzytelnianie — pierwsze podczas przyłączania się do sieci bezprzewodowej i kolejne w czasie ustanawiania połączenia VPN z siecią użytkową. Dzięki temu rozwiązaniu, jeśli ktoś uzyska dostęp do sieci bezprzewodowej, nie będzie mógł wyrządzić żadnych szkód, ponieważ nie będzie w niej żadnych zasobów, a dane przesyłane przez innych użytkowników będą szyfrowane za pomocą innego silnego algorytmu. Oczywiście trzeba się liczyć z tym, że nikt nie będzie chciał korzystać z takiej sieci z uwagi na jej skomplikowanie. Bezpieczeństwo jest przecież równowagą między wygodą użytkownika i paranoją. Oczywiście pracowałem w sieciach, w których wymagane było ustanawianie połączeń VPN w ramach połączenia bezprzewodowego WPA2-Enterprise.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Często większe znaczenie niż metody uwierzytelniania ma architektura sieci współdziałającej z segmentami bezprzewodowymi. Elementami istotnie wpływającymi na bezpieczeństwo są również VLAN-y gości, VLAN-y bezprzewodowych połączeń VoIP oraz ustawienia kontrolerów. Również w tym przypadku zalecam sięgnięcie po specjalistyczne opracowanie, ponieważ zagadnienia te nie zostały omówione w dalszej części rozdziału.

Konfiguracja punktu dostępowego

Większość osób konfiguruje punkty dostępowe za pomocą przeglądarek internetowych. Taki sposób działania wynika z tego, że urządzenia domowe pozwalają jedynie na dostęp za pomocą protokołu HTTP. Ponieważ wolę posługiwać się interfejsem wiersza poleceń, w dalszej części rozdziału przedstawię sposób wykorzystania instrukcji punktu dostępowego Cisco 1231 do utworzenia sieci bezprzewodowej złożonej z tylko jednego urządzenia WAP. Konfiguracja byłaby znacznie bardziej skomplikowana, gdybyśmy chcieli uwzględnić problemy roamingu i konfiguracji kontrolerów. W środowiskach korporacyjnych zazwyczaj wykorzystuje się **uproszczone punkty dostępowe** (LAP — *Lightweight Access Point*), które pracują pod nadzorem kontrolera sieci bezprzewodowej. Konfigurowanie środowiska bezprzewodowego za pośrednictwem kontrolera diametralnie różni się od rozwiązań przedstawionych w dalszej części rozdziału. Po raz kolejny zalecam więc zapoznanie się ze specjalistycznymi opracowaniami dotyczącymi tych zagadnień.

Konfiguracja omawiana w tym rozdziale nie należy do najbardziej skomplikowanych, gdyż jej celem jest przedstawienie ogólnej idei wykorzystania systemu IOS w podstawowych punktach dostępowych. Istnieje wiele znacznie lepszych sposobów przygotowania sieci bezprzewodowej, włącznie z zastosowaniem oddzielnych VLAN-ów, roamingiem itp. Wybierając rozwiązanie, kieruj się własnym osądem. Pamiętaj jednak, że proste mechanizmy zabezpieczające, takie jak WEP, nie stanowią dostatecznej ochrony sieci i nie należy ich stosować w środowiskach korporacyjnych.

Opisywane dalej rozwiązania odnoszą się do urządzenia Cisco 1231, które jest wykorzystywane do obsługi niewielkiego biura (z kilkoma pracownikami). Zakładam, że więcej punktów dostępowych nie jest potrzebnych, a wymagania firmy odnośnie do bezpieczeństwa też nie są wygórowane (biuro znajduje się na uboczu i jest chronione przez agresywne i wygłodniałe psy). Pierwsza czynność polega na włączeniu funkcji mostowania za pomocą polecenia `bridge irb`. Nie znoszę mostów w klasycznych sieciach, ale w przypadku punktu dostępowego włączenie wspomnianej funkcji jest konieczne. Dzięki temu wszystkie komputery bezprzewodowe będą pracowały w sieci tak, jakby były do niej bezpośrednio przyłączone.

```
Cisco-WAP(config)#bridge irb
```

Następne zadanie to ustawienie identyfikatora zestawu usług (SSID — *Service Set ID*). Wybrałem nazwę `Coolness`. Na podstawie wartości SSID stacja kliencka może zidentyfikować daną sieć bezprzewodową. Liczba identyfikatorów SSID przypisanych do jednego punktu dostępowego nie jest wstępnie ograniczona. Można nawet kojarzyć poszczególne nazwy z sieciami VLAN. Ja zdefiniuję tylko jeden identyfikator:

```
Cisco-WAP(config)#dot11 ssid Coolness
Cisco-WAP(config-ssid)#authentication open
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Polecenie `authentication open` informuje punkt dostępowy, że w dostępie do sieci nie jest wymagana żadna forma uwierzytelnienia. Gdyby potwierdzanie tożsamości użytkowników okazało się konieczne, moglibyśmy użyć polecenia `authentication shared` do zdefiniowania wspólnego klucza lub instrukcji `network eap` do włączenia mechanizmu LEAP. Ponieważ mechanizm WEP bazuje na wspólnym kluczu, który zapewnia jedynie bardzo ogólny sposób weryfikacji tożsamości, rozwiązania tego nie można nazwać uwierzytelnianiem.

Choć otwarte (`open`) uwierzytelnianie wydaje się złym pomysłem (i rzeczywiście takim jest) w sieciach pracujących pod nadzorem kontrolerów, często stosuje się je w sieciach VLAN przeznaczonych dla gości. Samo uwierzytelnianie jest wówczas realizowane przez kontroler. Jeśli korzystałeś kiedyś z hotelowej sieci bezprzewodowej, z pewnością musiałeś otworzyć stronę, na której trzeba było podać dane uwierzytelniające. Jest to właśnie ten rodzaj mechanizmu.

Teraz muszę skonfigurować pulę DHCP, która oprócz adresów IP zapewni klientom bezprzewodowym dane na temat serwera DNS i domeny. Aby uatrakcyjnić prezentację, ograniczę zakres adresów IP do przedziału od `192.168.1.100` do `192.168.1.149`. Zazwyczaj tego nie robię, ponieważ wolę wyznaczać pule adresowe ograniczone przestrzenią adresową podsieci. Wydzielenie takiego przedziału odzwierciedla jednak pewną konfigurację, z której kiedyś korzystałem. Ponadto większość osób nie lubi posługiwać się granicami podsieci. Najpierw zdefiniuję wykluczenie adresów spoza ustalonego przedziału. Jest to konieczne również z tego względu, że nie chcę korzystać z zakresu odpowiadającego podsieci.

```
Cisco-WAP(config)#ip dhcp excluded-address 192.168.1.1 192.168.1.109
Cisco-WAP(config)#ip dhcp excluded-address 192.168.1.150 192.168.1.255
```

Po zdefiniowaniu wykluczeń mogę utworzyć pulę adresową. Ponieważ zadanie jest wykonywane w systemie urządzenia bezprzewodowego, a takie urządzenie posługuje się czasami więcej niż jednym identyfikatorem SSID, definicja zakresu adresów IP musi zostać powiązana z konkretnym identyfikatorem sieci. Wspomniane powiązanie opisuje polecenie `ip dhcp pool` *`SSID`*. Pozostałe pokazane instrukcje nie powinny sprawiać kłopotu osobom, które parametryzowały kiedykolwiek serwer DHCP:

```
Cisco-WAP(config)#ip dhcp pool Coolness
Cisco-WAP(dhcp-config)#network 192.168.1.0 255.255.255.0
Cisco-WAP(dhcp-config)#default-router 192.168.1.1
Cisco-WAP(dhcp-config)#domain-name gad.net
Cisco-WAP(dhcp-config)#dns-server 192.168.1.200 192.168.1.205
Cisco-WAP(dhcp-config)#lease infinite
```

Kolejne zadanie polega na skonfigurowaniu interfejsu radiowego. Jak w każdym innym przypadku interfejs ma pewną nazwę — `Dot11Radio`. Ponieważ opisywane urządzenie jest wyposażone w jeden moduł radiowy, do jego oznaczenia służy ciąg `Dot11Radio0`.

```
Cisco-WAP(config)#interface Dot11Radio0
Cisco-WAP(config-if)#no ip address
```

Następna czynność to konfiguracja mechanizmu WEP. Jak wiadomo, jest on najprostszym mechanizmem zabezpieczającym, ale zawsze lepszym niż brak zabezpieczeń. W przypadku uruchamiania sieci bezprzewodowej w środowisku korporacyjnym zalecałbym zastosowanie nieco bardziej efektywnych technik. Tutaj jednak użyję 40-bitowego klucza WEP, wykorzystując do tego celu instrukcję `encryption key 1 size 40bit 1234567890 transmit-key`. Dziesięciocyfrowy ciąg odpowiada dokładnie 40 bitom. Instrukcja `encryption mode wep mandatory` zapewni wymuszenie szyfrowania w ramach wszystkich połączeń. Dostępne opcje to `mandatory`

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

(obowiązkowe) i `optional` (opcjonalne). Nigdy nie używaj wartości `optional`, bo po prostu nie będzie opcjonalne. Tak naprawdę nie powinieneś również korzystać z 40-bitowego klucza. Jeśli musisz polegać na algorytmie WEP, stosuj klucz 128-bitowy. A najlepiej zmień algorytm na WPA2.

```
Cisco-WAP(config-if)#encryption key 1 size 40bit 7 9C3A6F2CBFBC transmit-key
Cisco-WAP(config-if)#encryption mode wep mandatory
```

Gdybym chciał zastosować mechanizm WPA-PSK, musiałbym wprowadzić poniższe polecenia. Zwróć uwagę na to, że do szyfrowania wykorzystałem algorytm TKIP (ustawiany w konfiguracji interfejsu).

```
Cisco-WAP(config)#dot11 ssid Coolness
Cisco-WAP(config-ssid)#authentication open
Cisco-WAP(config-ssid)#authentication key-management wpa
Cisco-WAP(config-ssid)#guest-mode
Cisco-WAP(config-ssid)#wpa-psk ascii 7 To-jest-niezbyt-bezpieczny-klucz
Cisco-WAP(config)#int dot11Radio 0
Cisco-WAP(config-if)#encryption mode ciphers tkip
```

Aby włączyć mechanizm WPA2-PSK, mógłbym wykonać poniższe instrukcje. Są one niemal identyczne jak w przypadku ustawienia WPA-PSK. Jedyną różnicą jest włączenie szyfrowania AES. Pamiętaj, że algorytm AES jest dostępny w wybranych urządzeniach. Jeśli nie możesz go uruchomić, upewnij się, że dane urządzenie ma zaimplementowaną obsługę standardu WPA2.

```
Cisco-WAP(config)#dot11 ssid Coolness
Cisco-WAP(config-ssid)#authentication open
Cisco-WAP(config-ssid)#authentication key-management wpa
Cisco-WAP(config-ssid)#guest-mode
Cisco-WAP(config-ssid)#wpa-psk ascii 7 To-jest-niezbyt-bezpieczny-klucz
Cisco-WAP(config)#int dot11Radio 0
Cisco-WAP(config-if)#encryption mode ciphers aes-ccm
```

W kolejnym kroku trzeba przypisać interfejsowi `Dot11Radio0` identyfikator SSID `Coolness`:

```
Cisco-WAP(config-if)#ssid Coolness
```

W starszych modelach punktu dostępowego w pliku konfiguracyjnym może występować wpis:

```
speed basic-1.0 basic-2.0 basic-5.5 6.0 9.0 basic-11.0 12.0 18.0 24.0 36.0 48.0 54.0
```

Błędna treść tego polecenia prowadzi często do ograniczenia przepustowości punktu dostępowego. Na szczęście nie trzeba weryfikować tych ustawień i dopisywać jakichkolwiek wartości. Wystarczy po prostu zresetować ustawienie, wpisując instrukcję `speed default`:

```
Cisco-WAP(config-if)#speed default
```

Interfejs bezprzewodowy musi zostać dołączony do grupy mostu o numerze 1 i — jak zawsze w takich przypadkach — uruchomiony.

```
Cisco-WAP(config-if)#bridge-group 1
Cisco-WAP(config-if)#no shut
```

Po tej operacji można przystąpić do konfigurowania interfejsu ethernetowego. Ponieważ punkt dostępowy pracuje w sposób zbliżony do mostu, interfejs fizyczny nie musi dysponować adresem IP. Konieczne jest natomiast dodanie go za pomocą polecenia `bridge-group` do grupy mostu, która zostanie skonfigurowana później.

```
Cisco-WAP(config)#int g0
Cisco-WAP(config-if)#no ip address
Cisco-WAP(config-if)#bridge-group 1
Cisco-WAP(config-if)#no shut
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Interfejsy fizyczne są odpowiednio sparametryzowane. Nadszedł więc czas na konfigurację wirtualnego interfejsu mostu (BVI — *Bridge Virtual Interface*). To jemu zostanie przypisany adres IP punktu dostępowego.

```
Cisco-WAP(config-if)#interface BVI1
Cisco-WAP(config-if)#ip address 192.168.1.3 255.255.255.0
Cisco-WAP(config-if)#no shut
```

Funkcja mostu została zdefiniowana. Następny etap to ustawienie bramy domyślnej (za pomocą polecenia `ip default-gateway`):

```
Cisco-WAP(config-if)#ip default-gateway 192.168.1.1
```

Trzeba też umożliwić przekazywanie pakietów IP przez most:

```
Cisco-WAP(config)#bridge 1 route ip
```

Jeśli wszystkie zadania zostały wykonane poprawnie, to na tym etapie prac jednostki klienckie powinny mieć możliwość stowarzyszania się z punktem dostępowym. Aby to sprawdzić, wystarczy wykonać polecenie `show dot11 associations`:

```
Cisco-WAP#sho dot11 associations

802.11 Client Stations on Dot11Radio0:

SSID [Coolness] :

MAC Address    IP address     Device       Name      Parent      State
001e.5275.6ee0 192.168.1.124  unknown      -         self        Assoc
7c6d.62a4.6aad 192.168.1.120  unknown      -         self        Assoc
```

Kolejne zestawienie jest wynikiem wykonania polecenia `show interface` po przesłaniu kilku plików przez włączony punkt dostępowy. Jak widać na listingu, przeciążenie interfejsu bezprzewodowego nie stanowi żadnego problemu, szczególnie gdy korzysta z niego wielu użytkowników. Sieci bezprzewodowe są cenione za wygodę pracy, ale nie zapewniają wysokiej wydajności komunikacji.

```
Cisco-WAP#sho int dot11Radio 0
Dot11Radio0 is up, line protocol is up
  Hardware is 802.11G Radio, address is 0012.7f52.de80 (bia 0012.7f52.de80)
  MTU 1500 bytes, BW 54000 Kbit, DLY 1000 usec,
     reliability 255/255, txload 26/255, rxload 25/255
  Encapsulation ARPA, loopback not set
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:00:13, output 00:00:01, output hang never
  Last clearing of "show interface" counters 00:14:57
  Input queue: 0/75/15/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: fifo
  Output queue: 0/30 (size/max)
  5 minute input rate 5420000 bits/sec, 502 packets/sec
  5 minute output rate 5510000 bits/sec, 502 packets/sec
     377892 packets input, 523715591 bytes, 0 no buffer
     Received 38 broadcasts, 0 runts, 0 giants, 0 throttles
     0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
     0 input packets with dribble condition detected
     378073 packets output, 532095123 bytes, 0 underruns
     8 output errors, 0 collisions, 0 interface resets
     0 babbles, 0 late collision, 0 deferred
     0 lost carrier, 0 no carrier
     0 output buffer failures, 0 output buffers swapped out
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Filtrowanie adresów MAC

Jednym z najłatwiejszych sposobów zabezpieczenia sieci bezprzewodowej jest włączenie mechanizmu filtrowania adresów MAC stacji stowarzyszających się z punktem dostępowym. Wadą tego rozwiązania jest duży narzut administracyjny. Każde nowe urządzenie musi bowiem zostać ręcznie dodane do listy przechowywanej w punkcie dostępowym. W sieciach o większej liczbie punktów dostępowych, w których nie jest wykorzystywany kontroler sieci bezprzewodowej, administrator musi ponowić tę czynność w każdym urządzeniu WAP. Oczywiście w sieciach domowych taki sposób postępowania jest znacznie korzystniejszy niż uruchamianie i utrzymywanie serwera RADIUS. Rozwiązanie to nie należy jednak do najbezpieczniejszych z uwagi na możliwość podmienienia adresu MAC. Wiele urządzeń jest zresztą dostarczanych z naklejoną informacją o adresie MAC.

Aby utworzyć odpowiednią listę dostępu, musiałem najpierw sprawdzić wszystkie urządzenia bezprzewodowe i zarejestrować ich adresy MAC. Przygotowany wykaz posłuży do utworzenia standardowej listy dostępu, której numer musi się zawierać w przedziale od 700 do 799:

```
Cisco-WAP(config)#access-list ?
<1-99>            IP standard access list                  //Standardowa lista dostępu z adresami IP
<100-199>         IP extended access list                  //Rozszerzona lista dostępu z adresami IP
<1100-1199>       Extended 48-bit MAC address access list
                                                  //Rozszerzona lista dostępu z 48-bitowymi adresami MAC
<1300-1999>       IP standard access list (expanded range)
                                                  //Standardowa lista dostępu z adresami IP (dodatkowy zakres)
<200-299>         Protocol type-code access list           //Lista dostępu z kodami protokołów
<2000-2699>       IP extended access list (expanded range)
                                                  //Rozszerzona lista dostępu z adresami IP (dodatkowy zakres)
<700-799>         48-bit MAC address access list           //Lista dostępu z 48-bitowymi adresami MAC
dynamic-extended  Extend the dynamic ACL absolute timer    //Zwiększenie czasu dynamicznej listy ACL
```

Wszystkie zarejestrowane adresy MAC zapiszę więc na liście dostępu o numerze 701:

```
Cisco-WAP(config)#access-list 701 permit 0025.0046.3f26 0000.0000.0000
Cisco-WAP(config)#access-list 701 permit 0014.a524.5b7a 0000.0000.0000
Cisco-WAP(config)#access-list 701 permit 001d.a231.6568 0000.0000.0000
Cisco-WAP(config)#access-list 701 permit 7c6d.62a4.6aad 0000.0000.0000
Cisco-WAP(config)#access-list 701 permit 001e.a98b.0018 0000.0000.0000
```

Utworzony w ten sposób filtr adresów MAC muszę skojarzyć z punktem dostępowym za pomocą polecenia `dot1 association mac-list` *`lista-dostępu`*. Zatwierdzenie instrukcji powoduje zresetowanie interfejsu radiowego i uaktywnienie filtru.

```
Cisco-WAP(config)#dot11 association mac-list 700
*Mar 1 01:51:19.543: %LINK-5-CHANGED: Interface Dot11Radio0, changed state to reset
*Mar 1 01:51:19.555: %LINK-3-UPDOWN: Interface Dot11Radio0, changed state to up
```

Rozwiązywanie problemów

Przeanalizujmy kilka poleceń, które przydają się podczas rozwiązywania problemów w działaniu punktu dostępowego. Z doświadczenia wiem, że najwięcej kłopotów sprawia operacja stowarzyszania. Doskonale sprawdza się wówczas instrukcja `show dot11 associations`. Jeśli inne urządzenia Cisco stowarzyszą się z danym punktem dostępowym, polecenie wyświetli dodatkowe informacje na ich temat.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Cisco-WAP#sho dot11 associations

802.11 Client Stations on Dot11Radio0:

SSID [Coolness] :

MAC Address    IP address       Device        Name            Parent      State
0014.a524.5b7a 192.168.1.129    ccx-client    hp-laptop       self        Assoc
001d.a231.6568 192.168.1.135    CP-7921       SEP001DA2316568 self        Assoc
001e.a98b.0018 192.168.1.120    unknown       -               self        Assoc
7c6d.62a4.6aad 192.168.1.110    unknown       -               self        Assoc
```

Diagnozując problemy ze stowarzyszeniem, utrzymuj aktywną sesję monitora (za pomocą polecenia `terminal monitor`). Dzięki temu komunikaty konsolowe będą przesyłane do klienckiego programu telnet lub SSH. Oto przykład komunikatu konsolowego, który został wygenerowany po odłączeniu zasilania od jednostki bezprzewodowej:

```
*Mar 1 00:56:54.671: %DOT11-6-DISASSOC: Interface Dot11Radio0, Deauthenticating Station 0025.0046.3f26
↪Reason: Sending station has left the BSS
```
//Interfejs Dot11Radio0, uznanie stacji za nieuwierzytelnioną, powód: stacja opuściła strefę BSS

Poniżej widzimy komunikat wyświetlony po przywróceniu zasilania:

```
*Mar 1 00:58:25.607: %DOT11-6-ASSOC: Interface Dot11Radio0, Station 0025.0046.3f26 Associated
KEY_MGMT[NONE]  //Interfejs Dot11Radio0, stacja stowarzyszona
```

Dwie kolejne instrukcje są bardzo przydatne, gdy trzeba ustalić przyczynę problemu z uwierzytelnieniem jednostki bezprzewodowej:

```
debug dot11 aaa authenticator process
debug dot11 aaa authenticator state-machine
```

Jeśli został włączony mechanizm WPA lub WPA2 z odwołaniami do serwera RADIUS, niezbędne może się również okazać prześledzenie komunikacji z usługą RADIUS.

```
debug radius authentication
```

Informacje wynikowe poleceń debugowania mogą się okazać dość rozwlekłe. Dlatego firma Cisco zamieściła w dokumencie 50843 (*http://www.cisco.com/en/US/products/hw/wireless/ps430/products_tech_note09186a008024aa4f.shtml*) przykłady generowanych komunikatów i uzupełniła je wskazówkami na temat danych wartych szczególnej uwagi.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 30.

VoIP

Technologia realizacji połączeń głosowych w sieciach IP (VoIP — *Voice over IP*) jest największą rewolucją w sieci od czasu wprowadzenia protokołu IP. Pewnie część osób nie zgodzi się z tym stwierdzeniem, ale nie ma wątpliwości, że rozwiązania VoIP staną się bardzo istotnym elementem pracy sieciowej. Zobaczmy więc, na czym one polegają, w jaki sposób działają i jak można zbudować niewielki system telefoniczny, dysponując routerem, przełącznikiem i kilkoma telefonami IP.

Podobnie jak w poprzednich rozdziałach posłużę się urządzeniami Cisco. Oczywiście dostępne są również systemy innych producentów, ale jeśli czytasz tę książkę, to prawdopodobnie zaopatrujesz się w komponenty Cisco lub przygotowujesz się do egzaminu Cisco. Jeżeli zdecydowanie nie chcesz używać rozwiązań Cisco (a wielu menedżerów przyjmuje takie założenie, szczególnie w odniesieniu do telefonii), proponuję sprawdzenie otwartego projektu Asterisk. Użycie telefonów SIP (opisanych w podrozdziale na temat protokołu SIP) i serwera Linux pozwala na zbudowanie wydajnego systemu telefonicznego bez dużych nakładów finansowych. Rozwiązaniu temu została nawet poświęcona odrębna książka Jima Van Meggelena pt. *Asterisk: The Future of Telephony, Second Edition* (*http://oreilly.com/catalog/9780596009625*).

Telefonia internetowa to bardzo rozległe zagadnienie. Na temat jej wykorzystania i konfiguracji powstają całe książki. Podobnie jak poprzednio moim celem będzie więc zaprezentowanie podstaw funkcjonowania technologii i przedstawienie praktycznego przykładu jej zastosowania. Nie znajdziesz w tym rozdziale wszystkich informacji na temat technologii VoIP. Dowiesz się jednak z niego, w jaki sposób można utworzyć system telefonii biurowej, wykorzystując do tego celu aplikacje Call Manager Express oraz telefony Cisco 79xx.

Jak działają systemy VoIP

Z perspektywy sieci systemy VoIP realizują dwie funkcje — sterują połączeniem i przekazują głos. Sterowaniem połączeniami zajmiemy się w dalszej części podrozdziału. Przeanalizujmy najpierw strumienie danych głosowych.

Telefonia jest systemem, w którym pewna reprezentacja analogowego sygnału głosu rozmówcy jest przekazywana przez sieć do zdalnego aparatu telefonicznego, gdzie zostaje zamieniona na dźwięk. W początkowej fazie rozwoju telefonii do przekazywania głosu służył sygnał elektryczny. Z czasem transmisja analogowa została zastąpiona cyfrową, którą poddano multipleksacji w ramach łączy magistralnych, takich jak T1.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Więcej informacji na temat łączy T1 i zwielokrotniania w systemach telefonii znajduje się w rozdziale 21.

Gdy duże przedsiębiorstwa zaczęły instalować sieci transmisji danych, zauważyły, że utrzymywanie (i opłacanie) sieci danych i sieci telefonicznych jest nadmiernym obciążeniem. Pojawiła się więc idea cyfryzacji, pakietyzacji i przesyłania sygnału rozmównego za pomocą sieci IP. Gdyby udało się połączyć dwie sieci w jedną, firmy mogłyby znacząco obniżyć koszty działalności. Oczywiście powszechnie wiadomo, że technologia VoIP jest niezwykle fascynująca, ale z punktu widzenia menedżera po prostu umożliwia wprowadzenie oszczędności.

Problemem jest jednak jakość połączeń. Bez wdawania się w szczegółowe rozważania techniczne można stwierdzić, że ludzie przyzwyczaili się do wysokiej jakości przekazywanego sygnału. Być może nie zdajesz sobie z tego sprawy, ale dźwięk telefoniczny jest naprawdę **dobry**.

Telefonia komórkowa znacznie obniżyła nasze wymagania odnośnie do jakości połączeń. Jednak zanim telefony komórkowe i rozwiązania VoIP stały się dostępne, jakość sygnału głosowego była wyjątkowo dobra, co zwolennicy poprzedniego rozwiązania chętnie podkreślali w rozmowach z narzekającymi użytkownikami systemów VoIP.

Aby dorównać jakości tradycyjnego analogowego sygnału rozmównego, cyfrowy (pakietowy) strumień danych głosowych musi być przekazywany bardzo szybko i niezawodnie. Pakiety muszą docierać do odbiorcy w takiej samej kolejności, w jakiej zostały wyemitowane. Niestety spełnienie tych założeń w sieci IP nie jest łatwe. Zwiększenie szybkości przesyłania danych gwarantuje zastosowanie protokołu UDP. Mechanizm UDP nie uwzględnia jednak sprawdzania błędów. Nie zapewnia również dostarczania pakietów w odpowiedniej kolejności. Do wyeliminowania opisanych niedogodności stosuje się mechanizm QoS. Dzięki niemu pakiety rozmów VoIP są przekazywane w pierwszej kolejności (przed innymi pakietami), a jeśli są przesyłane, to nie można ich usuwać z sieci. Dodatkową trudnością jest obsługa dwóch strumieni danych w odwrotnych kierunkach. Połączenia głosowe wymagają przecież transmisji dupleksowej.

Protokół transmisji w czasie rzeczywistym (RTP — *Real-Time Protocol*) uwzględnia znaczniki czasu i numery sekwencyjne, które umożliwiają wykrycie utraty pakietów oraz sprawdzenie ich kolejności. Nie obejmuje jednak mechanizmów, które pozwoliłyby naprawić zauważony błąd.

Ustanawianie połączeń między dwoma punktami końcowymi wymaga **sterowania połączeniem**. Sterowanie połączeniem jest procesem negocjowania i ustanawiania połączeń głosowych lub audiowizualnych. Mechanizm sterowania połączeniem rozpoczyna swoje działanie w chwili wybrania numeru telefonu. Sam numer telefonu jest informacją potrzebną przełącznikom w sieci operatora telekomunikacyjnego do wyznaczenia trasy połączenia. W telefonii IP numery telefonów są zamieniane na adresy IP jednostek końcowych, które tworzą dwa strumienie danych czasu rzeczywistego.

Zgodnie z terminologią telefoniczną komunikacja związana ze sterowaniem jest realizowana w **kanale danych** (ang. *data channel*). Natomiast strumień RTP wyznacza **kanał transportowy** (ang. *bearer channel*).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zasadnicze dane rozmowy są przekazywane za pośrednictwem protokołu RTP. Sam protokół RTP jest przeznaczony do transportu różnych rodzajów informacji, w tym strumieni wizyjnych. Jeśli ktoś używa określenia **strumień RTP**, z pewnością ma na myśli przesyłanie głosu lub wizji, a nie sterowanie rozmową.

Analizując połączenia telefoniczne w ujęciu sieciowym, warto zapamiętać następującą regułę: sterowanie połączeniem powinno być przenoszone w protokole IP z priorytetem 3, natomiast strumień RTP powinien mieć ustawioną wartość pierwszeństwa w protokole IP na poziomie 5. Stosowny przykład znajduje się w rozdziale 32.

Protokoły

W opracowaniach na temat rozwiązań Cisco nieustannie przewija się określenie **H.323**. Często używane są również oznaczenia **H.225**, **G.711**, **G.729**, **SIP** oraz wiele innych skrótów literowo-cyfrowych. Wszystkie odnoszą się do protokołów transmisji głosu i wideo w telefonii IP. Przeanalizujmy bardziej popularne spośród nich:

H.323

H.323 jest zbiorczą specyfikacją (w zasadzie rekomendacją) opracowaną przez Międzynarodową Unię Telekomunikacyjną (ITU — *International Telecommunications Union*), która definiuje zasady transmisji głosu, wizji i danych na bazie protokołu IP. Sygnalizacja w rozwiązaniach H.323 wywodzi się z zalecenia ITU Q.931 uwzględnianego także w obwodach ISDN. Zalecenie H.323 opisuje przeznaczenie urządzeń takich jak terminal, strażnik (ang. *gatekeeper*), brama oraz wiele elementów brzegowych. Standard H.323 jest powszechnie stosowany w systemach VoIP i wideokonferencyjnych. Został zaprojektowany przez osoby zajmujące się telefonią, przez co osobom bez odpowiedniego przygotowania teoretycznego z zakresu telekomunikacji może się wydawać skomplikowany.

H.225

H.225 jest protokołem sterowania rozmową. Wchodzi w skład grupy specyfikacji H.323.

H.245

H.245 jest protokołem składowym grupy H.323 wykorzystywanym do transmisji informacji innych niż telefoniczne (na przykład związanych z szyfrowaniem, przepływem danych i zarządzaniem fluktuacją opóźnienia). Jest również stosowany do tunelowania protokołu H.225 oraz przekazywania sygnałów DTMF.

SIP

Protokół inicjowania sesji (SIP — *Session Initiation Protocol*) jest protokołem sygnalizacyjnym opracowanym przez organizację IETF. Zyskał dużą popularność w ciągu kilku ostatnich lat, głównie dzięki temu, że jest rozwiązaniem otwartym. W rezultacie stał się podstawowym mechanizmem sygnalizacji w otwartych systemach VoIP, takich jak Asterisk. Jest niezwykle wydajnym i w pełni konfigurowalnym rozwiązaniem. Bazuje na komunikatach tekstowych, które są znacznie łatwiejsze do interpretowania niż dane H.323. W latach 90. ubiegłego stulecia mogło to być problemem z uwagi na ograniczoną moc obliczeniową urządzeń o niskim poborze energii oraz niedużą przepustowość połączeń. Jednak obecnie (w czasach, gdy moc obliczeniowa i szybkość transmisji danych nie są ograniczeniem) zyskuje coraz większą popularność. Protokół SIP został zaprojektowany z myślą o pracy w sieciach IP i z pewnością będzie łatwiejszy do zrozumienia dla osób bez przegotowania telekomunikacyjnego.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

SCCP

Uproszczony protokół sterowania rozmową (SCCP — *Skinny Call Control Protocol*) jest rozwiązaniem firmy Cisco. Służy jako mechanizm sygnalizacyjny w systemach zarządzanych przez oprogramowanie Cisco Unified Communications Manager (znane wcześniej jako Cisco Call Manager). Protokół SCCP jest wykorzystywany przez telefony opisane w dalszej części rozdziału, w tym przez oprogramowanie Cisco IP Communicator (telefon programowy), adaptery telefonów analogowych (ATA — *Analog Telephone Adapter*) oraz inne elementy końcowe sieci telefonicznej Cisco. Protokół ten jest źródłem sporów między inżynierami i menedżerami. Osoby zarządzające korporacjami nie chcą kupować telefonów IP firmy Cisco z uwagi na wyższą cenę w porównaniu z urządzeniami SIP. Mechanizm SCCP ma jednak istotną zaletę — jest łatwy w implementacji i wnosi niewielki narzut transmisyjny. Zużywa więc mniej zasobów sieciowych. Jeśli w firmie działa 5000 telefonów IP rejestrujących się w menedżerze połączeń, wymieniane przez nie komunikaty powinny być krótkie i łatwe do przetworzenia.

Kodek

Określenie **kodek** pochodzi od słów **ko**der-**dek**oder. Choć nie odnosi się ono bezpośrednio do protokołu, wiele spośród wymienionych tutaj protokołów obejmuje funkcje kodeków. Kodeki odpowiadają za przekształcanie danych lub sygnałów jednego formatu w dane i sygnały innego formatu. Standardy kodeków to MPEG, MP3 (element standardu MPEG), Dolby Digital, G.711 lub G.729.

G.711

G.711 jest specyfikacją opisującą sposób przekształcania analogowego sygnału głosu w strumień cyfrowy. Algorytm G.711 jest mechanizmem bezstratnym, co oznacza, że jego użycie nie powoduje obniżenia jakości dźwięku. Spośród stosowanych powszechnie rozwiązań wydaje się najwierniej przekazywać głos, a brak kompresji sprawia, że nie wymaga zaawansowanych procesorów DSP. Niestety ceną za wysoką jakość jest większe użycie pasma (64 kb/s przy próbkowaniu z częstotliwością 8 kHz). Ponieważ przepływność strumienia wynikowego jest taka sama jak w klasycznej sieci telefonicznej, nie wnosi żadnych oszczędności w gospodarowaniu pasmem. Sprawdza się natomiast jako rozwiązanie sieci konwergentnych.

G.722

Dźwięk zakodowany w standardzie G.711 wydaje się dobry, ponieważ ma taką samą jakość jak transmisja głosu w klasycznej sieci telefonicznej (POTS — *Plain Old Telephone Service*). Brzmienie głosu zakodowanego w standardzie G.722 jest jednak jeszcze lepsze. Mechanizm ten zapewnia jeszcze wyższą jakość sygnału niż usługi POTS. Często jest nazywany standardem **szerokopasmowej transmisji audio**, ponieważ próbkowanie sygnału odbywa się z częstotliwością 16 kHz, a nie 8 kHZ jak w systemach G.711. Kodeki G.722 są implementowane w najbardziej rozbudowanych telefonach Cisco (na przykład w modelu 7575G). Wadą rozwiązania jest wykorzystanie jeszcze większej szerokości pasma niż w standardzie G.711. Z tego powodu powinno być stosowane jedynie w sieciach lokalnych, w których problem małej szerokości pasma nie występuje.

G.729

Standard G.729 definiuje kodek, który kompresuje sygnał głosowy w stosunku 8:1. Cieszy się dużą popularnością, ponieważ doskonale łączy możliwość przekazywania sygnałów o wysokiej jakości z koniecznością oszczędzania zasobów sieciowych. Wadą kodowania G.729 jest zwiększone zapotrzebowanie na moc obliczeniową procesorów DSP. Istnieje kilka

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

wariantów standardu G.729, z których najważniejsze to G.729a i G.729b. Pierwsza z odmian zapewnia tę samą jakość i ten sam stopień kompresji, ale bazuje na algorytmie wymagającym mniejszych zasobów procesora DSP. Z kolei wariant G.729b obejmuje mechanizm detekcji aktywności głosowej (VAD — *Voice Activity Detection*), który umożliwia częściowe lub całkowite zmniejszenie ruchu sieciowego w chwilach milczenia.

Power over Ethernet (PoE)

Mogłoby się wydawać, że zasilanie w kablu ethernetowym (PoE — *Power over Ethernet*) nie jest protokołem, a jednak nim jest. W systemach Cisco stosuje się dwie odmiany tego protokołu — mechanizm Cisco sprzed opracowania specyfikacji PoE oraz standardowe rozwiązanie 802.3af. Firma Cisco sprzedawała urządzenia PoE, zanim został opracowany odpowiedni standard, i dlatego wciąż wykorzystywane są systemy typu „prestandard". W nowych przełącznikach obsługiwane są obydwa protokoły, ale kupując urządzenia używane, warto zwrócić uwagę na rodzaj zaimplementowanego w nich mechanizmu. Jeśli masz starszy przełącznik Cisco i podłączysz do niego telefon innej firmy, rozwiązanie PoE nie zadziała. Na przykład prezentowany w dalszej części rozdziału przełącznik Cisco 7960G współdziała jedynie z urządzeniami PoE prestandard. Z kolei przełącznik 7971 obsługuje obydwa protokoły.

Terminologia telefoniczna

Ponieważ technologia VoIP jest formą telefonii, musisz poznać kilka określeń związanych z tym rodzajem komunikacji. Niektóre z nich bywają mylące, ale trzeba pamiętać, że przemysł telefoniczny ma już ponad 100 lat.

PBX

Centrala abonencka (PBX — *Private Branch Exchange*). Jest najważniejszym elementem systemu telefonicznego przedsiębiorstwa.

ACD

Automatyczny dystrybutor połączeń (ACD — *Automatic Call Distributer*). System ACD jest zazwyczaj rozbudowanym, specjalistycznym systemem telefonicznym instalowanym w centrali. Moduły ACD służą do zestawienia połączeń między agentami na podstawie specjalnie przygotowanych skryptów. Dostarczają również informacji na temat parametrów połączeń, w tym średniego czasu odpowiedzi, średniego czasu podtrzymania rozmowy, poziomu głośności itp.

ANI

Automatyczny identyfikator numeru (ANI — *Automatic Number Identifier*) odpowiada numerowi dzwoniącego rozmówcy. Obejmuje również pewne dodatkowe informacje, które nie są istotne dla omawianego zagadnienia. Identyfikator ANI ma funkcję zbliżoną (ale nie identyczną) do identyfikatora rozmówcy, który wyświetla się na telefonie. Nawet jeśli identyfikator rozmówcy zostanie zablokowany, ANI będzie dostępny.

DNIS

Usługa identyfikacji wybranego numeru (DNIS — *Dialed Number Identification Service*) jest mechanizmem udostępniającym numer wybrany przez dzwoniącego. Centrale telefoniczne często wykorzystują ten numer do przełączania rozmów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Hairpinning

Operacja hairpinningu polega na przekazaniu rozmowy lub wiadomości do urządzenia, z którego pochodzi. Wyobraź sobie dwa telefony przyłączone do jednego interfejsu menedżera połączeń CME. Jeśli jedno z urządzeń ustanowi połączenie z drugim, a połączenie będzie przekazywane przez router, strumień głosowy z telefonu A będzie dostarczany do routera, a następnie tym samym interfejsem do telefonu B. Taki sposób działania nazywa się hairpinningiem. Zazwyczaj staramy się go unikać i wymusić bezpośrednią komunikację między urządzeniami po wcześniejszym nawiązaniu połączenia. Rozwiązanie nie jest więc stosowane w sieciach IP, ale w tradycyjnej telefonii jest to dość częsta operacja.

FXO

Port FXO łączy system telefoniczny z linią zewnętrzną sieci POTS. W dalszej części rozdziału został opisany przykład przyłączenia systemu telefonii IP do zewnętrznej linii POTS. Port FXO można więc utożsamiać z linią POTS prowadzącą do zewnętrznej centrali.

FXS

Port FXS umożliwia podłączenie tradycyjnych telefonów analogowych. Zapewnia zasilanie i ton wybierania numeru, a także generuje sygnał dzwonienia. Każde gniazdko telefoniczne linii POTS jest portem FXS. Port FXS należy więc traktować jak linię POTS, do której można przyłączyć telefon.

Przygotowując testową sieć telefonii IP połączoną z liniami POTS i telefonami analogowymi, z pewnością zadasz sobie pytanie „Jakich portów powinienem użyć?". Telefony są przyłączane do portów FXS. Zewnętrzne linie POTS z kolei muszą zostać podłączone do portów FXO. Można również zapamiętać następującą regułę: telefony analogowe muszą być zasilane, a tradycyjne centrale telefoniczne zapewniają zasilanie. Porty FXS dostarczają zasilanie, a porty FXO nie.

Telefon programowy

Telefon programowy to aplikacja telefonii IP zainstalowana w danym urządzeniu (na przykład w komputerze stacjonarnym, laptopie, tablecie lub nawet smartfonie).

SRST

Telefonia zapasowa w zdalnym oddziale (SRST — *Survivable Remote Site Telephony*). Załóżmy, że masz dziesięć biur w różnych lokalizacjach. W każdym z biur pracuje dziesięć telefonów IP. Wszystkie telefony rejestrują się w centrali IP działającej w głównej siedzibie firmy. Gdy główna centrala IP stanie się niedostępna, połączenia z którymkolwiek z oddziałów będą niemożliwe do zrealizowania. Skonfigurowanie pomocniczej centrali IP w jednym z ośrodków pozwoliłoby złagodzić skutki awarii.

PLAR

Skrót PLAR oznacza bezpośrednią linię prywatną. Podniesienie słuchawki po jednej stronie linii powoduje dzwonienie na drugim końcu. Nie trzeba wymieniać żadnych numerów.

Terminologia telefoniczna firmy Cisco

PVDM

Skrót PVDM oznacza procesor sygnałowy przeznaczony do przetwarzania pakietów głosowych (ang. *Packet Voice Digital Signal Procesor*). Istnieje wiele odmian tych procesorów (w tym PVDM, PVDM2, PVDM3). Komponenty PVDM są procesorami DSP. Procesory DSP są

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

natomiast specjalizowanymi układami scalonymi, które przekształcają sygnały z jednego formatu na inny, kompresują lub dekompresują sygnały informacyjne albo w inny sposób zmieniają bądź analizują transmitowane sygnały. Mogą zamieniać sygnał analogowy na cyfrowy lub cyfrowy na analogowy (podobnie jak przetworniki DAC), ale mogą również przekształcać strumień danych jednego kodeka na strumień innego formatu. Budując taką sieć, jaka została opisana w tym rozdziale, nie będziesz potrzebować procesora PVDM. Telefony firmy Cisco mają wbudowane układy DSP, więc konwersja analogowo-cyfrowa i cyfrowo-analogowa jest przez nie obsługiwana wewnętrznie. Moduły PVDM przydają się do budowy długodystansowych połączeń magistralnych, w których zasadne wydaje się przekształcanie strumieni danego kodeka w strumień o większym stopniu kompresji.

Cisco Unified Communications Manager

Oprogramowanie Cisco Unified Communication Manager (CUCM) — nazywane wcześniej Cisco Call Manager (CCM) — jest odpowiednikiem centrali abonenckiej w sieci telefonii IP. Zmiana nazwy menedżera była podyktowana tym, że jego możliwości wykraczają daleko poza ustanawianie połączeń głosowych. Moduł ten jest zazwyczaj instalowany w serwerze lub kilku serwerach w głównej siedzibie firmy. Wiele osób nazywa go menedżerem połączeń (Call Manager) niezależnie od tego, jaki rodzaj oprogramowania został zainstalowany.

Cisco Unified Communications Manager Express

Wcześniejsza nazwa tego produktu to Call Manager Express. Jest to zubożona wersja oprogramowania CUCM przeznaczona do instalacji w systemie routera Cisco. Rozwiązanie CUCM Express (CUCME) jest często wykorzystywane w architekturze SRST. Zasady jego użycia zostały opisane w dalszej części rozdziału. Z pewnością zauważysz, że w wielu odniesieniach do modułu CUCME użyłem skrótu CME. Nie stosuję skrótu CUCME, ponieważ skutecznie łamie język.

Cisco Unity

Cisco Unity jest aplikacją poczty głosowej przeznaczoną do stosowania w sieciach korporacyjnych. Obejmuje takie oprogramowanie jak Microsoft Exchange, Active Directory oraz kontrolery domen. Rozwiązania Cisco Unity wykraczają poza tematykę książki.

Cisco Unity Connection

Cisco Unity Connection jest nowym rozwiązaniem firmy Cisco w dziedzinie poczty głosowej. Nie wymaga całej infrastruktury programowej firmy Microsoft, która jest wykorzystywana w systemie Cisco Unity. Platforma Cisco Unity Connection działa w systemu Linux z zainstalowaną bazą danych Informix. Służy do zarządzania powiadomieniami. W najnowszej wersji obsługuje pojedyncze skrzynki pocztowe, w których zapisywane są zarówno wiadomości tekstowe, jak i głosowe. Produkt ten nie został opisany w książce.

Cisco Unity Express

Oprogramowanie Cisco Unity Express (CUE) jest zubożoną wersją platformy Unity, przeznaczoną do zainstalowania w systemie routera obsługującego połączenia głosowe. System CUE wykorzystuje niewielki komputer wbudowany w moduły sieciowe lub karty AIM. Pojemność modułów jest różna zależnie od ich przeznaczenia. Choć w dalszej części rozdziału można znaleźć wiele odniesień do systemu CUE, jego konfiguracja nie została opisana. Przeanalizujemy natomiast zasady definiowania odwołań do systemu CUE z modułu CME (CUCME).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Typowe problemy telefonii IP

Przetwarzanie głosu w formie cyfrowej zawsze kojarzy się z kilkoma problemami. Co prawda nie są one powiązane jedynie z systemami VoIP, ale często można o nich usłyszeć od inżynierów VoIP. Przeanalizujmy je więc nieco dokładniej.

Opóźnienie

Opóźnienie jest wielkością odzwierciedlającą ilość czasu, jaka jest potrzebna na przesłanie pakietu przez sieć. W przypadku systemów VoIP odpowiada ona czasowi podawanemu przez polecenie `ping`. Jak wiadomo, polecenie `ping` mierzy opóźnienie podczas transmisji w dwóch kierunkach. Pewną szczególną jego wartością jest 300 ms. W telefonii IP obowiązuje bowiem zalecenie ITU-T G.114, zgodnie z którym opóźnienie w transmisji w jednym kierunku nie może przekroczyć 150 ms. W praktyce rozmówcy zaczynają narzekać dopiero po przekroczeniu czasu 250 ms. Mnie jednak irytuje nawet opóźnienie na poziomie 180 ms.

Czy zdarzyło Ci się kiedyś w rozmowie telefonicznej rozpocząć wypowiedź w tym samym czasie, w którym zaczął mówić Twój rozmówca? Jest to wynik opóźnienia. Połączenie w nowoczesnym systemie telefonii komórkowej przebiega w następujący sposób. Telefon A przekształca głos na sygnał cyfrowy, kompresuje go i przesyła do lokalnej stacji bazowej. Stacja bazowa odbiera strumień danych i przekazuje przez co najmniej jedną sieć (a niekiedy przez sieci wielu operatorów telekomunikacyjnych) do stacji znajdującej się w pobliżu telefonu B. Ta z kolei transmituje sygnał do telefonu B, który musi zdekompresować dane, zamienić je na sygnał audio i wyemitować przez głośnik. Na każdym z wymienionych etapów wprowadzane jest pewne opóźnienie. Mimo że wszystko dzieje się w czasie rzeczywistym, opóźnienie i tak jest zauważalne.

Opóźnienie ma zazwyczaj dwie przyczyny — opóźnienie propagacyjne i opóźnienie obliczeniowe. Wysyłanie pakietów VoIP z Polski do Japonii oznacza konieczność przesłania ich na odległość bliską połowie obwodu Ziemi. Zwiększenie szybkości podróżowania poszczególnych bitów w łączu nie jest możliwe, więc opóźnienie propagacyjne jest elementem, na który nie mamy wpływu, szczególnie jeśli wymaga transmisji w nielicznych kablach ułożonych na dnie oceanu.

Bezpośrednia odległość między punktami końcowymi połączenia nie zawsze odpowiada długości połączenia fizycznego. W połączeniu między Nowym Jorkiem i Kalifornią wydzielono wiele krótkich odcinków, które umożliwiają efektywne wykorzystanie pojemności łączy międzymiastowych. Dodatkowe łącza fizyczne są w pełni przezroczyste dla użytkowników końcowych. Niemniej zwiększenie całkowitej długości łącza objawia się zwiększonym opóźnieniem. Znam przypadki zmniejszenia opóźnień o połowę po tym, jak firma przekonała operatora telekomunikacyjnego do wytyczenia efektywniejszej trasy przez sieć.

Opóźnienie obliczeniowe wynika z konieczności przetworzenia pakietu przez urządzenie, które go przekazuje. Routery, procesory DSP, telefony komórkowe i centrale abonenckie wnoszą pewne dodatkowe opóźnienia obliczeniowe. Przekształcenie sygnału G.711 w strumień G.729 wymaga skompresowania danych. Operacja ta powoduje opóźnienie transmisji. Sposobem na zmniejszenie opóźnień obliczeniowych jest usunięcie urządzeń z trasy pakietów, zmniejszenie złożoności algorytmów kompresji, wymiana urządzeń na wydajniejsze itd. Wynikiem opóźnienia jest wydłużenie przerw w emisji dźwięku lub powstawanie efektu echa. Więcej informacji na ten temat zostało zamieszczonych w rozdziale 20.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Utrata pakietów3

Utrata pakietów w połączeniu VoIP oznacza duże problemy, które trzeba rozwiązać. Główną bronią w tej walce jest użycie mechanizmu QoS. Jeśli po poprawnej konfiguracji systemu QoS problem nadal występuje, najprawdopodobniej jego przyczyną jest niewystarczające pasmo lub niewłaściwe działanie urządzeń.

Pamiętaj, że w internecie nie obowiązują żadne reguły QoS! Telefonia internetowa jest tania, ale w zamian za jakość. Obwody telefoniczne gwarantują pewien poziom jakości usługi, który nie jest dostępny w internetowych rozwiązaniach VoIP. Zapewnienie odpowiedniej jakości transmisji jest możliwe w sieciach lokalnych, ale z chwilą gdy pakiet ją opuszcza i trafia do wielkiej internetowej chmury, jest na łasce dostawców usług internetowych, routerów i łączy, które często są mocno przeciążone.

Utrata więcej niż 1% pakietów w połączeniu głosowym oznacza istotne obniżenie jakości sygnału audio. Sama utrata pakietów jest zazwyczaj symptomem nadsubskrypcji łącza, niewłaściwej konfiguracji mechanizmu QoS lub uszkodzenia sprzętowego. Użytkownicy postrzegają ją jako **zrywaną rozmowę**. Jeśli liczba utraconych pakietów wzrośnie za bardzo, połączenie w ogóle zostanie przerwane. Problem może się również objawiać jako zniekształcenia głosu (dźwięk sprawia często wrażenie generowanego mechanicznie).

Fluktuacja opóźnienia

Fluktuacja opóźnienia również wiąże się z dostarczaniem pakietów na czas. Jeśli zostaną one przekazane w odpowiedniej kolejności, ale pięć pierwszych pakietów dotrze do odbiorcy bezpośrednio po sobie, a pięć kolejnych zostanie rozdzielonych pakietami FTP o dużych rozmiarach (lub, co gorsza, ramkami jumbo), stacja odbiorcza będzie miała duży problem z przekształceniem strumienia danych na użyteczny sygnał audio. Z technicznego punktu widzenia mamy w takim przypadku do czynienia z **fluktuacją opóźnienia międzypakietowego**.

Ponieważ problem dotyczy mikrosekundowych rozbieżności, większość nowoczesnych urządzeń VoIP jest wyposażanych w **bufor fluktuacji opóźnienia** (ang. *jitter buffer*), który minimalizuje opisany efekt. Zadanie bufora polega na przechowywaniu nadchodzących pakietów przez pewien czas, zanim nie zostaną poddane dalszemu przetwarzaniu. Takie rozwiązanie eliminuje fluktuację opóźnień, ale wprowadza dodatkowe opóźnienie transmisyjne. Pakiet musi przecież zostać zmagazynowany na pewien czas przed przekazaniem go do dalszego przetwarzania. W przypadku dużych wartości fluktuacji istnieje możliwość zwiększenia rozmiaru bufora. To jednak zazwyczaj wprowadza opóźnienie powodujące przekroczenie progu irytacji u użytkowników sieci.

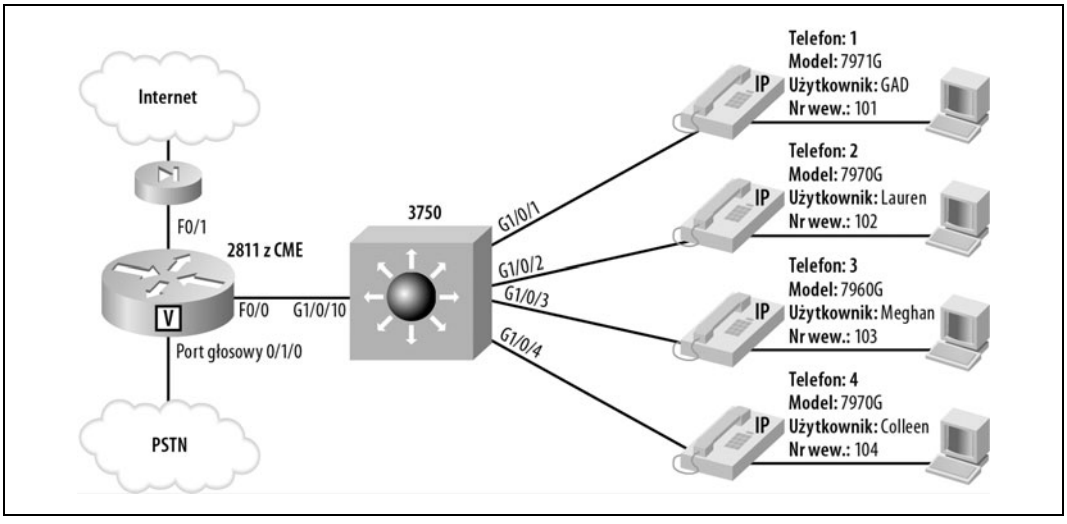
Fluktuacja opóźnienia jest szkodnikiem w świecie VoIP. Jeśli się zadomowiła, musisz ją wytępić. Najczęściej oznacza to poprawienie konfiguracji (lub włączenie) mechanizmu QoS.

Przykład systemu VoIP w małym biurze

Większość systemów VoIP bazujących na rozwiązaniach firmy Cisco składa się z punktów końcowych (telefonów), centrali abonenckiej (produktów UCM, Call Manager lub ich zubożonych wersji) oraz ewentualnie systemu poczty głosowej (Cisco Unity). Konfiguracja omawiana w dalszej

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

części rozdziału odpowiada przedstawionej na rysunku 30.1. Zaspokaja ona potrzeby małego biura konsultingowego. Wszelkie podobieństwa do rzeczywistych zdarzeń, a także osób żywych lub martwych, są zupełnie przypadkowe.



Rysunek 30.1. System VoIP małego biura bazujący na rozwiązaniu Cisco CME

Liczba możliwych konfiguracji systemu telefonii IP jest niemal nieograniczona. Prezentowany tutaj projekt jest tylko jednym z wielu sposobów wykonania zadania. Każde rozwiązanie, które spełnia założenia klienta, jest wydajne i nie ma luk w zabezpieczeniach, jest rozwiązaniem właściwym. Oczywiście są pewne sprawdzone techniki postępowania dotyczące szczególnie planów numeracji, ustanawiania połączeń i dodatkowych opcji, ale każda instalacja jest inna niż pozostałe.

Proponowane rozwiązanie bazuje na routerze Cisco 2811 wyposażonym w oprogramowanie UCM Express, kartę Cisco Unity Express AIM (AIM-CUE) oraz kartę o czterech portach FXO służącą do przyłączenia linii POTS. Router zawiera również procesor PVDM2-48 służący do konwersji strumieni audio, który nie będzie jednak wykorzystywany w omawianym przykładzie. Użyte telefony to Cisco 7971G, Cisco 7960G oraz dwa aparaty Cisco 7970G. Wszystkie są przyłączone do sieci za pośrednictwem przełącznika PoE Cisco 3750.

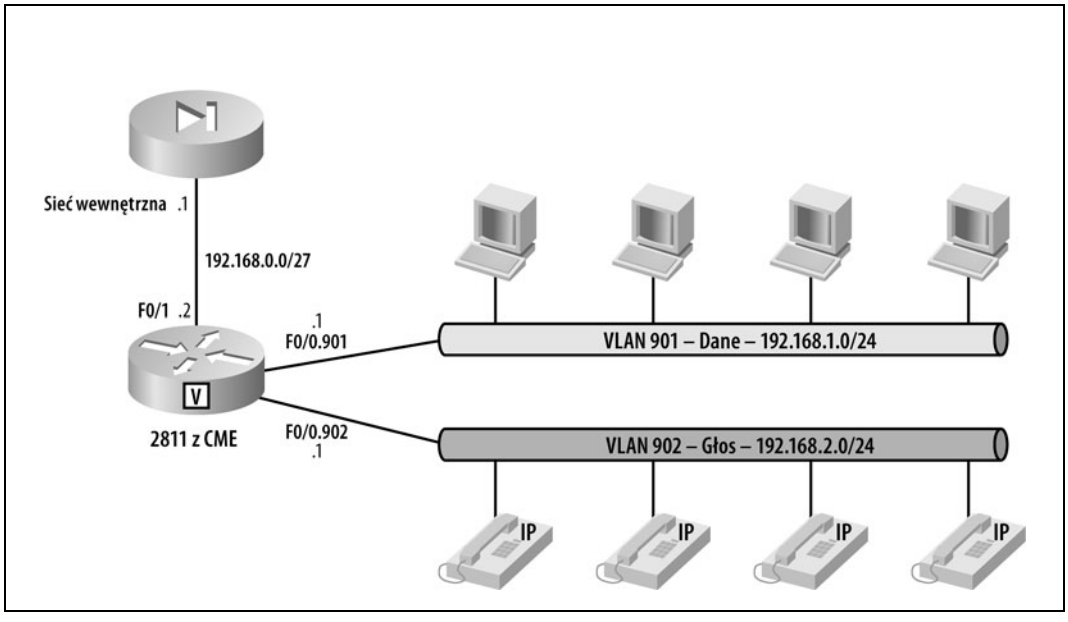
Jeśli budujesz sieć laboratoryjną, telefony nie mają znaczenia. Jeżeli jednak projektujesz ją na zamówienie klienta, pamiętaj, że są one najważniejszym składnikiem systemu. Użytkownik postrzega system przez pryzmat telefonu. Jeśli telefon nie spełni oczekiwań odbiorcy, cały system zostanie uznany za nieciekawy. Dużą rolę mogą w tym przypadku odegrać takie elementy jak podświetlany ekran czy podświetlane przyciski zajętości linii. Mój telefon (7960G) był bardzo tani, ponieważ firma wymieniała ten model na nowszy (7961G). Wydała tysiące dolarów na nowsze aparaty tylko dlatego, że w poprzednich nie było kontrolek informujących o zajętości linii.

Routery wykorzystywane w niewielkich sieciach często są uzupełniane o moduł przełącznika. Konfiguracja nie odbiega wówczas w żadnej mierze od przedstawionej na rysunku 30.1. Moduł przełącznika jest bowiem przyłączany wewnętrznie za pomocą gigabitowego interfejsu i działa identycznie jak oddzielne urządzenie 3750.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Sieci VLAN

Pierwsza czynność projektowa polega na zaplanowaniu sieci VLAN. Choć nic nie stoi na przeszkodzie, żeby przyłączyć telefon do tej samej sieci, w której są przekazywane dane, to nie polecam takiego rozwiązania. Zalecenia firmy Cisco odnośnie do budowy systemów VoIP zawierają wyraźne wytyczne, aby izolować ruch VoIP w oddzielnym VLAN-ie głosowym. Rozkład VLAN-ów został pokazany na rysunku 30.2.



Rysunek 30.2. Rozkład VLAN-ów w sieci VoIP

Omawiany przykład nie należy do najbardziej skomplikowanych, więc utworzyłem tylko dwa VLAN-y — jeden dla połączeń głosowych i jeden przeznaczony do transmisji danych.

```
SW-3750(config-vlan)#vlan 901
SW-3750(config-vlan)#name Dane
SW-3750(config-vlan)#vlan 902
SW-3750(config-vlan)#name Glos
```

Aby router mógł komunikować się z przełącznikiem, trzeba utworzyć między nimi łącze trunkowe. Najpierw w systemie przełącznika:

```
SW-3750(config)#interface GigabitEthernet1/0/10
SW-3750(config-if)#description [ R1-PBX F0/0 ]
SW-3750(config-if)#switchport trunk encapsulation dot1q
SW-3750(config-if)#switchport trunk allowed vlan 901,902
SW-3750(config-if)#switchport mode trunk
SW-3750(config-if)#speed 100
SW-3750(config-if)#duplex full
SW-3750(config-if)#no shut
```

Kolejne ustawienia wprowadzimy w routerze. Aby mógł się on komunikować z przełącznikiem za pośrednictwem łącza trunkowego, konieczne jest utworzenie podinterfejsów odpowiadających poszczególnym sieciom VLAN. Oto konfiguracja interfejsu fizycznego:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
R1-PBX(config)#interface FastEthernet0/0
R1-PBX(config-if)#description [ Trunk do 3750 ]
R1-PBX(config-if)#no ip address
R1-PBX(config-if)#duplex full
R1-PBX(config-if)#speed 100
```

Poniższe instrukcje ustawiają podinterfejs w VLAN-ie 901 (czyli VLAN-ie danych):

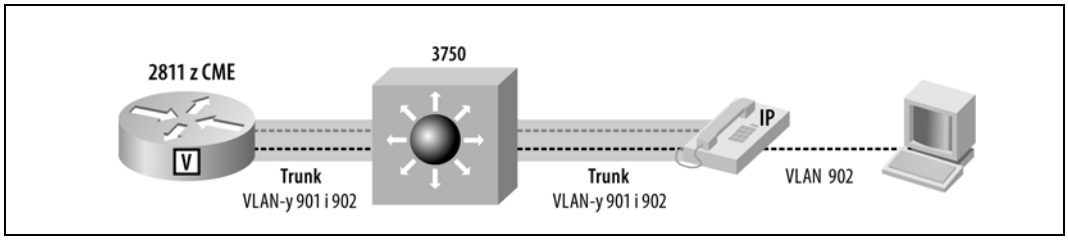
```
R1-PBX(config-if)#interface FastEthernet0/0.901
R1-PBX(config-subif)#description [ Dane ]
R1-PBX(config-subif)#encapsulation dot1Q 901
R1-PBX(config-subif)#ip address 192.168.1.1 255.255.255.0
```

Ostatni etap to konfiguracja podinterfejsu VLAN-u głosowego:

```
R1-PBX(config-subif)#interface FastEthernet0/0.902
R1-PBX(config-subif)#description [ Glos ]
R1-PBX(config-subif)#encapsulation dot1Q 902
R1-PBX(config-subif)#ip address 192.168.2.1 255.255.255.0
```

Żaden ze zdefiniowanych VLAN-ów nie ma nic wspólnego z technologią VoIP. Cała operacja sprowadziła się do utworzenia łącza trunkowego i podinterfejsów ethernetowych. Alternatywnie można było uznać VLAN 901 za VLAN natywny i skonfigurować tylko jeden podinterfejs 802.1Q jako interfejs VLAN-u głosowego.

Z rysunku prezentującego projekt sieci VLAN wynika, że telefony należą do jednej sieci VLAN, a stacje robocze do drugiej. Tymczasem schemat połączenia fizycznego przedstawia komputery połączone z telefonami. Takie połączenie jest możliwe dzięki wykorzystaniu dwuportowych przełączników wbudowanych w aparaty telefoniczne. Połączenie między telefonem a przełącznikiem ma formę łącza trunkowego, które przenosi ramki obydwu VLAN-ów. Drugi port jest przeznaczony do podłączenia komputera. Telefon zapewni połączenie z komputerem w ramach VLAN-u danych. Natomiast sam będzie komunikował się z siecią za pośrednictwem VLAN-u głosowego. Funkcje portów telefonu zostały pokazane na rysunku 30.3. Można je konfigurować na wiele ciekawych sposobów, ale to rozwiązanie jest zgodne z zaleceniami.



Rysunek 30.3. Podłączenie komputera i telefonu IP

Porty przełącznika

Telefony są podłączone do przełącznika za pomocą łączy trunkowych. Skonfigurujmy więc odpowiednio porty, z których korzystają. Załóżmy, że zostaną na nie zarezerwowane porty od 1 do 9. Polecenie `interface range` pozwoli na ustawienie wszystkich portów za jednym razem:

```
SW-3750(config)#int range g1/0/1 - 9
SW-3750(config-if-range)#description [ Telefon IP ]
SW-3750(config-if-range)#switchport trunk encapsulation dot1q
SW-3750(config-if-range)#switchport trunk native vlan 901
SW-3750(config-if-range)#switchport mode trunk
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
SW-3750(config-if-range)#switchport nonegotiate
SW-3750(config-if-range)#switchport voice vlan 902
SW-3750(config-if-range)#spanning-tree portfast
```

Wszystkie porty, do których przyłączone są telefony IP firmy Cisco, powinny zostać uwzględnione w mechanizmie QoS. Przełącznik powinien bowiem w szczególny sposób traktować pakiety pochodzące z aparatu telefonicznego. Wykorzystamy do tego celu polecenie interfejsu `auto qos`. Jak widać na poniższym listingu, nie ma ono wielu opcji. Wybierz jedną w zależności od rodzaju przyłączonego urządzenia.

```
SW-3750(config-if-range)#auto qos voip ?
cisco-phone       Trust the QoS marking of Cisco IP Phone
                                        //Zaufaj znakowaniu wprowadzonemu przez telefon IP Cisco
cisco-softphone   Trust the QoS marking of Cisco IP SoftPhone
                                        //Zaufaj znakowaniu wprowadzonemu przez aplikację Cisco IP SoftPhone
trust             Trust the DSCP/CoS marking  //Zaufaj znokowaniu DSCP/CoS
```

Ponieważ korzystam tylko z telefonów IP firmy Cisco, zastosuję opcję `cisco-phone`:

```
SW-3750(config-if-range)#auto qos voip cisco-phone
```

Jeśli wpisałeś polecenie po wcześniejszym podłączeniu telefonu, powinieneś zobaczyć na ekranie komunikat zbliżony do przedstawionego poniżej.

```
1w3d: %SWITCH_QOS_TB-5-TRUST_DEVICE_DETECTED: cisco-phone detected on port Gi1/0/1, port's configured
➥trust state is now operational.
            //Wykryto telefon Cisco na porcie Gi/1/0/1, skonfigurowany na porcie stan zaufania jest teraz obowiązujący
```

Kiedy następnym razem będziesz przeglądał plik konfiguracji bieżącej, nie przestrasz się całej strony poleceń `mls qos`. Wykonanie zadania w moim przełączniku spowodowało dodanie około 45 skomplikowanych instrukcji na samym początku pliku, więc żeby sprawdzić jakiekolwiek ustawienia, muszę najpierw przewinąć wstawiony kod. Przykładowy fragment pliku został zamieszczony w rozdziale 19. Włączenie mechanizmu AutoQoS spowodowało również zaśmiecenie konfiguracji każdego interfejsu, do którego przyłączony jest telefon:

```
srr-queue bandwidth share 10 10 60 20
srr-queue bandwidth shape 10 0 0 0
queue-set 2
mls qos trust device cisco-phone
mls qos trust cos
auto qos voip cisco-phone
```

Konfiguracja z pewnością nie wygląda przyjaźnie, ale w systemach VoIP nigdy nie należy ignorować mechanizmu QoS.

Praca po stronie przełącznika 3750 została zakończona. Pozostało jedynie skonfigurowanie VLAN-u i interfejsów sieci zewnętrznej znajdującej się pomiędzy routerem i zaporą sieciową. Ich ustawienia nie są jednak związane z systemem VoIP, więc nie będę się nimi zajmował.

Mechanizm QoS w routerze CME

Pewne ustawienia mechanizmu QoS muszą również zostać zdefiniowane na interfejsach routera. Postarajmy się załatwić sprawę w jak najmniej skomplikowany sposób. Zdefiniuję klasy połączeń sterujących oraz strumieni protokołu RTP przenoszącego głos. Przenoszenie innego ruchu pozostanie bez zmian. Bardziej rozbudowany przykład parametryzowania systemu QoS został przedstawiony w rozdziale 32. W implementowanym rozwiązaniu QoS zakładamy, że telefony

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

będą odpowiednio oznaczały generowane pakiety. To założenie jest poprawne, ponieważ korzystamy z telefonów firmy Cisco. W pierwszym kroku zdefiniujemy mapy klas (`class-map`) odpowiedzialne za identyfikację i klasyfikację ruchu:

```
R1-PBX(config)#class-map match-any Voice-Calls
R1-PBX(config-cmap)#description [---[ Polaczenia RTP ]---]
R1-PBX(config-cmap)#match ip dscp ef
R1-PBX(config-cmap)#match ip precedence 5
R1-PBX(config-cmap)#class-map match-any Voice-Control
R1-PBX(config-cmap)#description [---[ Sterowanie polaczeniem ]---]
R1-PBX(config-cmap)#match ip dscp af31
R1-PBX(config-cmap)#match ip precedence 3
```

Tak, wiersze `match` w każdej z klas są nadmiarowe, ale nie mogłem się powstrzymać. Będę już tak robił zawsze. Trudno pozbyć się starych nawyków. Następne zadanie polega na utworzeniu mapy polityki (`policy-map`) odpowiadającej mapom klas. Mapa polityki umożliwi ustanawianie równoległych połączeń G.711 z jednoczesnym zagwarantowaniem pasma (o nadmiernej szerokości) dla połączeń sterujących.

```
R1-PBX(config-cmap)#policy-map Ethernet-QoS
R1-PBX(config-pmap)#description [---[ Reguly przetwarzania pakietow klas ]---]
R1-PBX(config-pmap)#class Voice-Calls
R1-PBX(config-pmap-c)#priority 256 64
R1-PBX(config-pmap-c)#class Voice-Control
R1-PBX(config-pmap-c)#bandwidth 128
R1-PBX(config-pmap-c)#class class-default
R1-PBX(config-pmap-c)#fair-queue
```

Mapy polityki nie można przypisać do podinterfejsu należącego do interfejsu o wielu VLAN-ach. Konieczne jest zdefiniowanie zasad kształtowania ruchu przenoszonego przez interfejs, a następnie wskazanie polityki usług (`service-policy`) w ramach klasy `class-default`. Poniższy listing odpowiada za ograniczenie ruchu wychodzącego do 50 Mb/s (czyli do przepustowości połączenia internetowego) i nałożenie polityki usług:

```
R1-PBX(config)#policy-map SubInt-QoS
R1-PBX(config-pmap)#class class-default
R1-PBX(config-pmap-c)#shape average 50000000
R1-PBX(config-pmap-c)#service-policy Ethernet-QoS
```

Mapę polityki powiążę z interfejsami ethernetowymi:

```
R1-PBX(config-if)#int f0/0
R1-PBX(config-if)#service-policy output Ethernet-QoS
R1-PBX(config-if)#int f0/1
R1-PBX(config-if)#service-policy output Ethernet-QoS
```

Jeśli masz podinterfejsy, możesz powiązać z nimi politykę `SubInt-Qos`:

```
R1-PBX(config-if)#int f0/0.902
R1-PBX(config-subif)#service-policy output SubInt-QoS
```

Nie możesz jednak przypisać polityk usług odnoszących się do tego samego kierunku transmisji zarówno do interfejsu fizycznego, jak i zdefiniowanego w nim podinterfejsu. Router wygeneruje wówczas komunikat o błędzie:

```
R1-PBX(config-if)#service-policy output Ethernet-QoS
Attaching service policy to main and sub-interface in the same direction concurrently is not allowed
//Przypisanie polityki usług do głównego interfejsu i podinterfejsu o tym samym kierunku transmisji jest niedozwolone
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Usługa DHCP dla telefonów

Kolejna rzecz, którą trzeba zdefiniować po stronie routera, to zakres adresów IP przydzielanych przez usługę DHCP telefonom. Pamiętaj, że wszystkie telefony pracują w VLAN-ie 902, czyli w sieci `192.168.2.0/24`. Aby uniemożliwić pobieranie adresów IP należących do routerów i innych systemów sieciowych, zdefiniujmy pulę z zakresem od `.129` – `.222`

```
ip dhcp excluded-address 192.168.2.1 192.168.2.128
ip dhcp excluded-address 192.168.2.223 192.168.2.255
!
ip dhcp pool Phones
   network 192.168.2.0 255.255.255.0
   option 150 ip 192.168.2.1
   default-router 192.168.2.1
   dns-server 192.168.1.200 192.168.1.205
```

Instrukcja `option 150 ip 192.168.2.1` dostarcza telefonom informacji na temat adresu serwera TFTP, z którego mogą pobrać oprogramowanie i pliki konfiguracyjne. W kolejnym podrozdziale uruchomimy usługę TFTP na routerze 2811, więc jako adres serwera podałem adres IP VLAN-u głosowego tego routera. Poprawne byłoby również zastosowanie opcji 66, która pozwoliłaby na podanie nazwy zarejestrowanej w serwerach DNS.

Usługa TFTP

Podczas uruchamiania telefony IP wykorzystują protokół CDP do wymiany z przełącznikiem informacji o tym, jakie jest ich zapotrzebowanie na moc (jeśli są zgodne ze standardem PoE), którego VLAN-u powinny użyć itp. Następnie żądają przydzielenia adresu IP, określenia adresu serwera TFTP, a po zakończeniu procedury przyłączania pobierają z serwera TFTP pewne dodatkowe informacje.

Aby zagwarantować zgodną współpracę urządzeń, możemy zapisać w routerze odpowiednie pliki oprogramowania wbudowanego (firmware) i nakazać telefonom pobranie ich po ustanowieniu połączenia. Takie postępowanie daje gwarancję, że wszystkie urządzenia końcowe pracują pod kontrolą właściwego oprogramowania, co z kolei minimalizuje ryzyko wystąpienia problemów. Samo oprogramowanie jest zależne od modelu telefonu. Konieczne jest więc przygotowanie plików odpowiadających wszystkim typom urządzeń, które dodatkowo muszą być zgodne z wykorzystywaną wersją systemu CME lub UCME. W przypadku opisywanej sieci powinienem udostępnić pliki oprogramowania wbudowanego przeznaczone dla urządzeń Cisco 7970G i 7960G. Pliki są elementami składowymi pakietu CME.

Jeśli budujesz sieć testową, prawdopodobnie będziesz korzystał z wielu telefonów o różnych wersjach oprogramowania. Choć nie powinno być problemów z ich użyciem, warto nauczyć się pobierania i instalowania plików z oprogramowaniem wbudowanym. Firma Cisco utrudniła znacznie to zadanie osobom, które nie mają ważnego kontraktu. Jeśli masz telefony Cisco, ale nie możesz zainstalować w nich stosownych plików, nie przejmuj się. Prawdopodobnie i tak będą działały poprawnie. Chociaż doświadczyłem również pewnych problemów wynikających z niezgodności oprogramowania.

Nazwy plików nie zawsze są jednoznaczne. Na przykład plik oprogramowania wykorzystywany przeze mnie w urządzeniu 7960G ma nazwę *P00308010200*. Ponadto do każdego modelu

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

przypisanych jest wiele plików. Po załadowaniu odpowiednich plików do routera trzeba skonfigurować serwer TFTP tak, aby udostępniał je telefonom. Oto pliki przeznaczone dla modeli 7960G i 7970G:

```
tftp-server flash:apps70.9-1-1TH1-16.sbn
tftp-server flash:cnu70.9-1-1TH1-16.sbn
tftp-server flash:cvm70sccp.9-1-1TH1-16.sbn
tftp-server flash:dsp70.9-1-1TH1-16.sbn
tftp-server flash:jar70sccp.9-1-1TH1-16.sbn
tftp-server flash:SCCP70.9-1-1SR1S.loads
tftp-server flash:term70.default.loads
tftp-server flash:term71.default.loads
tftp-server flash:OS79XX.TXT
tftp-server flash:P00308010200.loads
tftp-server flash:P00308010200.bin
tftp-server flash:P00308010200.sbn
tftp-server flash:P00308010200.sb2
```

Jeśli pobierzesz pakiet CME/UCME z witryny Cisco Connection Online (CCO), wraz z plikami oprogramowania otrzymasz skrypt, który generuje odpowiednie polecenia serwera TFTP. Warto skorzystać z tego rozwiązania szczególnie w sieci firmowej — gdy będzie trzeba dodać nowy telefon Cisco, nie będziesz musiał szukać odpowiednich plików. Co prawda pamięć `flash:` i plik konfiguracyjny zostają zapełnione niepotrzebnymi wpisami, ale dzięki temu w przyszłości zaoszczędzisz trochę czasu.

Upewnij się, że wszystkie pliki zostały załadowane. Z dalszego opisu dowiesz się, skąd telefony „wiedzą", który z nich pobrać.

Usługa telefoniczna

System CME jest konfigurowany za pomocą polecenia `telephony-service`. Sama parametryzacja może być wykonana na dwa sposoby — ręcznie lub za pomocą skryptu:

```
R1-PBX(config)#telephony-service ?
setup    Start setup for Cisco Call Manager Express. Please refer to
         www.cisco.com/univercd/cc/td/doc/product/access/ip_ph/ip_ks/index.htm
         for full documentation.
         //Uruchomienie konfiguracji Cisco Call Manager Express. Pełna dokumentacja produktu znajduje się pod adresem
         //www.cisco.com/univercd/cc/td/doc/product/access/ip_ph/ip_ks/index.htm
<cr>
```

Jeżeli testujesz rozwiązanie w warunkach laboratoryjnych, powinieneś choć raz uruchomić skrypt `setup`. Jego zastosowanie w rzeczywistej sieci jest ograniczone, ale warto zobaczyć, jakie ustawienia są przez ten skrypt wprowadzane. Ponieważ należę do osób, które uwielbiają wiersz polecenia, wybiorę wariant trudniejszy.

W ramach trybu konfiguracji usługi telefonicznej wymienimy pliki oprogramowania wbudowanego, które powinny zostać pobrane przez poszczególne telefony. Najpierw jednak poinformujmy menedżera CME o obowiązku automatycznego wygenerowania plików konfiguracyjnych (na podstawie plików oprogramowania wbudowanego i konfiguracji routera). Służy do tego polecenie `create cnf-files`:

```
R1-PBX(config)#telephony-service
R1-PBX(config-telephony)#create cnf-files
Creating CNF files     //Tworzenie plików CNF
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Teraz możemy określić pliki firmware, które powinny zostać pobrane przez poszczególne modele telefonów:

```
R1-PBX(config-telephony)#load 7960-7940 P00308000900
Updating CNF files            //Aktualizacja plików CNF
CNF files updating complete   //Aktualizacja plików CNF zakończona
R1-PBX(config-telephony)#load 7970 SCCP70.9-1-1SR1S
Updating CNF files            //Aktualizacja plików CNF
CNF files update complete     //Aktualizacja plików CNF zakończona
```

Komunikat informujący o aktualizacji plików CNF jest następstwem wykonania polecenia `create cnf-files`. Gdyby polecenie to nie zostało wcześniej wprowadzone, po zatwierdzeniu instrukcji `load` na ekranie nie pojawiłby się żaden dodatkowy komunikat. Kolejność operacji nie jest istotna. Wprowadzenie polecenia `load` jako pierwszego oznaczałoby, że pliki konfiguracyjne zostaną wygenerowane później (po wykonaniu polecenia `create cnf-files`). Jeśli jest taka potrzeba, zawsze można ponownie wygenerować pliki konfiguracyjne, wprowadzając tę samą instrukcję:

```
R1-PBX(config-telephony)#create cnf-files
Creating CNF files     //Tworzenie plików CNF
```

Następnym zadaniem będzie narzucenie pewnych wartości granicznych na system. Ograniczmy maksymalną liczbę telefonów do 16, a liczbę wybieranych numerów do 128.

```
R1-PBX(config-telephony)#max-ephones 16
R1-PBX(config-telephony)#max-dn 128
```

Po co ustalać maksymalną liczbę telefonów? Ograniczenie liczby urządzeń w statycznym środowisku firmowym zapobiega rejestrowaniu nieuprawnionych telefonów. Wierz mi lub nie, ale to jest poważny problem. Jeśli zmniejszysz tę liczbę później, otrzymasz następujący komunikat:

```
R1-PBX(config-telephony)#max-ephones 8
Please remove ephone 9 to 16 by re-configuring or reloading the system!!
//Usuń telefony od 9 do 16, rekonfigurując lub restartując system!!
```

Ustawienie wartości `max-dn` ułatwia opracowanie planu numeracji. Obydwa parametry domyślnie mają wartość zero.

Kolejnym krokiem jest przypisanie systemowi telefonicznemu adresu IP VLAN-u głosowego, który został wydzielony na danym routerze (`int f0/0.902`):

```
R1-PBX(config-telephony)#ip source-address 192.168.2.1 port 2000
CNF files updating complete    //Aktualizacja plików CNF zakończona
```

Ponieważ zmiana ma bardzo duże znacznie dla poprawnego działania telefonów, wprowadzenie instrukcji spowodowało automatyczne zaktualizowanie plików konfiguracyjnych. Również tym razem jest to konsekwencja wykonania wcześniej polecenia `create cnf-files`.

Zajmijmy się teraz zdefiniowaniem komunikatu systemowego. Komunikat ten będzie wyświetlany we wszystkich telefonach. Zazwyczaj obejmuje on po prostu nazwę firmy, ale możesz podać dowolną treść.

```
R1-PBX(config-telephony)#system message GAD Technology
```

Ustawmy strefę czasową telefonów. Domyślnie system telefoniczny posługuje się czasem uniwersalnym pobieranym z routera (niezależnie od strefy czasowej zdefiniowanej w routerze). Wykonanie zadania sprowadza się do wybrania numeru strefy z listy:

```
R1-PBX(config-telephony)#time-zone ?
  <1-53> select timezone name used by IP phones (offset in minutes)
                    //Wybierz strefę czasową obowiązującą w telefonach IP (przesunięcie jest wyrażone w minutach)
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
1 Dateline Standard Time -720
2 Samoa Standard Time -660
3 Hawaiian Standard Time -600
4 Alaskan Standard/Daylight Time -540
5 Pacific Standard/Daylight Time -480
6 Mountain Standard/Daylight Time -420
7 US Mountain Standard Time -420
8 Central Standard/Daylight Time -360
9 Mexico Standard/Daylight Time -360
10 Canada Central Standard Time -360
11 SA Pacific Standard Time -300
12 Eastern Standard/Daylight Time -300
13 US Eastern Standard Time -300
```

W moim przypadku odpowiednim ustawieniem jest 13:

```
R1-PBX(config-telephony)#time-zone 13
```

Upewnij się, że router korzysta z poprawnego serwera NTP lub kilku takich serwerów (więcej informacji na ten temat znajduje się w rozdziale 38.). W przypadku systemu telefonicznego poprawne ustawienie czasu jest bardzo ważne, ponieważ informacje o czasie są wyświetlane w każdym telefonie i prezentowane w raportach. Poza tym wiele osób korzysta z telefonu jak z zegara.

Zastosowany w przykładzie router został wyposażony w karty AIM-CUE, które udostępniają usługę poczty głosowej (w ramach systemu CUE). Konfiguracja platformy CUE wykracza poza ramy tematyczne książki. Niemniej konfiguracja urządzenia PBX powinna uwzględniać numer wewnętrzny umożliwiający połączenie użytkownika z systemem poczty głosowej. Poniższa instrukcja sprawi, że system CUE będzie dostępny pod numerem 222:

```
R1-PBX(config-telephony)#voicemail 222
```

Ponieważ router pełni funkcję centrali abonenckiej, podniesienie słuchawki powinno spowodować wyemitowanie sygnału gotowości do wybrania numeru. Aby wybrać numer spoza planu numeracji zdefiniowanego w centrali, trzeba poprzedzić go odpowiednią cyfrą. Zazwyczaj jest nią 0, ale wiele firm wybiera również cyfrę 9. W drugim przypadku występuje pewien konflikt z numerami alarmowymi. W prezentowanym przykładzie celowo wybierzemy drugą wersję i spróbujemy rozwiązać problem nakładania się numerów.

```
R1-PBX(config-telephony)#secondary-dialtone 9
```

Teraz możemy określić zasady przekazywania rozmów między numerami wewnętrznymi. System CME obsługuje cztery tryby przekazywania rozmów: `blind`, `full-blind`, `full-consult` oraz `local-consult`. W zasadzie najważniejszy jest podział na tryby: `blind` i `consult`.

Przekazywanie typu blind

Wyobraź sobie, że odbierasz rozmowę, która powinna zostać przełączona do Stefana Nudy z księgowości. Musisz zadzwonić do Stefana, którego nie lubisz z powodu jego pasjonujących opowieści o grze na tubie. Jeśli przełączysz rozmowę bez uprzedniego powiadomienia Stefana o tym fakcie, przekażesz połączenie w trybie `blind`.

Przekazywanie typu consult

Wyobraź sobie, że odebrałeś rozmowę do Ani Barbi z marketingu. Od miesięcy szukasz pretekstu, żeby z nią porozmawiać. Teraz masz szansę. Przekazujesz rozmowę na jej numer wewnętrzny, ona odbiera i wymieniacie uprzejmości. Po chwili wsłuchiwania się w jej

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

anielski głos, informujesz ją o rozmowie. Gdy odłożysz słuchawkę, połączenie zostanie przekazane pod nowy numer. Możliwość przedstawienia rozmówcy ujmującym barytonem zwiększa szansę na ponowną rozmowę z Anią w przyszłości.

W omawianym systemie dostępne są następujące opcje:

```
R1-PBX(config-telephony)#transfer-system ?
blind          Perform blind call transfers (without consultation) with single
               phone line using Cisco proprietary method
               //Przełącz w ciemno (bez konsultacji) z wykorzystaniem jednej linii i niestandardowej
               //metody przełączania (firmy Cisco)
full-blind     Perform call transfers without consultation using H.450.2 or
               SIP REFER standard methods
               //Przełącz bez konsultacji z użyciem techniki H.450.2 lub standardowych metod SIP REFER
full-consult   Perform H.450.2/SIP call transfers with consultation using
               second phone line if available, fallback to full-blind if second
               line unavailable. This is the recommended mode for most systems.
               See also 'supplementary-service' commands under 'voice service
               voip' and dial-peer.
               //Przełącz z użyciem techniki H.450.2/SIP z konsultacją, wykorzystując drugą linię, gdy jest dostępna.
               //Użyj trybu full-blind, jeśli druga linia jest niedostępna. To rozwiązanie jest zalecane w przypadku
               //większości systemów. Sprawdź również polecenia 'supplementary-service' w grupie 'voice service voip'
               //oraz polecenie dial-peer
local-consult  Perform call transfers with local consultation using second
               phone line if available, fallback to blind for non-local
               consultation/transfer target. Uses Cisco proprietary method.
               //Przełącz z lokalną konsultacją, używając drugiej linii, gdy jest dostępna. Użyj trybu
               //blind w przypadku nielokalnych urządzeń docelowych. Wykorzystaj metodę firmy Cisco;
```

Zazwyczaj stosuję technikę `full-consult`, ponieważ jego działanie jest zgodne z oczekiwaniem większości użytkowników. Niestety czasami trzeba porozmawiać ze Stefanem, ale w dłuższej perspektywie jest to korzystne rozwiązanie, bo kiedyś w końcu odbierzesz rozmowę do Ani.

```
R1-PBX(config-telephony)#transfer-system full-consult
```

Plan numeracji

Plany numeracji mogą być bardzo proste albo niezwykle skomplikowane. Staraj się, aby były maksymalnie przejrzyste i logiczne. Przedstawione tu przykładowe rozwiązanie jest bardzo proste, ponieważ obejmuje tylko cztery telefony. Każdy użytkownik otrzyma numer wewnętrzny o wartości większej niż 101. Ponieważ w firmie pracują tylko cztery osoby, numer 111 zostanie przypisany linii zewnętrznej. Będzie również służył jako numer skrzynki głosowej dla osób, które nie wiedzą, z kim chcą rozmawiać.

Jeśli budujesz system obejmujący większą liczbę lokalizacji, musisz utworzyć nieco bardziej skomplikowany plan. Pamiętaj o rozwoju firmy. Niezależnie od tego, ile biur ma w danej chwili, z czasem ich liczba może wzrosnąć o rząd wielkości. Jeśli jest ich 10, zaplanuj 100. Jeśli jest ich 130, zaplanuj 1000. Dlaczego? Ponieważ użytkownikom łatwiej będzie zapamiętać długie numery na początku funkcjonowania systemu, niż dostosować się do zmian w przyszłości. Z doświadczenia wiem, że pracownicy szybciej zmieniają system Windows na Mac, niż zapamiętują nowe numery wewnętrzne. Ludzie nie lubią zmian, szczególnie gdy odnoszą się one do telefonii.

Logiczne plany numeracji ułatwiają definiowanie reguł wybierania. Umożliwiają także szybsze zapamiętanie schematu podziału numerów. Jeśli w każdym ze 100 oddziałów pracuje pięciu użytkowników, rozważ użycie sześciocyfrowych numerów wewnętrznych w formacie OOOUUU

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

(litera O odpowiada cyfrze oddziału, natomiast U reprezentuje cyfrę użytkownika). Numer 501002 oznaczałby numer wewnętrzny 2 w oddziale 501. Pracownicy oddziału 501 będą wówczas wybierali po prostu 2, natomiast osoby spoza oddziału będą musiały wprowadzić ciąg cyfr 501002.

Lista użytkowników przykładowego systemu oraz odpowiadające im numery wewnętrzne zostały pokazane poniżej:

```
GAD      101
Lauren   102
Meghan   103
Colleen  104
General  111
```

Nie ma tu niczego skomplikowanego, ale skalowanie może się okazać problemem. Zaproponowane rozwiązanie sprawdzi się w niewielkiej firmie, która nigdy się nie rozrośnie. Co się stanie, gdy do systemu zostanie dołączony 11. użytkownik? Okaże się, że będzie musiał otrzymać numer o wartości większej niż odpowiadająca ogólnej skrzynce pocztowej. Dla użytkowników nie ma to znaczenia, ale administratora może przyprawić o ból głowy podczas definiowania reguł wybierania numerów.

Porty głosowe

Do routera jest przyłączona linia POTS. Za obsługę linii POTS odpowiadają cztery porty karty FXO zainstalowanej w slocie HWIC o numerze 1. Skonfiguruję tylko jeden z nich:

```
R1-PBX(config)#voice-port 0/1/0
R1-PBX(config-voiceport)#connection plar 111
R1-PBX(config-voiceport)#caller-id enable
```

Polecenie `caller-id` nie powinno budzić wątpliwości. Umożliwia ono przekazywanie do telefonów identyfikatora osoby dzwoniącej. Instrukcja `connection plar` odpowiada za przekazywanie rozmów przychodzących pod numer 111. Zauważ, że nie określiłem nigdzie numeru telefonu. Nie jest to potrzebne, ale żeby zachować przejrzystość konfiguracji, możemy zawrzeć numer w opisie portu:

```
R1-PBX(config-voiceport)#description 555-123-4567
```

W rozwiązaniach korporacyjnych najczęściej doprowadzane jest łącze T1 PRI. Niewielkie biuro (opisywane w tym przykładzie) nie potrzebuje jednak 24 linii zewnętrznych, więc zastosowanie linii POTS będzie właściwym rozwiązaniem.

Konfiguracja telefonów

Telefony IP firmy Cisco mogą realizować wiele funkcji, ale obowiązkowo muszą pozwalać użytkownikom na odbieranie rozmów i ustanawianie połączeń. Linie telefonu są powiązane z przyciskami. Oczywiście w telefonie jest mnóstwo przycisków, ale te, które mam na myśli, są ułożone wzdłuż prawej krawędzi wyświetlacza (są to **przyciski linii** zgodnie z dokumentacją Cisco). Każdy z nich jest przypisany do pewnego numeru wewnętrznego, czyli elementu `ephone-DN` pliku konfiguracyjnego. Załóżmy, że do pierwszego przycisku będziemy przypisywali właściciela linii, a do trzech kolejnych pozostałych użytkowników systemu. Wyznaczymy również przycisk powiązany z linią ogólną (`general`). Rozkład przycisków został pokazany na rysunku 30.4. Zauważ, że w lewej dolnej części ekranu wyświetla się komunikat systemowy (`GAD Technology`).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Rysunek 30.4. Telefon Cisco 7970. Przyciski linii są umieszczone po prawej stronie aparatu

Pierwsza czynność polega na zdefiniowaniu linii telefonicznych. Każdej linii odpowiada wpis `ephone-dn`. Instrukcja `ephone-dn` pełni wiele funkcji. Przede wszystkim umożliwia określenie numeru wewnętrznego. Ponadto odpowiada za konfigurację przycisków telefonu. Zobaczmy jej działanie w praktyce.

Ephone-DN

Ponieważ jestem szefem, sobie przypiszę pierwszy numer wewnętrzny. Wykorzystam do tego celu instrukcję `ephone-dn`:

```
R1-PBX(config)#ephone-dn 1 dual-line
R1-PBX(config-ephone-dn)#
```

Słowo kluczowe `dual-line` umożliwia przyjmowanie i ustanawianie połączeń w czasie prowadzenia innej rozmowy. Zazwyczaj włączam funkcję obsługi dwóch rozmów na wszystkich liniach wewnętrznych.

Kolejny krok to określenie numeru wewnętrznego, który został przypisany danemu użytkownikowi. Nie zawsze jest to tylko numer wewnętrzny, ale w tym przypadku tak powinno być. Ponieważ jestem szefem i chcę używać pierwszego numeru wewnętrznego, moja linia otrzyma numer 101:

```
R1-PBX(config-ephone-dn)#number 101
```

Dzięki słowu kluczowemu `secondary` do jednej linii można przypisać większą liczbę numerów:

```
R1-PBX(config-ephone-dn)#number 101 secondary 9735554333
```

W rozważanym przykładzie nie jest to konieczne, więc poprzestaniemy na wpisaniu instrukcji `number 101`.

Dana linia wewnętrzna może zostać opisana za pomocą kilku poleceń: `label`, `description` oraz `name`. Instrukcja `label` (etykieta) wyznacza tekst wyświetlany obok przycisku. Polecenie `descrip↪tion` definiuje nazwę prezentowaną w górnej części wyświetlacza i powiązaną z pierwszym

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

przyciskiem. Z kolei instrukcja `name` ustawia nazwę użytkownika prezentowaną na wyświetlaczach innych telefonów w czasie ustanawiania połączenia z danego telefonu. Na rysunku 30.4 widać, że pierwszy przycisk jest opisany jako `GAD`. Oto dlaczego:

```
R1-PBX(config-ephone-dn)#label GAD
```

W górnej części wyświetlacza (widocznego na rysunku 30.4) znajduje się ciąg `Office`. Jest to nazwa przypisana do pierwszej linii za pomocą polecenia `description`:

```
R1-PBX(config-ephone-dn)#description Office
```

W początkowej fazie konfiguracji zdefiniowałem numer 222 odpowiadający poczcie głosowej. Ustawię więc automatyczne przekazywanie rozmów do systemu poczty głosowej. Zadanie to wymaga także zdefiniowania dwóch wartości czasu — czasu oczekiwania przed przekazaniem, gdy linia jest zajęta, oraz czasu oczekiwania przed przekazaniem, kiedy nikt nie odbiera rozmowy. Obydwie wartości są ustawiane za pomocą instrukcji `call-forward`:

```
R1-PBX(config-ephone-dn)#call-forward busy 222
R1-PBX(config-ephone-dn)#call-forward noan 222 timeout 10
```

Liczba zapisana za słowem kluczowym `tiemout` odpowiada liczbie sekund oczekiwania na odebranie rozmowy. Przełączenie nie musi odnosić się do systemu poczty głosowej. W ten sam sposób można przełączyć rozmowę pod inny numer wewnętrzny.

Konfiguracja pozostałych linii jest podobna. Poniżej znajduje się fragment pliku opisujący wszystkie cztery linie.

```
ephone-dn 1 dual-line
 number 101
 label GAD
 description Office
 name Office
 call-forward busy 222
 call-forward noan 222 timeout 10
!
ephone-dn 2 dual-line
 number 102
 label Lauren
 description Lauren
 name Lauren
 call-forward busy 222
 call-forward noan 222 timeout 10
!
ephone-dn 3 dual-line
 number 103
 label Meghan
 description Meghan
 name Meghan
 call-forward busy 222
 call-forward noan 222 timeout 10
!
ephone-dn 4 dual-line
 number 104
 label Colleen
 description Colleen
 name Colleen
 call-forward busy 222
 call-forward noan 222 timeout 10
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kolejnym zadaniem jest zdefiniowanie sekcji `ephone-dn` odpowiadającej linii ogólnej (`general`). W opisywanym systemie wykorzystywana jest tylko jedna linia zewnętrzna, która zostanie przypisana do wszystkich telefonów. Dzięki temu po wybraniu numeru firmy wszystkie telefony zaczną dzwonić (podświetlą się też odpowiednie przyciski).

Jak już wcześniej wspomniałem, telefonię IP można konfigurować na nieskończenie wiele sposobów. Być może prezentowane rozwiązanie nie jest właściwe dla Twojego systemu. Każda instalacja jest przecież inna. Zachęcam więc wszystkich do poszukiwania informacji na temat mechanizmów VoIP i samodzielnego testowania ich funkcji.

Zakładam, że numer linii zewnętrznej to 111. Zachowanie zgodności z planem numeracji jest w tym przypadku szczególnie ważne, o czym świadczą dalsze przykłady.

```
ephone-dn 10 dual-line
 number 111
 label Outside-Line
 name Outside Line
 call-forward busy 222
 call-forward noan 222 timeout 15
```

Z doświadczenia wiem, że nawet jeśli do firmy jest doprowadzona tylko jedna linia POTS, trzeba ją skonfigurować jako linię podwójną (`dual-line`). W przeciwnym przypadku, gdy ktoś z pracowników będzie łączył się z pocztą głosową za pośrednictwem podanego numeru wewnętrznego, linia będzie zajęta. Mimo że nie jest ustanowione żadne połączenie z siecią zewnętrzną. Rozmowy przychodzące będą wówczas przełączane do systemu poczty głosowej zgodnie z ustawieniem `call-forward busy` lub wyemitowany zostanie sygnał zajętości (w przypadku braku instrukcji). Włączenie trybu `dual-line` powoduje, że nadejście połączenia (w chwili gdy użytkownik jest połączony z pocztą głosową) zostanie zasygnalizowane dzwonkiem i wyświetleniem identyfikatora rozmówcy.

Przywoływanie

Przywoływanie jest użyteczną funkcją, którą łatwo się konfiguruje. Wystarczy określić specjalny numer, którego wybranie spowoduje uaktywnienie linii we wszystkich telefonach przygotowanych do odbierania takich rozmów. Po uaktywnieniu linii w telefonie zostanie włączony głośnik, co umożliwi dzwoniącemu przekazanie informacji grupie osób. Choć można w ten sposób zdefiniować większą liczbę grup, rozwiązanie przykładowe ograniczymy do jednej grupy obejmującej wszystkie telefony.

Cała operacja sprowadza się do utworzenia sekcji `ephone-dn` i zapisania w niej specjalnego polecenia `paging ip`. Zadaniem instrukcji jest wyznaczenie adresu multiemisji oraz numeru portu, które będą wykorzystywane do przywoływania. Poniższe rozwiązanie aktywuje system przywoływania po wybraniu numeru 555.

```
R1-PBX(config)#ephone-dn 20
R1-PBX(config-ephone-dn)#number 555
R1-PBX(config-ephone-dn)#description Przywolanie wszystkich telefonow
R1-PBX(config-ephone-dn)#paging ip 239.1.1.10 port 2000
```

W kolejnym podrozdziale została przedstawiona metoda przyłączania telefonów do grupy przywoławczej.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ephone

Po sparametryzowaniu linii wewnętrznych i linii zewnętrznej trzeba się zająć konfiguracją samego telefonu. Ustawienia telefonu ethernetowego są definiowane za pomocą polecenia `ephone`. Podobnie jak w przypadku instrukcji `ephone-dn` konieczne jest określenie identyfikatora aparatu (dzięki temu powiązania między telefonami są przejrzyste).

```
R1-PBX(config)#ephone 1
R1-PBX(config-ephone)#description GAD
```

Wyróżnikiem telefonu jest jego adres MAC zapisany na naklejce przyklejonej do spodniej części aparatu. Oczywiście można go też ustalić za pomocą różnych narzędzi sieciowych:

```
R1-PBX(config-ephone)#mac-address 0019.AA96.D096
```

Przypisanie telefonu do utworzonej uprzednio grupy przywoławczej sprowadza się do wpisania instrukcji `paging-dn` wraz z identyfikatorem tej grupy:

```
R1-PBX(config-ephone)#paging-dn 20
```

Określenie typu telefonu pozwoli systemowi na wybranie odpowiednich plików oprogramowania wbudowanego, które zostaną dostarczone do aparatu w czasie rejestracji:

```
R1-PBX(config-ephone)#type 7971
```

Nadszedł w końcu czas, aby zająć się przyciskami telefonu. Najprostsza forma konfiguracji polega na podaniu numeru przycisku i powiązaniu go z identyfikatorem `ephone-dn`. Zapis `1:1` oznacza powiązanie numeru wewnętrznego zdefiniowanego w sekcji `ephone-dn 1` z przyciskiem o numerze `1`. Przyciski są ponumerowane kolejno od góry do dołu. Liczba przycisków przeznaczonych do konfiguracji zależy od modelu aparatu. W urządzeniu 7971G jest osiem przycisków, z których wykorzystam pięć.

Polecenie konfiguracji przycisku ma wiele ciekawych opcji. Oto rezultat wyświetlenia podpowiedzi:

```
R1-PBX(config-ephone)#button ?
  LINE button-index:dn-index pairs example 1:2 2:5   //LINIA pary indeks-przycisku:index-dn, przykład: 1:2 2:5
Configuration line:button with separator feature options:
                                                     //Ustawienie linia:przycisk z opcjami w formie speratorów:
: normal phone lines                                 //standardowe linie telefoniczne
       example    button 1:2 2:5                     //przykład
s silent ring, ringer muted, call waiting beep muted
                          //ciche dzwonienie, dzwonek wyciszony, wyciszone powiadomienie o rozmowie oczekującej
       example    button 1s2 2s5                     //przykład
b silent ring, ringer muted, call waiting beep not muted
                        //ciche dzwonienie, dzwonek wyciszony, niewyciszone powiadomienie o rozmowie oczekującej
       example    button 1b2 2b5                     //przykład
f feature ring                                       //opcje dzwonka
       example    button 1f2 2f5                     //przykład
       see also 'no dnd feature-ring'                //zobacz również 'no dnd feature-ring'
m monitor line, silent ring, call waiting display suppressed
                                //monitorowanie linii, cichy dzwonek, zabronione wyświetlanie rozmowy oczekującej
       example    button 1m2 2m5                     //przykład
       see also 'transfer-system full-consult dss'   //zobacz również 'transfer-system full-consult dss'
w watch line (BLF), watch the phone offhook status via the phone's primary ephone-dn
 //obserwacja linii, monitorowanie stanu widełek słuchawki za pośrednictwem podstawowego identyfikatora ephone-dn telefonu
       example    button 1w2 2w5                     //przykład
o overlay lines, combine multiple lines per physical button
                                                     //nałożenie linii, powiązanie kilku linii z jednym przyciskiem
       example    button 1o2,3,4,5                   //przykład
c overlay call-waiting, combine multiple lines per physical button
                        //nałożenie rozmów oczekujących, powiązanie kilku linii z jednym przyciskiem
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
        example    button 1c2,3,4,5                        //przykład
        see also 'huntstop channel' for ephone-dn dual-line
                        //zobacz również 'hunstop channel' w odniesieniu do instrukcji ephone-dn dual-line
x expansion/overflow, define additional expansion lines that are
                                                          //rozszerzenie/przekazanie, definiuje dodatkowe linie,
        used when the primary line for an overlay button is
                        // które zostaną użyte, gdy linia podstawowa przycisku jest zajęta przez prowadzoną rozmowę
        occupied by an active call
        Expansion works with 'button o' and not with 'button c'
                        //rozszerzenie działa z poleceniem 'button o' i nie działa z poleceniem 'button c'
        example    button 4o21,22,23,24,25                 //przykład
                   button 5x4
                   button 6x4
Different separator options may be used for each button  //każdemu przyciskowi można przypisać różne opcje
        example    button 1:2 2s5 3b7 4f9 5m22 6w10        //przykład
```

W różnych wersjach systemów CME dostępne są różne opcje. To zestawienie pochodzi z oprogramowania CME 8.0. Choć jego treść nie jest szczególnie pasjonująca, wyjaśnia dość dokładnie znaczenie poszczególnych opcji. W dalszej części prezentacji zamierzam wykorzystać tylko dwie z nich (*normal* i *monitor*). Tryb *normal* zapewnia standardową pracę przycisku — rozmowa przychodząca powoduje uruchomienie dzwonka telefonu i podświetlenie przycisku (przy założeniu, że telefon ma podświetlane przyciski).

Przypisanie dwóm telefonom tego samego numeru wewnętrznego prowadzi do ciekawej sytuacji. Jeśli w obydwu urządzeniach zdefiniowane zostanie standardowe działanie przycisków, obydwa telefony będą sygnalizowały przychodzącą rozmowę za pomocą dzwonka. Jeśli przyciski danego aparatu są podświetlane, dzwonieniu towarzyszy miganie pomarańczowej diody. Po odebraniu rozmowy przez inny aparat światło zmienia się na czerwone.

Ustawienie przycisku jako przycisku monitorowania sprawia, że przychodząca rozmowa nie wyzwala dzwonka. Kontrolka zmienia kolor na czerwony, z chwilą gdy słuchawka innego aparatu zostanie podniesiona. Zdarzenia są także odnotowywane na ekranie wyświetlacza. Na przykład, jeśli w skrzynce głosowej znajduje się wiadomość przeznaczona dla danego numeru wewnętrznego, obok przycisku będzie wyświetlona ikona koperty. Przeanalizujmy następujący przykład. Numer wewnętrzny mojej linii to 101. Gdy ktoś wybierze ten numer, mój telefon zacznie dzwonić. Telefony innych osób (monitorujące ten numer) pozostaną wyciszone. Niemniej zostanie na nich wyświetlone powiadomienie o tym, że mój telefon dzwoni.

W każdym telefonie linia główna zostanie przypisana do standardowego przycisku. Pozostałe będą objęte monitoringiem. Na końcu zdefiniujemy przycisk linii ogólnej i skrzynki głosowej. Oto konfiguracja przycisków mojego telefonu:

```
R1-PBX(config-ephone)#button 1:1 2m2 3m3 4m4 5:10
```

Odpowiada ona za sparametryzowanie przycisków zgodnie z rysunkiem 30.4.

System dba o to, by jeden identyfikator `ephone-dn` nie został przypisany do dwóch przycisków. Próba powiązania identyfikatora `3` z przyciskami `1` i `4` kończy się takim komunikatem o błędzie:

```
R1-PBX(config-ephone)#button 1:3 4m3
 duplicate dn [3] assignment: button 4, 3   //zduplikowane powiązanie dn [3]: przycisk 4, 3
   only button 1:3 accepted                 //zaakceptowany jedynie przycisk 1:3
```

Opisaną zasadę konfiguracji trzeba zastosować w odniesieniu do pozostałych telefonów biura.

Telefon 1.

```
ephone 1
 description GAD
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
 mac-address 0019.AA96.D096
 paging-dn 20
 type 7971
 button 1:1 2m2 3m3 4m4 5:10
```

Telefon 2.

```
ephone 2
 description Lauren
 mac-address 001D.4595.7FE5
 paging-dn 20
 type 7970
 button 1:2 2m1 3m3 4m4 5:10
```

Telefon 3.

```
ephone 3
 description Meghan
 mac-address 0012.4362.C4D2
 paging-dn 20
 type 7960
 button 1:3 2m1 3m2 4m4 5:10
```

Telefon 4.

```
ephone 4
 description Colleen
 mac-address 000D.BC50.FC2C
 paging-dn 20
 type 7960
 button 1:4 2m1 3m2 4m3 5:10
```

Każdy telefon będzie dzwonił w wyniku wybrania przypisanego mu numeru wewnętrznego lub w przypadku pojawienia się rozmowy przychodzącej na linii POTS (przycisk `5` jest powiązany z identyfikatorem `ephone-dn 10`).

PTT

Jeśli korzystasz z telefonu bezprzewodowego, takiego jak Cisco 7921G, możesz opisać sposób działania przycisku PTT (ang. *Push To Talk*), tak aby inicjował przywołanie. Identyfikator przywołania został wcześniej skonfigurowany. Oto stosowny fragment kodu:

```
ephone-dn 20
 number 555 no-reg primary
 description Page All Phones
 paging ip 239.1.1.10 port 2000
```

Taka konfiguracja zapewnia przywołanie po wybraniu numeru 555. Aby analogiczna operacja była inicjowana po naciśnięciu przycisku PTT w telefonie 7921G, trzeba powiązać ten przycisk z nową wartością `ephone-dn`, która z kolei będzie się odwoływała do identyfikatora systemu przywołania. Nie jest to proste rozwiązanie, ale nie ja pisałem program.

Najpierw trzeba utworzyć nową sekcję `ephone-dn`, która będzie się odnosiła do sekcji przywołania. Odpowiada za to instrukcja `intercom`, której funkcję można opisać słowami: „gdy ktoś wybierze 556, wybierz 555”.

```
R1-PBX(config)#ephone-dn 21
R1-PBX(config-ephone-dn)# number 556
R1-PBX(config-ephone-dn)# name IP-Paging
R1-PBX(config-ephone-dn)# intercom 555 label "PTT Paging"
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Aby przycisk działał poprawnie w telefonie 7921G, musimy wykorzystać kod XML konfigurujący opcje przycisku PTT. W tym celu konieczne jest wykonanie instrukcji `service phone thumbButton1` w sekcji `telephony-service`. Poniższe polecenia informują telefon o obowiązku wybrania numeru przypisanego do przycisku `6`, gdy naciśnięty zostanie przycisk PTT:

```
R1-PBX(config)#telephony-service
R1-PBX(config-telephony)#service phone thumbButton1 PTTH6
```

Zgodnie z tą techniką przycisk PTT telefonu jest powiązany z przyciskiem `6`. Jeśli dysponujesz większą liczbą telefonów 7921G i nie chcesz stosować tego samego rozwiązania we wszystkich, sprawdź w dokumentacji Cisco, jakie są alternatywne metody konfiguracji.

Ostatnia czynność polega na powiązaniu przycisku `6` z identyfikatorem `ephone-dn 21`:

```
R1-PBX(config)#ephone 2
R1-PBX(config-ephone)#button 6:21
```

Kontrolki oczekujących wiadomości

Mimo że opis nie dotyczy konfiguracji poczty głosowej, obsługujący ją moduł AIM-CUE został zainstalowany i trzeba określić zasady interakcji z nim. Jedną z chętnie wykorzystywanych funkcji poczty głosowej jest funkcja powiadamiania o oczekujących wiadomościach (MWI — *Message Waiting Indicator*).

Mechanizm MWI współdziała z identyfikatorem `ephone-dn`. Oprogramowanie Unity Express wybiera określony numer (dołączony do numeru wewnętrznego) w celu włączenia kontrolki sygnalizacyjnej. Gdy trzeba ją wyłączyć, wybiera inny numer (również powiązany z numerem wewnętrznym). Skonfigurowałem system CUE w taki sposób, aby wybierał numer 8000 w celu włączenia kontrolki oraz 8001 w celu wyłączenia. W instrukcjach występują kropki, których znaczenie zostało opisane w podrozdziale „Reguły wybierania” na stronie 534. Na potrzeby tej analizy wystarczy zapamiętać, że kropka w numerze oznacza „pasuje każda cyfra”.

```
R1-PBX(config)#ephone-dn 40
R1-PBX(config-ephone-dn)#number 8000...
R1-PBX(config-ephone-dn)#mwi on
```

Analogiczna sekcja opisująca wyłączenie kontrolki:

```
R1-PBX(config)#ephone-dn 41
R1-PBX(config-ephone-dn)#number 8001...
R1-PBX(config-ephone-dn)#mwi off
```

Jednym z ciekawszych sposobów na przetestowanie mechanizmu jest wybranie jednego z numerów. Podejdź do telefonu, podnieś słuchawkę i wybierz numer 8000xxx, zastępując *xxx* numerem wewnętrznym telefonu. Kontrolka powinna się zaświecić. Po wybraniu 8001xxx zgaśnie.

Stanowczo odradzam wybieranie kodów MWI przypisanych kolegom z pracy. Spowoduje to wyświetlenie informacji o wiadomości, której nie ma. Przyglądanie się zakłopotaniu i irytacji kolegi z powodu nieistniejącej wiadomości wcale nie jest śmieszne. Naprawdę, nie rób tego.

Między innymi z powodu takiego sposobu działania mechanizmu MWI nie zdecydowałem się na przypisanie linii zewnętrznej numeru wewnętrznego złożonego ze wszystkich dziesięciu cyfr. Gdybym tak zrobił, konieczna byłaby bardziej skomplikowana reguła wybierania lub nawet kilka reguł o różnej długości numeru.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Reguły wybierania

Reguły wybierania (ang. *dial peer*) sprawiają najwięcej kłopotów osobom, które zaczynają pracę z systemami VoIP. Definicje tych reguł powinny być jak najmniej skomplikowane, co niestety wymaga czasu, którego inżynierowie zazwyczaj nie mają. Widziałem horrendalnie skomplikowane reguły liczone w setkach, które po przemyśleniu można było zredukować do mniej niż dziesięciu. Reguły wybierania można postrzegać jako trasy statyczne — im więcej wpisów uda się zagregować, tym lepszy będzie wynik końcowy.

Reguły wybierania są kolejnym elementem technologii VoIP, którego liczba ustawień wydaje się nieskończona. Zawsze istnieje wiele metod realizacji jednego zadania. Jeśli więc nie podoba Ci się rozwiązanie zaproponowane przeze mnie albo sądzisz, że mógłbyś je wykonać lepiej, nie wahaj się. Działaj.

Istnieje wiele rodzajów reguł wybierania, które służą do różnych celów:

```
R1-PBX(config)#dial-peer ?
cor          Class of Restriction                    //Klasa ograniczenia
data         Data type                               //Transmisja danych
hunt         Define the dial peer hunting choice     //Definiowanie wyboru alternatywnego numeru
inbound      Define the inbound options              //Definiowanie opcji rozmów przychodzących
outbound     Define the outbound options             //Definiowanie opcji rozmów wychodzących
search       Define dial peer search service         //Definiowanie usługi wyszukiwania jednostek
terminator   Define the address terminate character  //Definiowanie znaku kończącego adres
voice        Voice type                              //Połączenie głosowe
```

Jedyne opcje spośród wymienionych, którymi jesteśmy zainteresowani, to te dotyczące połączeń głosowych. Reguły połączeń głosowych również dzielą się na wiele kategorii, a rodzaj wybranego mechanizmu zależy od rodzaju łącza:

```
R1-PBX(config)#dial-peer voice 10 ?
mmoip    Multi Media Over IP       //Multimedia w protokole IP
pots     Telephony                 //Telefonia
voatm    Voice over ATM            //Głos w protokole ATM
vofr     Voice over Frame Relay    //Głos w protokole Frame Relay
voip     Voice over IP             //Głos w protokole IP
```

Interesują nas jedynie rozwiązania odnoszące się do technologii POTS i VoIP. Zatem przyjrzyjmy się im bliżej.

POTS

Reguły wybierania związane z technologią POTS odwołują się do linii POTS. Odkrywcze, prawda? W systemie telefonicznym małego biura dostępna jest tylko jedna linia POTS, co nie zmienia faktu, że trzeba wyznaczyć dla niej określone reguły wybierania numerów, żeby można było nawiązywać połączenia z rozmówcami spoza firmy. Zgodnie z wcześniejszym założeniem ustanowienie połączenia wychodzącego wymaga poprzedzenia numeru cyfrą 9. Zacznijmy więc pracę od wpisania jednej z najmniej skomplikowanych definicji:

```
R1-PBX(config)#dial-peer voice 1 pots
R1-PBX(config-dial-peer)# destination-pattern 9T
R1-PBX(config-dial-peer)# port 0/1/0
```

Instrukcja `destination-pattern` (wzorzec dopasowania numerów docelowych) zawiera informację: „dopasuj dziewiątkę, a później rejestruj wszystkie cyfry aż do upłynięcia dopuszczalnego czasu". Wyobraź sobie, że podnosisz słuchawkę i wybierasz numer, aby zamówić 400 egzemplarzy tej

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

książki. W trakcie wybierania numeru okazuje się, że nie pamiętasz czterech ostatnich cyfr. Trzymając cały czas słuchawkę, szukasz numeru w notatkach. System przez pewien czas czeka na kolejne cyfry, ale w końcu zaczyna emitować sygnał zajętości. Oczekiwanie wynikało z uwzględnienia w regule wybierania **przerwy międzycyfrowej** (ang. *interdigit timer*). Parametr ten informuje system o tym, jak długo powinien oczekiwać na wprowadzenie kolejnej cyfry. Litera `T` w regule wybierania oznacza, że numer może mieć dowolną długość.

Czas przerwy jest ustawiany w ramach konfiguracji portu głosowego:

```
R1-PBX(config)#voice-port 0/1/0
R1-PBX(config-voiceport)#timeouts interdigit ?
<0-120> seconds   //<0 – 120> sekund
```

Zostawienie reguły w takiej postaci pozwoliłoby na wykonywanie rozmów wychodzących. Problemem jest jednak to, że połączenie byłoby inicjowane po upływie czasu przerwy międzycyfrowej (która domyślnie wynosi dziesięć sekund). Na załadowanie strony nie czekam dłużej niż dwie sekundy. Nie zamierzam więc czekać dziesięć sekund na zainicjowanie połączenia. Dlatego zmodyfikujemy nieco poprzednią regułę.

Zacznijmy od wymuszenia na użytkowniku wybrania cyfry 9 oraz dziesięciocyfrowego numeru telefonu:

```
R1-PBX(config)#dial-peer voice 100 pots
R1-PBX(config-dial-peer)#destination-pattern 9[2-9].........
R1-PBX(config-dial-peer)#port 0/1/0
R1-PBX(config-dial-peer)#forward-digits 10
```

Na dwie rzeczy warto zwrócić szczególną uwagę. Pierwszą jest instrukcja `destination-pattern` z parametrem `9[2-9].....` (dziewięć kropek). Dlaczego przedział jest zdefiniowany jako 2 – 9? Ponieważ numery kierunkowe w Stanach Zjednoczonych zawsze rozpoczynają się od cyfry z zakresu od 2 do 9. Instrukcja `forward-digits 10` informuje router o obowiązku wysłania jedynie dziesięciu ostatnich cyfr (czyli odrzucenia cyfry 9).

Czytelnikom spoza Stanów Zjednoczonych należą się przeprosiny. Na świecie wykorzystywanych jest tak wiele formatów numerów telefonicznych, że uwzględnienie ich wszystkich w prezentowanych przykładach byłoby niemożliwe. Skoncentrowałem się więc na rozwiązaniach obowiązujących w USA. Oto kilka informacji na ich temat:

- Format numerów telefonicznych to 1+NPA-NXX-NumerAbonenta.
- Numery telefonów są zazwyczaj zapisywane w formacie (NPA) NXX-9999, gdzie 9999 oznacza numer abonenta.
- Cyfra 1 jest często wykorzystywana do zainicjowania połączeń międzymiastowych.
- Zgodnie z wytycznymi ITU cyfra 1 jest również numerem kierunkowym do Stanów Zjednoczonych.
- Człon NPA (kod obszaru) składa się z trzech cyfr, z których pierwsza zawiera się w przedziale od 2 do 9.
- Człon NXX (kod centrali) składa się z trzech cyfr, z których pierwsza zawiera się w przedziale od 2 do 9.
- Cyfra 0 służy do ustanowienia połączenia z operatorem lub do zainicjowania połączenia z pomocą operatora.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Poniżej przedstawiona została reguła POTS obsługująca wybieranie numerów w formacie 9+1+10 cyfr:

```
R1-PBX(config)#dial-peer voice 110 pots
R1-PBX(config-dial-peer)#destination-pattern 91..........
R1-PBX(config-dial-peer)#port 0/1/0
R1-PBX(config-dial-peer)#forward-digits 11
```

A co z numerem alarmowym 911? Trzeba zdefiniować jedną lub dwie reguły wybierania odnoszące się do tego przypadku. Pierwszy z listingów opisuje wybieranie numeru 9+911:

```
R1-PBX(config)#dial-peer voice 120 pots
R1-PBX(config-dial-peer)#destination-pattern 9911
R1-PBX(config-dial-peer)#port 0/1/0
R1-PBX(config-dial-peer)#forward-digits 3
```

Kolejny jest przeznaczony dla osób, które w panice wybierają jedynie 911:

```
R1-PBX(config)#dial-peer voice 130 pots
R1-PBX(config-dial-peer)#destination-pattern 911
R1-PBX(config-dial-peer)#port 0/1/0
R1-PBX(config-dial-peer)#forward-digits all
```

VoIP

Reguły VoIP, jak nietrudno się domyślić, są wykorzystywane w połączeniach ustanawianych z jednostkami dysponującymi adresami IP lub jednostkami o zbliżonym sposobie działania. Interfejsami docelowymi nie mogą być interfejsy ATM lub Frame Relay, dla których są przeznaczone osobne kategorie reguł.

Najprostsza reguła wybierania powoduje ustanowienie połączenia z określonym adresem IP. Posłuży nam ona do opisania odwołań do systemu poczty głosowej. Liczba instrukcji została ograniczona do niezbędnego minimum:

```
dial-peer voice 2 voip
 destination-pattern 222
 session target ipv4:192.168.1.11
 codec g711ulaw
```

Jeśli rozmowa jest kierowana do numeru `222`, należy ją przekazać na adres `192.168.1.11` i zastosować kodek G.711. Oczywiście powyższy listing można uzupełnić o wiele różnych opcji, co będziemy robili, konfigurując mechanizm SIP. Niemniej podstawowa reguła wybierania w ramach systemu VoIP wygląda tak jak na listingu.

Wzorce reguł wybierania

Reguły wybierania są interpretowane zgodnie z precyzyjnymi zasadami, które zostały opisane w różnych źródłach. Niemniej podstawowe informacje na ich temat zostały zamieszczone również w tym rozdziale.

Doskonały poradnik na temat wzorców dopasowania w regułach wybierania (w systemie IOS) nosi tytuł *Understanding Inbound and Outbound Dial Peers Matching on IOS Platforms* i jest dostępny na stronie portalu Cisco pod adresem *http://www.cisco.com/en/US/tech/tk652/tk90/technologies_tech_note09186a008010fed1.shtml*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Reguły połączeń wychodzących

Określenie **połączenia wychodzące** odnosi się do połączeń wykonywanych z konfigurowanego urządzenia. Jedną z podstawowych zasad, o których trzeba pamiętać, definiując reguły wybierania, jest zasada stanowiąca, że **najdłuższe dopasowanie bezpośrednie wygrywa**. Przeanalizujmy dwie reguły:

```
dial-peer voice 3 voip
 destination-pattern 222...
 session target ipv4:10.10.10.10
!
dial-peer voice 4 voip
 destination-pattern 2222..
 session target ipv4:20.20.20.20
```

Gdyby wybierany numer składał się z cyfr 222111, za pasującą uznana zostałaby reguła `3`. W przypadku numeru 222211 lepsza okazałaby się reguła `4`, mimo że ciąg cyfr pasuje także do reguły `3`.

Trzeba jednak pamiętać, że w przypadku niejawnego porównania za najkorzystniejsze uznawane jest najkrótsze dopasowanie. Oto stosowny przykład:

```
dial-peer voice 10 voip
 destination-pattern 2..
 session target ipv4:10.10.10.10
!
dial-peer voice 20 voip
 destination-pattern 2......
 session target ipv4:20.20.20.20
```

Numer 2111111 zostanie dopasowany do reguły `10`, choć wydawałoby się, że reguła `20` będzie odpowiedniejsza.

Reguły połączeń przychodzących

Reguły wybierania dla połączeń przychodzących są weryfikowane w następującej kolejności:

1. Wybrany numer jest porównywany z instrukcją `incoming called-number`.
2. Numer dzwoniący jest porównywany z instrukcją `answer-address`.
3. Numer dzwoniący jest porównywany z instrukcją `destination-pattern`.
4. Sprawdzana jest konfiguracja portu.
5. Stosowana jest domyślna reguła wybierania o numerze 0.

Przykłady przedstawione w tym rozdziale są bardzo ogólne, więc wykorzystana została tylko instrukcja `destination-pattern`. Dzięki niej wszystkie rozmowy przychodzące są automatycznie przekazywane na jeden numer wewnętrzny. Odpowiada za to polecenie `plar 111` portu głosowego.

SIP

Protokół SIP stał się bardzo popularny z wielu powodów. Jednym z nich z pewnością jest to, że ma on charakter otwartego standardu. Wielu operatorów telekomunikacyjnych oferuje obecnie połączenia SIP typu trunk, które zaspokajają potrzeby telefoniczne firm w ramach istniejących łączy internetowych.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Urządzenia końcowe SIP mogą pełnić funkcje jednostki klienckiej, serwera lub obydwu jednocześnie. Jak nietrudno się domyślić, telefon IP jest zazwyczaj klientem SIP. Klientem jest również zbudowany na bazie oprogramowania CME system PBX, ponieważ aby realizować połączenia wychodzące, musi zarejestrować się na serwerze utrzymywanym w sieci dostawcy usług internetowych. Telefon SIP może również pełnić funkcję serwera. W ramach systemu SIP wyróżnia się cztery rodzaje serwerów:

Serwer pośredniczący
: Jeśli klient nie dysponuje adresem odbiorcy, serwer ten (często określany jako proxy SIP) wysyła żądania w jego imieniu. Serwer pośredniczący pełni rolę skomputeryzowanego dyspozytora sieci telefonicznej.

Serwer przekierowania
: Serwery przekierowania odsyłają do jednostek klienckich informacje o przeniesieniu odbiorcy lub usunięciu go z sieci.

Serwer rejestracji
: Telefony (jednostki klienckie) muszą się rejestrować, a zadaniem serwera rejestracji jest przetwarzanie żądań rejestracji. W rozwiązaniach IP PBX firmy Cisco telefony rejestrują się z użyciem adresów MAC. Z kolei jednostki klienckie SIP rejestrują adresy IP.

Serwer lokalizacji
: Przetwarzaniu żądania rejestracji przez serwer rejestracji towarzyszy zapisanie w serwerze lokalizacji położenia jednostki klienckiej.

Zazwyczaj wszystkie wymienione usługi są implementowane w formie pojedynczego serwera w celu ułatwienia pracy użytkownikom końcowym. Jeśli nie budujesz własnego systemu SIP, najprawdopodobniej będziesz miał styczność jedynie z serwerem pośredniczącym. W przeciwnym przypadku będziesz musiał wprowadzić kilka adresów serwerów SIP (tak jak to zostało przedstawione w dalszej części rozdziału), które zazwyczaj będą miały tę samą wartość.

Ustanawianie połączenia SIP nie jest skomplikowaną operacją. Oto lista komunikatów, które są wymieniane w czasie tej procedury:

REGISTER
: Zarejestrowanie lokalizacji klienta w serwerze.

INVITE
: Komunikat SIP `INVITE` jest wysyłany podczas inicjowania połączenia z telefonem zdalnym.

ACK
: Potwierdzenie odebrania komunikatu SIP.

CANCEL
: Przerwanie operacji `INVITE` przed ustanowieniem połączenia (odpowiednik odłożenia słuchawki w czasie, gdy telefon po drugiej stronie ciągle dzwoni).

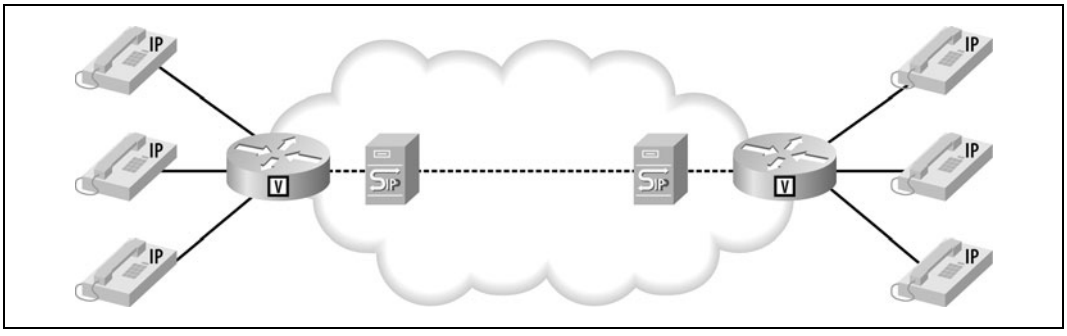
INFO
: Sygnalizacja w czasie trwania rozmowy.

OPTIONS
: Pobranie wykazu funkcji serwera SIP.

BYE
: Zakończenie połączenia. Gdy rozmówca odkłada słuchawkę, telefon wysyła komunikat `BYE`.

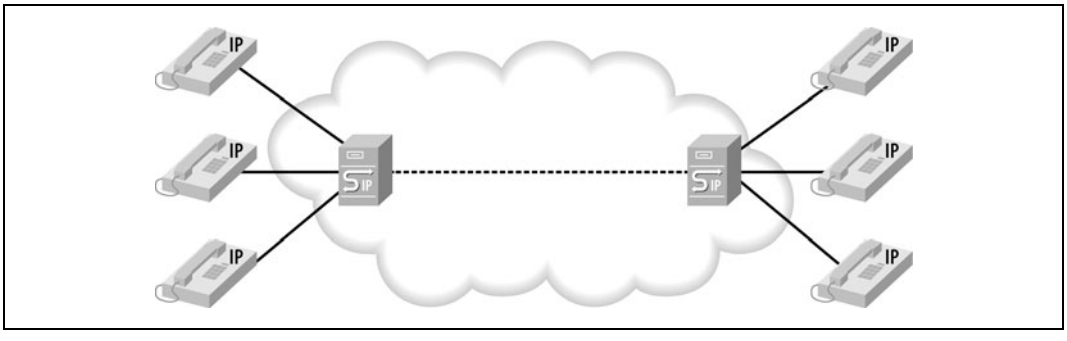
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Na rysunku 30.5 przedstawiono sposób połączenia dwóch małych biur, z których każde jest wyposażone w system Cisco CME i dysponuje połączeniem trunkowym SIP zapewnianym przez dostawcę usług internetowych. Każdy z telefonów rejestruje się w lokalnym routerze CME za pomocą protokołu SCCP. Każdy router rejestruje klienta SIP w serwerze SIP dostawcy usług internetowych.



Rysunek 30.5. Dwa biura z systemami CME i łączami SIP

Wielu dostawców usługi SIP oferuje również w pełni funkcjonalne centrale PBX, co eliminuje konieczność instalowania systemu CME w routerze PBX w sieci firmowej. Z tego powodu coraz więcej firm ogranicza etaty dla takich osób jak my. Przykład konfiguracji SIP niewymagającej systemu CME został pokazany na rysunku 30.6.



Rysunek 30.6. Dwa biura korzystające z telefonów SIP bez systemu CME

W takim rozwiązaniu telefony rejestrują się bezpośrednio w serwerze pośredniczącym dostawcy usługi SIP. Ponieważ system CME nie jest wykorzystywany, protokół SCCP okazuje się zbędny.

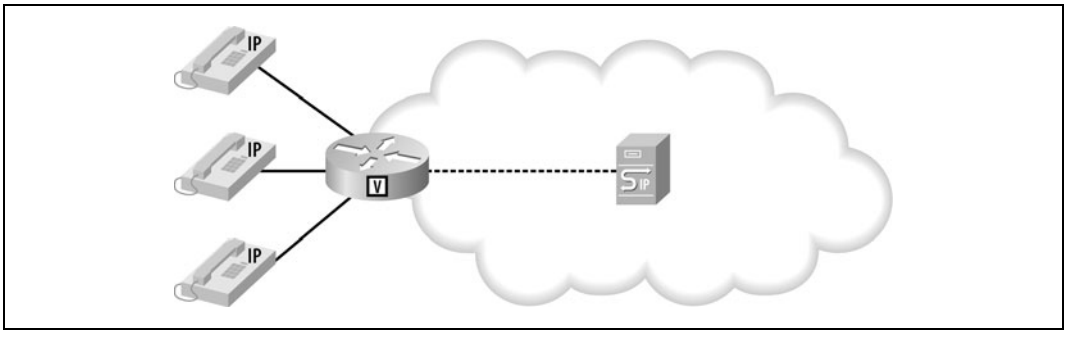
Wszystkie funkcje PBX (od obsługi numerów wewnętrznych po pocztę głosową) są realizowane przez urządzenia operatora. Konfiguracja urządzeń końcowych może zostać przeprowadzona

Jeśli budujesz lub testujesz sieć bazującą całkowicie na protokole SIP i korzystasz z telefonów IP firmy Cisco, konieczne może się okazać zapisanie w nim specjalnego oprogramowania wbudowanego (firmware'u) przeznaczonego do obsługi mechanizmu SIP. Więcej informacji na ten temat znajduje się na stronach portalu Cisco. Wiele nowych telefonów Cisco ma zaimplementowaną obsługę protokołu SIP. Pamiętaj jednak, żeby to sprawdzić.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

za pośrednictwem przeglądarki internetowej przez niewyszkolonych użytkowników. Jeśli korzystasz z usług firm takich jak Vonage lub Magic Jack, Twoja sieć jest elementem opisywanego modelu. Takie rozwiązanie spodobało się wielu menedżerom, co potwierdzają dostawcy usługi SIP, którzy działają na rynku usług telekomunikacyjnych od kilku lat. Znam firmy z tej branży, które zatrudniają po 100 osób.

Prezentowany w tym rozdziale przykład zostanie uzupełniony o łącze trunkowe SIP między chmurą operatora i systemem CME. Załóżmy, że otrzymaliśmy jedno lub dwa takie łącza i trzeba skonfigurować system telefonii tak, aby wykorzystywał je do przenoszenia rozmów przychodzących i wychodzących. Sieć wynikowa powinna odpowiadać przedstawionej na rysunku 30.7.



Rysunek 30.7. Łącze trunkowe SIP obsługujące linie zewnętrzne

Pierwsza czynność polega na pozyskaniu od dostawcy informacji na temat samych łączy trunkowych. Potrzebne dane to: numer telefonu, nazwa użytkownika oraz hasło. Nazwa użytkownika i hasło są niekiedy określane jako **nazwa uwierzytelniająca** (ang. *authname*) i **klucz** (ang. *key*). Dwa łącza trunkowe SIP opisywane w przykładzie mają następujące parametry:

Dane serwera pośredniczącego:
SIP-Proxy: sbc.sip-provider.com

Łącze 1:
Numer: 9735554222
Nazwa uwierzytelniająca: 557333333
Klucz: 123456789

Łącze 2:
Numer: 9735554333
Nazwa uwierzytelniająca: 608222222
Klucz: 123456789

Uzyskane informacje pozwolą urządzeniom na uwierzytelnienie się w serwerze pośredniczącym podczas rejestracji łączy przeznaczonych do przekazywania połączeń przychodzących i wychodzących.

Każdy z dostawców usługi SIP może korzystać z innego sposobu przyłączania sieci firmowej. Podstawowe zasady funkcjonowania telefonii internetowej i protokołów SIP są niezmienne. Jednak nie wszystkie omawiane tutaj rozwiązania nadają się do zastosowania w dowolnym łączu SIP. Niemniej przeanalizowanie przykładowej sieci ułatwia ustalenie, jakie elementy powinny zostać zmienione, aby dany system telefoniczny współdziałał z usługami dostawcy.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Musimy się zająć tylko kilkoma sekcjami konfiguracji mechanizmu SIP. Pierwsza z nich to `vois service voip`:

```
voice service voip
 allow-connections h323 to h323
 allow-connections h323 to sip
 allow-connections sip to h323
 allow-connections sip to sip
 no supplementary-service sip moved-temporarily
 no supplementary-service sip refer
```

Jeśli budujesz użytkowy system na podstawie informacji zamieszczonych w tym rozdziale, upewnij się, że wyłączyłeś reguły wybierania skojarzone z portem głosowym i linią POTS. Mogą one przeszkadzać w bieżącej konfiguracji (konfiguracji mechanizmu SIP).

Początkowe polecenia są dość oczywiste. Przedostatnia instrukcja odpowiada za wyłączenie wysyłania komunikatów SIP 302 `moved temporarily` (tymczasowo przeniesiony), dzięki czemu przekazywanie rozmów w systemie CME działa lepiej. Ostatnie polecenie odnosi się do systemu CME i powoduje wysyłanie komunikatu `SIP REFER` podczas przekazywania rozmów w ramach CME. Ponieważ większość dostawców usługi SIP nie obsługuje komunikatu `SIP REFER`, ustawienie to wymusza operację hairpinningu rozmowy.

Zapewnienie podstawowych funkcji połączeń SIP wymaga zrealizowania trzech zadań: poinformowania systemu CME o lokalizacji serwera pośredniczącego SIP, uwierzytelnienia rozmów wychodzących oraz zarejestrowania numerów, tak aby dostawca SIP wiedział, gdzie ma przekazywać rozmowy kierowane do firmy. Wszystkie działania należy wykonać w trybie konfiguracyjnym `sip-ua`. Najpierw zajmijmy się serwerem pośredniczącym:

```
R1-PBX(config)#sip-ua
R1-PBX(config-sip-ua)#registrar dns:sbc.sip-provider.com expires 3600
R1-PBX(config-sip-ua)#sip-server dns:sbc.sip-provider.com
```

Zwróć uwagę na to, że serwery SIP są zdefiniowane za pomocą nazw domenowych. Można je także zapisać, używając adresów IPv4, IPv6 lub nazw DNS. Jeśli wpiszesz niepoprawną wartość adresu, zobaczysz na ekranie wyjątkowy komunikat o błędzie:

```
R1-PBX(config-sip-ua)#registrar gad
Malformed SIP Registrar address. Valid formats are:   //Błędny adres serwera rejestracji SIP. Poprawne formaty to:
^((dns:.*)|(ipv4:[0-9]?[0-9]?[0-9]?\.[0-9]?[0-9]?[0-9]?\.[0-9]?[0-9]
?[0-9]?\.[0-9]?[0-9]?[0-9]?(:[0-9]+)?)|(ipv6:\[([0-9A-Fa-f.:])+\](:[0-9]+)?))$
```

Nie ma w nim nic strasznego, jeśli umiesz się posługiwać wyrażeniami regularnymi. Jednak osoby, które nie znają wyrażeń regularnych, mogą się czuć zakłopotane. Wykorzystanie nazw domenowych wiąże się z obowiązkiem zdefiniowania adresów serwerów DNS:

```
R1-PBX(config)#ip name-server 192.168.1.200
R1-PBX(config)#ip name-server 192.168.1.205
```

Jeśli podobnie jak ja wyłączasz odwzorowania DNS na routerach, żeby nie marnować czasu, gdy system stara się zamienić błędnie wpisane polecenia na adresy IP, powinieneś przywrócić działanie tej funkcji:

```
R1-PBX(config)#ip domain-lookup
```

Poniższe polecenie umieszcza nazwę zawartą w poleceniu `registrar` w **nagłówku przekierowania** protokołu SIP wykorzystywanym podczas przekazywania rozmów. Standardowo system CME używa do tego celu nazwy domenowej routera:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
R1-PBX(config-sip-ua)#host-registrar
```

Kolejnym zadaniem jest konfiguracja uwierzytelniania połączeń wychodzących:

```
R1-PBX(config-sip-ua)#authentication username 608222222 password 123456789
```

Zauważ, że nigdzie nie zostały wymienione numery telefonów. To było moje pierwsze spostrzeżenie, gdy uruchomiłem protokół SIP w systemie CME. Oczywiście u innych operatorów SIP konfiguracja może wyglądać inaczej. Numery telefonów mogą na przykład pełnić rolę nazw użytkownika.

Konieczne jest również zapisanie danych uwierzytelniających, które umożliwią systemowi CME zarejestrowanie się w sieci dostawcy:

```
R1-PBX(config-sip-ua)#credentials username 608222222 password 123456789 realm voip.sip-provider.com
```

W tym momencie prace nieco się komplikują. Dysponujemy dwoma łączami trunkowymi SIP, więc trzeba zarejestrować obydwa. Dzięki temu rozmówcy będą mogli dzwonić pod obydwa numery. Spróbujmy zarejestrować drugie z nich:

```
R1-PBX(config-sip-ua)#credentials username 557333333 password 123456789 realm voip.sip-provider.com
```

Nie ma błędów!

Na czym więc polega problem? Rozwiązanie to działa poprawnie jedynie w systemach IOS 15.0(1)XA5 i późniejszych, co oznacza, że korzystam z oprogramowania CME 8.0. W czasie testów aktualizowałem czterokrotnie oprogramowanie routera, aby zainstalować funkcje niezbędne do prawidłowej pracy protokołu SIP. Rozpocząłem projekt z systemem CME w wersji 3.3. Oprogramowanie działało poprawnie aż do czasu skonfigurowania więcej niż jednego łącza trunkowego. Żeby ustawienie zadziałało właściwie, zmieniałem wersje z 3.3 przez 4.0 i 7.1 aż do 8.0. Największym problemem okazało się zarejestrowanie dwóch łączy trunkowych SIP od tego samego dostawcy. Nie jest to typowe ustawienie, ponieważ nic nie stoi na przeszkodzie, aby w ramach jednego łącza trunkowego realizowanych było wiele połączeń przychodzących i wychodzących (chyba że w danej firmie potrzebnych jest kilka numerów telefonicznych dla rozmów przychodzących). Znacznie częściej dwa łącza trunkowe SIP są prowadzone do różnych dostawców, gdyż zwiększa to niezawodność systemu.

Gdyby łącza pochodziły od różnych dostawców, pracowałyby w różnych domenach (ang. *realm*).

Jedną z zalet uruchomienia oprogramowania CME 8.0 lub jego późniejszej wersji jest możliwość skorzystania ze słowa kluczowego `number` w instrukcji `credentials`. Zauważyłem jednak pewien problem w komunikacji z siecią dostawcy. Gdy dodałem numery telefonów, łącza nie rejestrowały się poprawnie. Musiałem wpisać nazwę użytkownika w polu numeru, żeby mechanizm zadziałał.

```
R1-PBX(config-sip-ua)#credentials number 608222222 username 608222222 password 123456789 realm
↪voip.sip-provider.com
R1-PBX(config-sip-ua)#credentials number 557333333 username 557333333 password 7 123456789 realm
↪voip.sip-provider.com
```

Ponownie przypominam, że różni dostawcy mogą rozwiązywać ten problem w różny sposób. Niekiedy konieczne jest przetestowanie kilku ustawień i wybranie najlepszego w danej sytuacji.

Aby sprawdzić, czy dane uwierzytelniające zostały zaakceptowane, wystarczy wykonać polecenie `show sip-ua register status`:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
R1-PBX#sho sip-ua register status
Line         peer         expires(sec) registered  P-Associated-URI
============ ============ ============ =========== ================
101          20001        1857         yes
102          20002        1857         yes
103          20003        1857         yes
104          20004        1857         yes
557333333    -1           2340         yes
608222222    -1           2145         yes
```

System został zaprojektowany w taki sposób, aby łącza trunkowe SIP były wykorzystywane jedynie podczas komunikacji z jednostkami poza biurem. Poszczególne telefony nie muszą się więc rejestrować na serwerach dostawcy. Zgodnie z domyślnymi ustawieniami będą jednak próbowały wykonać to zadanie i w zależności od konfiguracji sieci dostawcy SIP oraz rodzaju wykupionej usługi może im się to udać. Aby nie dopuścić do takiej sytuacji, wprowadziłem w sekcji `ephone-dn` instrukcję uniemożliwiającą rejestrację w ramach systemu SIP. Zakaz rejestracji numerów wewnętrznych na serwerach dostawcy usługi SIP zapewnia słowo kluczowe `no-reg` polecenia `number`. Dodane ustawienie odnosi się jedynie do połączeń SIP. Standardowa rejestracja SCCP działa bez przeszkód:

```
R1-PBX(config-ephone-dn)#ephone-dn 1 dual-line
R1-PBX(config-ephone-dn)#number 101 no-reg
```

Jeśli zdecydujesz się na przypisanie numerowi wewnętrznemu numeru SIP, możesz wykluczyć z rejestracji tylko ten numer. Nie wybrałem takiego rozwiązania z powodów opisanych wcześniej i związanych z mechanizmem MWI. Niemniej jeśli chcesz wdrożyć podobną konfigurację, wypróbuj następujące ustawienia:

```
R1-PBX(config-ephone-dn)#ephone-dn 1 dual-line
R1-PBX(config-ephone-dn)#number 101 secondary 9735554333 no-reg primary
```

Po czterokrotnym powtórzeniu czynności w odniesieniu do kolejnych numerów wewnętrznych informacje statusowe o rejestracji SIP stały się znacznie bardziej przejrzyste.

```
R1-PBX#sho sip-ua register status
Line         peer         expires(sec) registered  P-Associated-URI
============ ============ ============ =========== ================
557333333    -1           2340         yes
608222222    -1           2145         yes
```

W dalszej części rozdziału przedstawiłem kilka timerów i liczników powtórzeń, które sprawdziły się w mojej sieci. Wszystkie są związane z wymianą komunikatów SIP, a ich przeznaczenie nie powinno budzić wątpliwości.

```
R1-PBX(config-sip-ua)#retry invite 2
R1-PBX(config-sip-ua)#retry response 3
R1-PBX(config-sip-ua)#retry bye 2
R1-PBX(config-sip-ua)#retry cancel 2
```

W przykładach konfiguracji systemu CME publikowanych w internecie wykorzystywane są różne wartości timerów. Może się wydawać dziwne, że wiele z komunikatów SIP trzeba powtarzać, ale nie wynika to z utraty pakietów. Często analizowałem przypadki, w których zaproszenie SIP (`INVITE`) wracało jako „nieautoryzowane", a powtórzona wersja komunikatu była akceptowana bez żadnych zastrzeżeń.

```
R1-PBX(config-sip-ua)#retry register 10
R1-PBX(config-sip-ua)#timers connect 100
R1-PBX(config-sip-ua)#timers register 250
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

SIP i NAT

Protokół SIP nie zawsze poprawnie współpracuje z mechanizmem NAT. W komunikacie SIP uwzględniony jest adres IP urządzenia nadawczego. Ponieważ często jest to adres prywatny, urządzenie odbiorcze nie może go wykorzystać w procesie routingu.

Aby umożliwić korzystanie z mechanizmu NAT w telefonii SIP, wystarczy wpisać następujące polecenia systemu CME:

```
R1-PBX(config)#sip-ua
R1-PBX(config-sip-ua)#nat symmetric role active
R1-PBX(config-sip-ua)#nat symmetric check-media-src
R1-PBX(config-sip-ua)#connection-reuse
```

Polecenie `connection-reuse` sekcji `sip-ua` jest ukrywane w każdej wersji systemu IOS (włącznie z 15.0(1)XA5), z której korzystałem.

Oto cała konfiguracja sekcji `sip-ua`:

```
sip-ua
 credentials number 608222222 username 608222222 password 7 023510541B4B3B294D5A48 realm
 ↪voip.sip-provider.com
credentials number 557333333 username 557333333 password 7 023510541B4B3B294D5A48 realm
 ↪voip.sip-provider.com
 authentication username 608222222 password 7 023510541B4B3B294D5A48
 nat symmetric role active
 nat symmetric check-media-src
 no remote-party-id
 no redirection
 retry invite 2
 retry response 3
 retry bye 2
 retry cancel 2
 retry register 10
 timers connect 100
 timers register 250
 registrar dns:sbc.sip-provider.com expires 3600
 sip-server dns:sbc.sip-provider.com
 connection-reuse
 host-registrar
```

Skoro łącza trunkowe SIP zostały sparametryzowane, możemy się zająć konfiguracją routera. Dzięki temu będzie mógł odbierać i ustanawiać połączenia. Będziemy musieli opracować dwie reguły wybierania dla połączeń VoIP. Najpierw jednak konieczne jest wykonanie kilku czynności przygotowawczych. Staram się we wszystkich połączeniach wykorzystywać kodek G.711, ponieważ zależy mi na wysokiej jakości dźwięku. Jeśli jego użycie nie jest możliwe, stosuję kodowanie G.729r8. Do opisu preferencji służy klasa `voice`. Odwołanie do niej zostanie zawarte w regule wybierania numerów.

```
R1-PBX(config-dial-peer)#voice class codec 1
R1-PBX(config-class)#codec preference 1 g711ulaw
R1-PBX(config-class)#codec preference 2 g729r8
```

Teraz mogę utworzyć regułę wybierania dla przychodzących połączeń SIP:

```
R1-PBX(config)#dial-peer voice 89 voip
R1-PBX(config-dial-peer)#description [ Przychodzace polaczenia SIP ]
```

W regule trzeba zawrzeć informację o tym, że wykorzystywany jest protokół SIP oraz że wszystkie pasujące do niej połączenia powinny być realizowane przy użyciu sekcji `sip-ua`, która została

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

zdefiniowana wcześniej. Może się to wydawać dziwne, ale serwer SIP działający w routerze „wie", w jaki sposób transkodować rozmowę do standardu H.323, więc może wysyłać strumienie danych do telefonów Cisco.

```
R1-PBX(config-dial-peer)#session protocol sipv2
R1-PBX(config-dial-peer)#session target sip-server
```

Następny wiersz wyznacza kryterium dopasowania, które odpowiada każdej przychodzącej rozmowie.

```
R1-PBX(config-dial-peer)#incoming called-number .%
```

To rozwiązanie nie musi się sprawdzać w każdym przypadku. W moim wydawało się działać doskonale do czasu, gdy uruchomiłem obsługę poczty głosowej i mechanizm MWI. Ponieważ moduł AIM-CUE wykorzystuje protokół SIP do wysyłania poleceń MWI, generowane powiadomienia pasowały do reguł wybierania SIP i powodowały dzwonienie telefonu za każdym razem, gdy zmieniał się stan kontrolki. Usunąłem błąd, dodając instrukcję `called-number` ze wzorcem pasującym do każdego numeru dłuższego od numeru MWI. Reguła mechanizmu MWI zawierała wzorzec `800[12]...`.Wystarczyło więc zagwarantować dopasowanie każdego ciągu dłuższego niż siedem cyfr. Ponieważ mój dostawca usług SIP gwarantuje, że wszystkie rozmowy przychodzące mają numer złożony z dziesięciu cyfr, korekta doskonale spełniła swoje zadanie.

```
R1-PBX(config-dial-peer)#incoming called-number ........+
```

W tej samej sekcji trzeba zdefiniować odwołanie do klasy `voice`, która odpowiada za wykorzystanie kodeków G.711 i G.729.

```
R1-PBX(config-dial-peer)#voice-class codec 1
```

Kolejne polecenie umożliwia przekształcenie sygnałów DTMF w **opisowe zdarzenia telefoniczne** (ang. *named telephone events*). W kompresowanych strumieniach rozmów (takich jak G.729) tony DTMF ulegają zniekształceniu. Przedstawiona instrukcja umożliwia natomiast przekształcenie ich w tekstowe wartości liczbowe przesyłane za pomocą pakietów SIP:

```
R1-PBX(config-dial-peer)#dtmf-relay rtp-nte
```

System detekcji aktywności głosowej (VAD — *Voice Activity Detection*) odpowiada za wstrzymywanie wysyłania pakietów, gdy router zarejestruje ciszę po stronie nadawcy. Technika ta pozwala na zaoszczędzenie pasma, ale jednocześnie powoduje u odbiorcy wrażenie, że rozmowa została przerwana. Rozwiązaniem problemu jest sztuczne wygenerowanie po stronie odbiorcy szumu symulującego normalne tło rozmowy. Z założenia staram się unikać mechanizmu VAD, chyba że jest on niezbędny.

```
R1-PBX(config-dial-peer)#no vad
```

Reguła wybierania dla połączeń przychodzących jest skonfigurowana, więc teoretycznie powinniśmy mieć możliwość odbierania rozmów. Niestety tak nie jest. Mimo poprawnego zarejestrowania łączy trunkowych SIP każda próba ustanowienia połączenia kończy się wygenerowaniem komunikatu o „niedostępności numeru". Taka sytuacja z pewnością jest bardzo stresująca dla osób, które dopiero zaczynają pracę z systemami SIP, ale problem można łatwo rozwiązać. Oto wynik wykonania polecenia `debug ccsip message` obejmujący informacje o nieudanej rozmowie (więcej informacji na ten temat znajduje się w podrozdziale „Rozwiązywanie problemów" na stronie 549).

```
Feb 3 19:07:35.084: //-1/xxxxxxxxxxxx/SIP/Msg/ccsipDisplayMsg:
Sent:
SIP/2.0 404 Not Found
Via: SIP/2.0/UDP 208.73.146.95:5060;branch=z9hG4bK6gmsg8300gk1vmsmj4g0.1
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
From: "DONAHUE GARY " <sip:+197355512120192.168.101.113:5060>;tag=gK0a3f2f72
To: <sip:9735554333@192.168.101.100:5060>;tag=3DF80A8-23C0
Date: Thu, 03 Feb 2011 19:07:35 GMT
Call-ID: pcst1296760055008367193100@192.168.201.113
CSeq: 1 INVITE
Allow-Events: telephone-event
Server: Cisco-SIPGateway/IOS-12.x
Reason: Q.850;cause=1
Content-Length: 0
```

Łącze trunkowe SIP zostało zarejestrowane i komunikat SIP `INVITE` dotarł do routera. Jednak mimo wszystko dostawca wysyła do osoby dzwoniącej informację o niedostępności telefonu. Problem wynika z tego, że nie zdefiniowaliśmy żadnej sekcji `ephone-dn`, która zawierałaby dany numer telefonu. Jeśli dodamy wspomnianą sekcję i skojarzymy ją z przyciskiem telefonu, kolejna próba zakończy się wyemitowaniem dźwięku dzwonka.

```
ephone-dn 22 dual-line
 number 9735554333 no-reg primary
 label SIP-Line-1
 name SIP-Line-1
```

Nie chciałem wprowadzać opisanej zmiany z kilku powodów. Po pierwsze, konfiguruję system telefoniczny małego biura o niedużej liczbie numerów wewnętrznych. Wolałbym, aby wszystkie telefony wyświetlały jedynie numery wewnętrzne, bez numeru zewnętrznego. Po drugie, numer telefonu nie pasuje do założonego planu numeracyjnego (w którym wszystkie numery wewnętrzne mają format 1xx). Jeśli utworzę skrzynkę poczty głosowej przeznaczoną dla linii zewnętrznej, kontrolki sygnalizujące nadejście wiadomości nie będą działać.

Po raz kolejny muszę wspomnieć o tym, że zadanie można wykonać na wiele sposobów. W niektórych przypadkach powiązanie przycisku z rzeczywistym numerem telefonu może się okazać właściwym rozwiązaniem. Zachęcam więc do przetestowania stosownej konfiguracji. Podobnie wykorzystanie wcześniejszego przykładu z numerami dodatkowymi pozwoliłoby na odbieranie połączeń SIP w ramach zdefiniowanego numeru wewnętrznego.

Zamiast dodać nową sekcję `ephone-dn` odnoszącą się do łączy trunkowych SIP, postanowiłem utworzyć kolejne numery wewnętrzne.

```
ephone-dn 32 dual-line
 number 132 no-reg primary
 label SIP-4332
 call-forward busy 222
 call-forward noan 222 timeout 15
!
ephone-dn 33 dual-line
 number 133 no-reg primary
 label SIP-4333
 call-forward busy 222
 call-forward noan 222 timeout 15
```

I znowu pojawia się problem. Komunikat SIP `INVITE` jest kierowany pod numery 9735554333 lub 973555422. W jaki sposób rozmowa miałby więc zostać przekazana pod nowe numery wewnętrzne? Rozwiązanie zapewnia funkcja nazywana **translacją połączeń głosowych** (ang. *voice translation*).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Dzięki translacji połączeń głosowych możemy zastąpić dowolny numer połączenia przychodzącego lub wychodzącego. Reguły translacji przypominają wyrażenia regularne. Poniższa instrukcja odpowiada za zamianę dowolnego numeru na wartość 133:

```
R1-PBX(config)#voice translation-rule 11
R1-PBX(cfg-translation-rule)#rule 1 /.*/ /133/
```

Po utworzeniu reguły translacji trzeba ją przypisać do profilu translacji i odpowiednio nazwać:

```
R1-PBX(config)#voice translation-profile SIP-Inbound
R1-PBX(cfg-translation-profile)#translate called 11
```

Następnie profil translacji powinien zostać skojarzony z regułą wybierania:

```
R1-PBX(config)#dial-peer voice 89 voip
R1-PBX(config-dial-peer)#translation-profile incoming SIP-Inbound
```

Pamiętasz, co mówiłem na temat niedziałającego mechanizmu MWI? Translacja jest jedną z przyczyn niedogodności. Gdy moduł AIM-CUE wybiera numer 80000101 w celu dostarczenia komunikatu MWI do telefonu, wykorzystuje protokół SIP. Numer spełnia kryteria dopasowania reguły wybierania (.%) i zostaje zamieniony na wartość 133. To z kolei powoduje dzwonienie wszystkich telefonów.

Teraz rozmowa przychodząca jest poprawnie przekazywana do wszystkich telefonów, którym przypisano identyfikator ephone-dn o wartości 33. Oto przebieg tej samej rozmowy, ale z włączoną translacją połączeń przychodzących:

```
Feb 3 19:29:03.668: //-1/xxxxxxxxxxxx/SIP/Msg/ccsipDisplayMsg:
Sent:
SIP/2.0 100 Trying
Via: SIP/2.0/UDP 208.73.146.95:5060;branch=z9hG4bK8v1qso00903gmpgqi501.1
From: "DONAHUE GARY " <sip:+19735551212@192.168.101.112:5060>;tag=gK025eb403
To: <sip:9735554333@192.168.101.100:5060>
Date: Thu, 03 Feb 2011 19:29:03 GMT
Call-ID: pcst1296761343515383102821 0@192.168.201.112
CSeq: 1 INVITE
Allow-Events: telephone-event
Server: Cisco-SIPGateway/IOS-12.x
Content-Length: 0


Feb 3 19:29:03.672: //-1/xxxxxxxxxxxx/SIP/Msg/ccsipDisplayMsg:
Sent:
SIP/2.0 180 Ringing
Via: SIP/2.0/UDP 208.73.146.95:5060;branch=z9hG4bK8v1qso00903gmpgqi501.1
From: "DONAHUE GARY " <sip:+19735551212@192.168.101.112:5060>;tag=gK025eb403
To: <sip:9735554333@192.168.101.100:5060>;tag=3F32A10-2138
Date: Thu, 03 Feb 2011 19:29:03 GMT
Call-ID: pcst1296761343515383102821 0@192.168.201.112
CSeq: 1 INVITE
Allow: INVITE, OPTIONS, BYE, CANCEL, ACK, PRACK, UPDATE, REFER, SUBSCRIBE, NOTIFY,
INFO, REGISTER
Allow-Events: telephone-event
Remote-Party-ID: <sip:133@192.168.0.3>;party=called;screen=no;privacy=off
Contact: <sip:608222222@192.168.0.3:5060>
Server: Cisco-SIPGateway/IOS-12.x
Content-Length: 0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Treść listingu zmienia się zależnie od wersji oprogramowania. Gdy korzystałem z systemu CME 4, po zastosowaniu translacji ciąg `To: <sip:9735554333...` był przedstawiany jako `<sip:133....`

Skoro połączenia przychodzące są obsługiwane poprawnie, zajmijmy się regułami połączeń wychodzących. Z pewnością domyślasz się, że również połączenia wychodzące powinny zostać objęte mechanizmem translacji. Próba wybrania numeru zewnętrznego z aparatu o dowolnym numerze wewnętrznym kończy się niepowodzeniem, ponieważ komunikat SIP `INVITE` zawiera numer wewnętrzny, który nie jest poprawnie rozpoznawany przez system dostawcy. W konsekwencji operacja kończy się zwróceniem informacji o „nieautoryzowanym połączeniu". Konieczne jest więc utworzenie reguły translacji, która zapewni zastąpienie wszystkiego, co wygląda jak numer telefonu dzwoniącego, nazwą użytkownika jednego z łączy trunkowych SIP.

```
R1-PBX(config)#voice translation-rule 1
R1-PBX(cfg-translation-rule)#rule 1 /^.*/ /608222222/
```

Kolejny problem wynika z założenia, że każdy numer zewnętrzny powinien zostać poprzedzony cyfrą 9. Wymóg ten dotyczy wszystkich połączeń, również tych, które są przekazywane w łączu trunkowym SIP. Konieczne jest więc usunięcie początkowej cyfry 9 z numerów połączeń wychodzących:

```
R1-PBX(config)#voice translation-rule 2
R1-PBX(cfg-translation-rule)#rule 1 /^9/ //
```

Teraz trzeba utworzyć nowy profil translacji, który uwzględni obydwie reguły:

```
R1-PBX(config)#voice translation-profile SIP-Outbound
R1-PBX(cfg-translation-profile)#translate calling 1
R1-PBX(cfg-translation-profile)#translate called 2
```

Translacja została zdefiniowana. Czas utworzyć regułę wybierania. Można ją było zdefiniować przed opracowaniem zasad translacji, ale i tak nie działałaby poprawnie. Polecenia są analogiczne jak w przypadku połączeń przychodzących z tą różnicą, że ustawienie musi zostać zapisane w instrukcji `destination-pattern`. Kryterium dopasowania numeru docelowego jest spełnione, jeśli numer składa się z 11 cyfr i rozpoczyna się od cyfry 9.

```
R1-PBX(config)#dial-peer voice 90 voip
R1-PBX(config-dial-peer)#translation-profile outgoing SIP-Outbound
R1-PBX(config-dial-peer)#destination-pattern 9..........
R1-PBX(config-dial-peer)#session protocol sipv2
R1-PBX(config-dial-peer)#session target sip-server
R1-PBX(config-dial-peer)#dtmf-relay rtp-nte
R1-PBX(config-dial-peer)#codec g711ulaw
R1-PBX(config-dial-peer)#no vad
```

Dotychczasowa konfiguracja sprawdza się podczas ustanawiania połączeń z rodziną i znajomymi, ale trzeba także uwzględnić przypadki wybierania numerów alarmowych. Oczywiście istnieje kilka sposobów ich ustawienia w zależności od tego, czy dostawca ma uruchomioną usługę E-911, czy nie. Jeśli ma (tak jak w opisywanym przypadku), wystarczy utworzyć plan numeracji, który uwzględni we wzorcu dopasowania numerów docelowych wartość `911`. Pierwsza reguła wybierania odnosi się do numeru 911 niepoprzedzonego cyfrą 9.

```
R1-PBX(config)#dial-peer voice 91 voip
R1-PBX(config-dial-peer)#translation-profile outgoing SIP-Outbound
R1-PBX(config-dial-peer)#destination-pattern 911
R1-PBX(config-dial-peer)#session protocol sipv2
R1-PBX(config-dial-peer)#session target sip-server
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
R1-PBX(config-dial-peer)#dtmf-relay rtp-nte
R1-PBX(config-dial-peer)#codec g711ulaw
R1-PBX(config-dial-peer)#no vad
```

Potrzebne będzie jeszcze jedno kryterium dopasowania odpowiadające przypadkom, w których użytkownik **będzie** pamiętał o wprowadzeniu cyfry 9 przed zasadniczym numerem. Cała konfiguracja przebiega niemal identycznie. Inny jest tylko identyfikator reguły wybierania oraz wzorzec dopasowania numeru docelowego. Aby zaoszczędzić miejsca, przedstawiam jedynie instrukcję wyznaczającą wzorzec dopasowania:

```
R1-PBX(config-dial-peer)#destination-pattern 9911
```

Kolejny wzorzec umożliwia wybieranie numerów 411 i 911 (a także 211, 311, 511, 611, 711 i 811):

```
R1-PBX(config-dial-peer)#destination-pattern 9[2-9]11
```

Ten z kolei zapewnia realizację połączeń międzynarodowych zainicjowanych w Stanach Zjednoczonych:

```
R1-PBX(config-dial-peer)#destination-pattern 9011T
```

Na tym etapie system jest gotowy do przyjmowania i wykonywania połączeń za pośrednictwem łącza trunkowego SIP. Jeśli użytkownik zdalny wybierze przydzielony firmie numer, zadzwonią wszystkie telefony przypisane do numeru wewnętrznego 133. Jeśli którykolwiek z pracowników firmy wprowadzi poprawny numer zewnętrzny, zostanie ustanowione połączenie wychodzące realizowane za pośrednictwem łącza trunkowego SIP. Dzwonienie między numerami wewnętrznymi nie jest w żaden sposób ograniczone, podobnie jak korzystanie z poczty głosowej.

Rozwiązywanie problemów

Rozwiązywanie problemów w działaniu systemów VoIP i SIP oraz oprogramowania CME nie jest szczególnie skomplikowane, ale jak w przypadku każdej innej funkcji urządzeń Cisco trzeba znać polecenia, które wygenerują użyteczne zestawienia.

Rejestracja telefonów

Poprawne zarejestrowanie telefonu powoduje wygenerowanie komunikatu konsolowego, który można przechwycić w połączeniu telnet lub SSH po wpisaniu instrukcji `terminal monitor`. Poniższy listing został wygenerowany po włączeniu telefonu programowego:

```
Feb 3 01:44:52.667: %IPPHONE-6-UNREGISTER_NORMAL: ephone-10:SEP001C42D1565D IP:192.168.10.1 Socket:7
↪DeviceType:Phone has unregistered normally.
Feb 3 01:45:32.264: %IPPHONE-6-REG_ALARM: 25: Name=SEP001C42D1565D Load= 7.0.5.0 Last=Initialized
Feb 3 01:45:32.268: %IPPHONE-6-REGISTER: ephone-10:SEP001C42D1565D IP:10.211.55.11 Socket:8
↪DeviceType:Phone has registered.
```

Kolejny wpis pojawił się po przerwaniu działania programu bez poprawnego zakończenia jego pracy (jak w przypadku odłączenia telefonu od gniazdka):

```
Feb 3 02:08:03.692: %IPPHONE-6-UNREGISTER_ABNORMAL: ephone-10:SEP001C42D1565D IP:192.168.10.1 Socket:7
↪DeviceType:Phone has unregistered abnormally.
```

Aby sprawdzić, czy telefon został poprawnie zarejestrowany, można wykonać polecenie `show ephone`. Z poniższego zestawienia wynika, że dwa telefony ethernetowe nie zostały zarejestrowane, a trzeci poprawnie wykonał procedurę rejestracji. Dodatkowo wyświetlane są informacje o tym, które przyciski zostały skonfigurowane oraz jakie numery są do nich przypisane:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
R1-PBX#sho ephone
ephone-1[0] Mac:0019.AA96.D096 TCP socket:[-1] activeLine:0 whisperLine:0 UNREGISTERED
mediaActive:0 whisper_mediaActive:0 startMedia:0 offhook:0 ringing:0 reset:0
reset_sent:0 paging 0 debug:0 caps:0
IP:0.0.0.0* 0 Unknown 0 keepalive 0 max_line 0 available_line 0
paging-dn 20
Preferred Codec: g711ulaw
Lpcor Type: none

ephone-2[1] Mac:001D.4595.7FE5 TCP socket:[-1] activeLine:0 whisperLine:0 UNREGISTERED
mediaActive:0 whisper_mediaActive:0 startMedia:0 offhook:0 ringing:0 reset:0
reset_sent:0 paging 0 debug:0 caps:0
IP:0.0.0.0* 0 Unknown 0 keepalive 0 max_line 0 available_line 0
paging-dn 20
Preferred Codec: g711ulaw
Lpcor Type: none

ephone-3[2] Mac:0012.4362.C4D2 TCP socket:[2] activeLine:0 whisperLine:0 REGISTERED
in SCCP ver 11/9 max_streams=0
mediaActive:0 whisper_mediaActive:0 startMedia:0 offhook:0 ringing:0 reset:0
reset_sent:0 paging 0 debug:0 caps:8
IP:192.168.2.131 * 49916 Telecaster 7960 keepalive 87 max_line 6 available_line 6
button 1: cw:1 ccw:(0 0)
  dn 3  number 103 CH1   IDLE       CH2    IDLE        shared with monitor-ring
button 2: cw:1 ccw:(0 0)
  dn 1  number 101 CH1   DOWN       CH2    DOWN        monitor-ring shared
button 3: cw:1 ccw:(0 0)
  dn 2  number 102 CH1   IDLE       CH2    IDLE        monitor-ring shared
button 4: cw:1 ccw:(0 0)
  dn 4  number 104 CH1   IDLE       CH2    IDLE        monitor-ring shared
button 6: cw:1 ccw:(0 0)
  dn 10 number 111 CH1   IDLE       CH2    IDLE        shared
paging-dn 20
Preferred Codec: g711ulaw
Lpcor Type: none
```

TFTP

Podglądanie sesji TFTP pozwala na zorientowanie się, jakie działania są podejmowane przez telefony podczas rejestracji. Włączenie podglądu zdarzeń związanych z protokołem TFTP sprowadza się do wykonania instrukcji `debug tftp events`:

```
R1-PBX#debug tftp events
TFTP Event debugging is on   //Rejestracja zdarzeń TFTP jest włączona
```

Ponowne uruchomienie telefonu programowego powoduje wygenerowanie następujących komunikatów:

```
R1-PBX(config-ephone)#restart
restarting 001C.42D1.565D
Feb 3 01:34:30.291: %IPPHONE-6-UNREGISTER_NORMAL: ephone-10:SEP001C42D1565D IP:192.168.10.1 Socket:6
DeviceType:Phone has unregistered normally.
Feb 3 01:34:30.335: TFTP: Looking for CTLSEP001C42D1565D.tlv
Feb 3 01:34:30.451: TFTP: Looking for CTLSEP001C42D1565D.tlv
Feb 3 01:34:30.615: TFTP: Looking for CTLSEP001C42D1565D.tlv
Feb 3 01:34:30.795: TFTP: Looking for SEP001C42D1565D.cnf.xml
Feb 3 01:34:30.951: TFTP: Looking for CTLSEP001C42D1565D.tlv
Feb 3 01:34:31.143: TFTP: Looking for XMLDefault.cnf.xml
Feb 3 01:34:31.643: TFTP: Opened system:/its/vrf1/XMLDefault.cnf.xml, fd 10,
size 3088 for process 210
Feb 3 01:34:32.451: TFTP: Finished system:/its/vrf1/XMLDefault.cnf.xml, time 00:00:00 for process 210
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Narzędzie to jest bardzo użyteczne w przypadku dodawania dzwonków lub obrazów pulpitu. Pierwsze wybranie opcji zmiany tła pulpitu w telefonie 7970 powoduje pobranie z odpowiedniego katalogu serwera pliku *List.xml*. Dowodem na to jest poniższy zapis.

```
Feb 3 01:37:56.976: TFTP: Looking for Desktops/320x212x12/List.xml
Feb 3 01:37:56.980: TFTP: Opened flash:Desktops/320x212x12/List.xml, fd 10, size 1116 for process 210
Feb 3 01:37:57.400: TFTP: Finished flash:Desktops/320x212x12/List.xml, time 00:00:00 for process 210
```

Reguły wybierania

Weryfikacja sposobu działania reguł wybierania w starszych wersjach systemu IOS oznaczała konieczność wprowadzenia polecenia `debug dialpeer`. W nowszych wersjach systemu polecenie to zostało zastąpione instrukcją `debug voip dialpeer all`:

```
R1-PBX#debug dialpeer
This CLI command is now 'debug voip dialpeer all'
                                    //To polecenie CLI zostało zastąpione przez 'debug voip dialpeer all'
R1-PBX#debug voip dialpeer all
voip dialpeer all debugging is on   //Debugowanie wybierania w połączeniach VoIP jest włączone
```

Oczywiście można rejestrować tylko wybrane zdarzenia z grupy `dialpeer`:

```
R1-PBX#debug voip dialpeer ?
 all       Enable all debugs        //Włączenie pełnego debugowania
 default   Enable default debugs    //Włączenie domyślnego mechanizmu debugowania
 detail    detail debug             //Szczegółowe debugowanie
 error     major call and software errors debug
                                    //Debugowanie ważniejszych błędów połączenia i oprogramowania
 function  function debug           //Debugowanie funkcji
 inout     inout debug              //Debugowanie funkcji wejścia-wyjścia
 <cr>
```

Ta forma podglądu zdarzeń powoduje wygenerowanie znacznej ilości informacji:

```
Feb 3 21:08:20.935: //-1/xxxxxxxxxxxx/DPM/dpMatchPeersCore: Calling Number=608222222, Called
↪Number=608222222, Peer Info Type=DIALPEER_INFO_SPEECH
Feb 3 21:08:20.939: //-1/xxxxxxxxxxxx/DPM/dpMatchPeersCore: Match Rule=DP_MATCH_DEST; Called
↪Number=608222222
Feb 3 21:08:20.939: //-1/xxxxxxxxxxxx/DPM/dpMatchPeersCore: No Outgoing Dial-peer Is Matched;
↪Result=NO_MATCH(-1)
Feb 3 21:08:20.939: //-1/xxxxxxxxxxxx/DPM/dpMatchSafModulePlugin: dialstring=608222222, saf_enabled=1,
↪saf_dndb_lookup=1, dp_result=-1
Feb 3 21:08:20.939: //-1/xxxxxxxxxxxx/DPM/dpMatchPeersMoreArg: Result=NO_MATCH(-1)
Feb 3 21:08:20.939: //-1/xxxxxxxxxxxx/DPM/dpAssociateIncomingPeerCore: Calling Number=+19735551212,
↪Called Number=, Voice-Interface=0x0, Timeout=TRUE, Peer Encap Type=ENCAP_VOIP, Peer Search
↪Type=PEER_TYPE_VOICE, Peer Info Type=DIALPEER_INFO_SPEECH
```

Użycie polecenia `show dialpan number` *ciąg-cyfr* pozwala na sprawdzenie, czy dana reguła wybierania działa i w jaki sposób dopasowuje numery. Wykonanie polecenia nie powoduje wybrania jakiegokolwiek numeru. Dostarcza jedynie informację o tym, czy podany numer spełnił kryteria dopasowania reguły. Wynik jest dość rozbudowany, ale znalezienie odpowiedzi nie powinno stanowić problemu. Oto przykład użycia instrukcji podczas konfiguracji reguł linii POTS:

```
R1-PBX#sho dialplan number 92111111111
Macro Exp.: 92111111111

VoiceEncapPeer100
        peer type = voice, system default peer = FALSE, information type = voice,
        description = `',
        tag = 100, destination-pattern = `9[2-9].........',
        voice reg type = 0, corresponding tag = 0,
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
        allow watch = FALSE
        answer-address = `', preference=0,
        CLID Restriction = None
        CLID Network Number = `'
[fragment usinięty]
```

A oto wynik wykonania tego samego polecenia przy włączonych regułach połączeń VoIP:

```
R1-PBX#sho dialplan number 92111111111
Macro Exp.: 92111111111

VoiceOverIpPeer90
        peer type = voice, system default peer = FALSE, information type = voice,
        description = `',
        tag = 90, destination-pattern = `9.........',
        voice reg type = 0, corresponding tag = 0,
        allow watch = FALSE
        answer-address = `', preference=0,
        CLID Restriction = None
        CLID Network Number = `'
```

SIP

Aby sprawdzić, czy punkty końcowe protokołu SIP zostały zarejestrowane, należy wykonać polecenie `show sip-ua register status`. W prezentowanym przykładzie protokół SIP jest stosowany tylko w łączach trunkowych, a w definicjach `ephone-dn` zostało użyte słowo kluczowe `no-reg`. Poniższe zestawienie jest potwierdzeniem tego stanu. Gdyby w systemie pracowały telefony SIP, ich numery również byłyby widoczne.

```
R1-PBX#sho sip-ua reg status
Line         peer         expires(sec) registered  P-Associated-URI
============ ============ ============ =========== ================
557333333    -1           2175         yes
608222222    -1           2172         yes
```

Aby przeanalizować wszystkie komunikaty wysyłane i odbierane w protokole SIP, trzeba wykonać polecenie `debug ccsip messages`. Jest to doskonałe narzędzie, które prezentuje każdy bit komunikacji SIP od rejestracji telefonów i łączy trunkowych po wymianę komunikatów `INVITE`, `CANCEL`, `BYE` oraz wszystkich odpowiedzi na nie. W przypadku jakichkolwiek problemów z protokołem SIP zawsze dostarcza szczegółowych informacji na temat realizowanych zadań.

```
R1-PBX#debug ccsip messages
SIP Call messages tracing is enabled  //Śledzenie komunikatów w ramach połączeń SIP jest włączone
```

Najlepszym dowodem na to, że zestawienie wynikowe może być bardzo szczegółowe, jest poniższy listing. Został on wygenerowany w chwili, gdy wybrałem numer telefonu i odłożyłem słuchawkę przed odebraniem rozmowy przez drugą osobę.

```
*Feb 3 01:17:07.740: //-1/xxxxxxxxxxxx/SIP/Msg/ccsipDisplayMsg:
Sent:
INVITE sip:9735554471@sbc.sip-provider.com:5060 SIP/2.0
Via: SIP/2.0/UDP 192.168.0.3:5060;branch=z9hG4bKB1761
From: <sip:9735551212@sbc.sip-provider.com>;tag=383B4-0
To: <sip:9735554471@sbc.sip-provider.com>
Date: Thu, 03 Feb 2011 01:17:07 GMT
Call-ID: 26F3EF99-2E6A11E0-800D80A2-733EA81A@192.168.0.3
Supported: 100rel,timer,resource-priority,replaces,sdp-anat
Min-SE: 1800
Cisco-Guid: 0563478376-0778703328-2148040866-1933486106
User-Agent: Cisco-SIPGateway/IOS-12.x
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Allow: INVITE, OPTIONS, BYE, CANCEL, ACK, PRACK, UPDATE, REFER, SUBSCRIBE, NOTIFY,
INFO, REGISTER
CSeq: 102 INVITE
Max-Forwards: 70
Timestamp: 1296695827
Contact: <sip:9735551212@192.168.0.3:5060>
Expires: 180
Allow-Events: telephone-event
Proxy-Authorization: Digest username="9735551212",realm="voip.sip-provider.com",
uri="sip:9735554471@sbc.sip-provider.com:5060",
response="48fab7d8750961aa91e43f11879ddd4f",nonce="9967d6506474325b254afb6d6bc1baf4",
algorithm=md5
Content-Type: application/sdp
Content-Disposition: session;handling=required
Content-Length: 264

v=0
o=CiscoSystemsSIP-GW-UserAgent 8383 1800 IN IP4 192.168.0.3
s=SIP Call
c=IN IP4 192.168.0.3
t=0 0
m=audio 18674 RTP/AVP 0 101
c=IN IP4 192.168.0.3
a=rtpmap:0 PCMU/8000
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-16
a=ptime:20
a=direction:active

*Feb 3 01:17:07.844: //-1/xxxxxxxxxxxx/SIP/Msg/ccsipDisplayMsg:
Received:
SIP/2.0 100 Trying
Via: SIP/2.0/UDP 192.168.0.3:5060;branch=z9hG4bKB1761
From: <sip:9735551212@sbc.sip-provider.com>;tag=383B4-0
To: <sip:9735554471@sbc.sip-provider.com>
Call-ID: 26F3EF99-2E6A11E0-800D80A2-733EA81A@192.168.0.3
CSeq: 102 INVITE
Timestamp: 1296695827

*Feb 3 01:17:11.004: //-1/xxxxxxxxxxxx/SIP/Msg/ccsipDisplayMsg:
Sent:
CANCEL sip:9735551212@sbc.sip-provider.com:5060 SIP/2.0
Via: SIP/2.0/UDP 192.168.0.3:5060;branch=z9hG4bKB1761
From: <sip:9735551212@sbc.sip-provider.com>;tag=383B4-0
To: <sip:9735551212@sbc.sip-provider.com>
Date: Thu, 03 Feb 2011 01:17:07 GMT
Call-ID: 26F3EF99-2E6A11E0-800D80A2-733EA81A@192.168.0.3
CSeq: 102 CANCEL
Max-Forwards: 70
Timestamp: 1296695831
Reason: Q.850;cause=16
Content-Length: 0
*Feb 3 01:17:11.112: //-1/xxxxxxxxxxxx/SIP/Msg/ccsipDisplayMsg:
Received:
SIP/2.0 200 OK
Via: SIP/2.0/UDP 192.168.0.3:5060;branch=z9hG4bKB1761
From: <sip:9735551212@sbc.sip-provider.com>;tag=383B4-0
To: <sip:9735551212@sbc.sip-provider.com>;tag=aprq1smnhp0-l46jaq00000c6
Call-ID: 26F3EF99-2E6A11E0-800D80A2-733EA81A@192.168.0.3
CSeq: 102 CANCEL
Timestamp: 1296695831
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
*Feb 3 01:17:11.136: //-1/xxxxxxxxxxxx/SIP/Msg/ccsipDisplayMsg:
Received:
SIP/2.0 487 Request Terminated
Via: SIP/2.0/UDP 192.168.0.3:5060;branch=z9hG4bKB1761
From: <sip:9735551212@sbc.sip-provider.com>;tag=383B4-0
To: <sip:9735551212@sbc.sip-provider.com>;tag=131481110
Call-ID: 26F3EF99-2E6A11E0-800D80A2-733EA81A@192.168.0.3
CSeq: 102 INVITE
Timestamp: 1296695827
Contact: <sip:9735551212@208.73.146.95:5060;transport=udp>
Content-Length: 0
User-Agent: Cisco-SIPGateway/IOS-12.x

*Feb 3 01:17:11.144: //-1/xxxxxxxxxxxx/SIP/Msg/ccsipDisplayMsg:
Sent:
ACK sip:9735551212@sbc.sip-provider.com:5060 SIP/2.0
Via: SIP/2.0/UDP 192.168.0.3:5060;branch=z9hG4bKB1761
From: <sip:9735551212@sbc.sip-provider.com>;tag=383B4-0
To: <sip:9735551212@sbc.sip-provider.com>;tag=131481110
Date: Thu, 03 Feb 2011 01:17:07 GMT
Call-ID: 26F3EF99-2E6A11E0-800D80A2-733EA81A@192.168.0.3
Max-Forwards: 70
CSeq: 102 ACK
Allow-Events: telephone-event
Content-Length: 0
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 31.

Podstawy QoS

Mechanizmy gwarantowania jakości usług (QoS — *Quality of Service*) są implementowane po to, aby przepływ danych nie spowodował takiej zajętości łącza, która uniemożliwiałaby innym strumieniom korzystanie z tego samego łącza. Pamiętaj, że łącza sieci WAN są połączeniami szeregowymi. Bity wprowadzane po jednej stronie są więc odbierane na drugim końcu połączenia w takiej samej kolejności, w jakiej zostały wysłane. Niezależnie od tego, czy łącze zostało wykonane zgodnie ze standardem T1 i ma przepustowość 1,5 Mb/s, czy jest zgodne ze standardem OC3 i ma przepustowość 155 Mb/s, w danym czasie przesyłany jest tylko jeden bit.

Mechanizm QoS umożliwia nadanie określonym rodzajom ruchu wyższego priorytetu niż innym. Po sklasyfikowaniu ruch o najwyższym priorytecie jest przekazywany w pierwszej kolejności, natomiast strumienie o niższym priorytecie podlegają kolejkowaniu. Najważniejszym zadaniem systemu QoS jest określenie, które strumienie powinny otrzymać wyższy priorytet w dostępie do łącza.

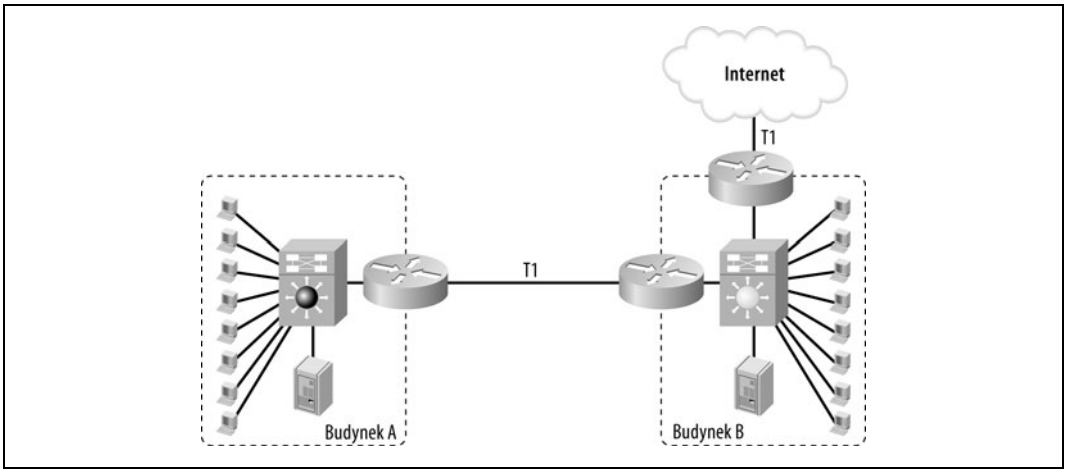
Na rysunku 31.1 przedstawiono dwa budynki połączone za pomocą linii T1. Budynek B jest również przyłączony do internetu za pomocą łącza T1. W każdym budynku pracują serwery oraz stacje robocze stu użytkowników. Dane serwerów są replikowane między lokalizacjami codziennie w godzinach pracy. Użytkownicy korzystają z telefonów IP, a komunikacja między budynkami jest relatywnie częsta. Pracownicy mogą korzystać z internetu.

Jedyna trasa poza sieć budynku A wiedzie do budynku B (łączem T1). Co się stanie, gdy wszyscy użytkownicy pracujący w budynku A zechcą skorzystać z pojedynczego łącza w tym samym czasie? Przepustowość łącza wynosi 1,5 Mb/s, a każdy użytkownik dysponuje połączeniem z siecią o wydajności 100 Mb/s (lub nawet 1 Gb/s).

Dobry projektant sieci nie dopuściłby do zbudowania systemu o takich ograniczeniach, a administrator (jeśli takowy jest zatrudniony) zadbałby o to, żeby problem był mniej dotkliwy. Niemniej wciąż wiele sieci ma podobną architekturę i czeka, aż ktoś ją zmieni. Małe firmy nie zatrudniają administratorów sieci, więc opisywane tutaj problemy mogą się zdarzać w praktyce.

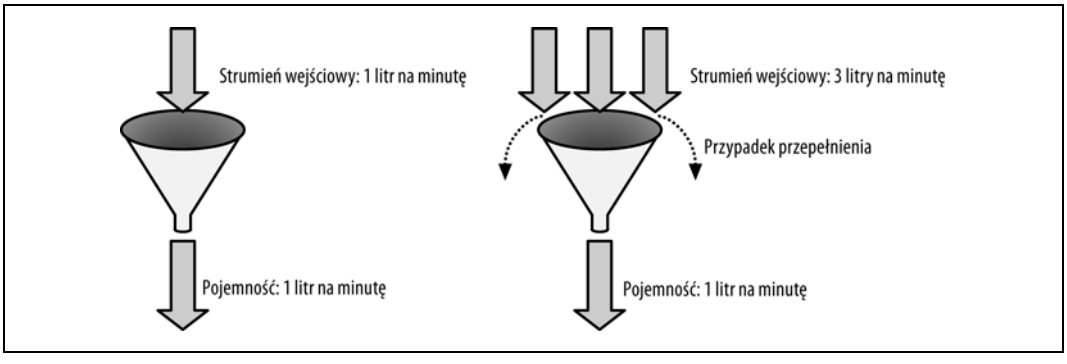
Załóżmy, że 30 osób spośród 100 użytkowników zaczyna w tym samym czasie pobierać darmowy utwór promujący płytę z najlepszymi solówkami gitarowymi. Przyjmijmy, że 20 osób przeszukuje zasoby księgarni internetowej. Kolejne 20 osób pobiera najnowsze aktualizacje do systemów operacyjnych, a pozostałe 30 osób korzysta z telefonii IP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 31.1. Sieć w dwóch budynkach

Problemem jest to, że łącze T1 nie gwarantuje dostatecznie dużej przepustowości, aby przenieść cały opisany ruch. Wyobraź sobie, że jest to wąż z lejkiem na górze (zobacz rysunek 31.2). Wąż zapewnia przepływ jednego litra cieczy w ciągu minuty. Dopóki ilość przelewanej cieczy nie przekracza jednego litra na minutę, nie występują żadne problemy w działaniu układu.



Rysunek 31.2. Przypadek przepełnienia

Załóżmy, że postanowiłeś zaszaleć i zamierzasz wlać trzy litry cieczy w czasie jednej minuty. Określona średnica węża sprawia, że na jego końcu ciecz wypływa w tempie jednego litra na minutę. Co więc się stanie z pozostałymi dwoma litrami? Ciecz najpierw wypełni lejek, a następnie przeleje się przez jego górną krawędź i rozleje się po podłodze. Wystąpi **przepełnienie**.

Prawa rządzące łączem T1 są podobne, ale zamiast cieczy przekazywane są pakiety (które w przeciwieństwie do cieczy mają właściwości nieciągłych obiektów logicznych). Połączenie T1 umożliwia przekazywanie danych z szybkością 1 500 000 bitów na sekundę. Gdy częstotliwość dostarczania bitów wzrośnie powyżej tej wartości, interfejs routera będzie buforować kolejne pakiety (podobnie jak lejek z cieczą) aż do zapełnienia bufora. Później dochodzi do przepełnienia bufora i interfejs zaczyna odrzucać pakiety. Dane są wówczas tracone bezpowrotnie.

Pewne rodzaje pakietów są ważniejsze od innych. Można je kategoryzować na różne sposoby, na przykład na podstawie protokołów: UDP, TCP, FTP, HTTP i VoIP. Poszczególne pakiety

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

mają różne cechy i obowiązują między nimi różne zależności czasowe. Niektóre protokoły wymagają dostarczania pakietów w odpowiedniej kolejności. Natomiast inne są szczególnie wrażliwe na utratę pakietów. Przeanalizujmy niektóre z protokołów i zobaczmy, czym się różnią.

TCP
: Protokół TCP uwzględnia mechanizm alarmowania stacji nadawczej o utracie lub uszkodzeniu pakietów i obowiązku ponowienia transmisji. Dzięki temu aplikacje bazujące na protokole TCP nie są wrażliwe na utratę pakietów. Ponadto ich działanie w mniejszym stopniu zależy od czasu transmisji niż w przypadku aplikacji UDP.

UDP
: Protokół UDP nie obejmuje żadnych mechanizmów sprawdzania błędów ani informowania o utracie pakietów. Z tego względu aplikacje bazujące na protokole UDP muszą uwzględniać ryzyko utraty danych podczas transmisji.

HTTP
: HTTP jest protokołem bazującym na mechanizmie TCP. Większość aplikacji HTTP nie jest zależna od czasu. Dłuższe oczekiwanie na załadowanie rysunku podczas przeglądania strony (często w czasie największego ruchu) nie jest zazwyczaj postrzegane jako problem.

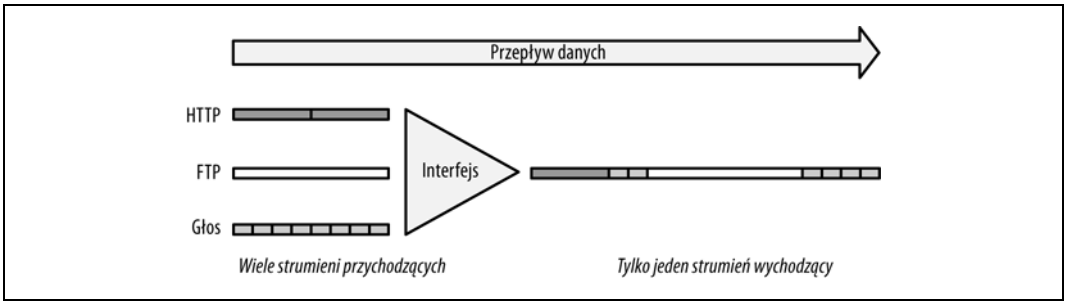
FTP
: FTP jest protokołem bazującym na mechanizmie TCP. Nie jest to rozwiązanie czasu rzeczywistego. Nie jest również wrażliwe na opóźnienia. Ewentualna utrata pakietów podczas transmisji wiąże się z wydłużeniem czasu przesyłania pliku (czas ten jest potrzebny na retransmisję danych), ale nie stanowi większego problemu dla użytkownika.

Telnet i SSH
: Protokoły telnet i SSH bazują na protokole TCP. Choć mogą się wydawać rozwiązaniami czasu rzeczywistego, wcale nimi nie są. Ewentualna retransmisja pakietów objawia się dłuższym czasem odpowiedzi na naciskanie klawiszy. Nie wpływa natomiast na wynik operacji.

VoIP
: Rozwiązania VoIP wykorzystują protokół UDP jako podstawę działania protokołu RTP (służącego do przenoszenia głosu) oraz protokół TCP w procesie sterowania rozmową. Technologia VoIP wymaga wysokiej niezawodności oraz dużych przepustowości. Nie toleruje również dostarczania pakietów w niewłaściwej kolejności. Zastosowanie mechanizmu UDP może się więc wydać nielogiczne, ponieważ nie gwarantuje on niezawodnego dostarczania danych. Znajduje jednak zastosowanie w rozwiązaniach VoIP, ponieważ wnosi mniejszy narzut obliczeniowy niż protokół TCP i charakteryzuje się mniejszym zapotrzebowaniem na pasmo. Zysk na szybkości jest bardzo istotny. Problemy niezawodności można natomiast rozwiązać za pomocą mechanizmów QoS. W praktyce wdrażanie technologii VoIP jest jednym z głównych powodów uruchamiania mechanizmów QoS.

Przy założeniu, że interfejs działa zgodnie z zasadą FIFO (pierwszy element wchodzący jest pierwszym elementem wychodzącym), pakiety mogą być emitowane przez interfejs w sposób nieciągły. Na rysunku 31.3 przedstawiono trzy rodzaje przekazywanych informacji: HTTP, FTP i głos. Dane głosowe są szczególnie wrażliwe na dostarczanie niezgodnie z kolejnością nadawania. Problemem jest również utrata pakietów. Pamiętaj, że połączenia głosowe są realizowane z wykorzystaniem protokołu UDP. Brak niezawodnego mechanizmu transportowego powoduje, że ewentualne przepełnienie bufora i utrata pakietu ma charakter ostateczny. Skutkiem tego jest pogorszenie jakości rozmów, irytacja użytkowników i niezadowolenie menedżerów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 31.3. Pakiety przesyłane przez interfejs

Mechanizm QoS może złagodzić negatywne skutki opisanych problemów (a nawet je wyeliminować). Konfiguracja systemu QoS z pewnością nie należy do najłatwiejszych z uwagi na mnogość opcji. Jeśli zdecydujesz się na zdawanie egzaminu CCIE, będziesz musiał poznać je wszystkie, a dodatkowo opracować techniki ich nadużywania. Większość testowych problemów nie ma pokrycia w rzeczywistości. Dlatego rozwiązania opisywane w dalszej części rozdziału będą dotyczyły bardziej typowych przypadków.

W rzeczywistym świecie systemy QoS są stosowane z wielu względów. Niekiedy z powodu wykorzystywania mediów strumieniowych (takich jak głos i wideo), które wymagają małego opóźnienia i szybkiego dostarczania danych. Czasami przyczyną jest przeciążenie łącza i chęć usprawnienia transmisji bez konieczności wykupywania droższych usług. Zdarzają się również sytuacje, w których nowe łącze zostało wykupione, ale realizacja zamówienia zajmuje na przykład 120 dni. Mechanizm QoS może wówczas pomóc w efektywniejszym wykorzystaniu istniejącego połączenia, zanim nowa usługa stanie się dostępna.

Rozwiązania QoS znajdują zastosowanie także w nieco bardziej wyrafinowanych konfiguracjach. Możesz na przykład nadać połączeniom telnet i SSH wyższy priorytet niż innym formom ruchu sieciowego. Jeśli Twoją sieć zaatakuje wirus, zawsze będziesz mógł się połączyć ze swoimi routerami (za pomocą protokołu telnet lub SSH). Analogicznie, gdy chcesz się przypodobać szefowi, możesz podnieść priorytet generowanego przez niego ruchu (oczywiście nie kosztem własnego), dzięki czemu wzrośnie komfort jego pracy. Oczywiście wymienione przykłady nie są rzeczywistymi problemami, ale wielokrotnie byłem proszony o przygotowanie podobnych mechanizmów w sieciach moich klientów. Często dział wykonawczy musi mieć zagwarantowane wydajniejsze połączenia sieciowe niż pozostałe działy firmy. Zdarza się też, że po usłyszeniu o mechanizmach QoS menedżerowie żądają „lepszego traktowania" ich danych w sieci. Nie zalecam takiego rozwiązania, ale jest ono możliwe do zrealizowania.

Ogólnie rzecz ujmując, każdy ruch, który można zidentyfikować, można również przetwarzać z odpowiednim priorytetem. Ustalenie, czy jest to wykonalne, należy do Twoich obowiązków służbowych, a samo wdrażanie mechanizmu bywa niekiedy zabawne.

Rodzaje mechanizmów QoS

Określenie „mechanizm QoS" pasuje do każdej z licznego zbioru funkcji, które służą do ograniczania pasma, gwarantowania przepustowości lub wyznaczania priorytetu określonego rodzaju ruchu. Z technicznego punktu widzenia QoS jest pewnym zbiorem protokołów i idei. Terminy

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

CoS (klasyfikowanie usług), QoS, kolejkowanie, oznaczanie, nakładanie polityki lub nawet kształtowanie ruchu w zasadzie opisują to samo zadanie.

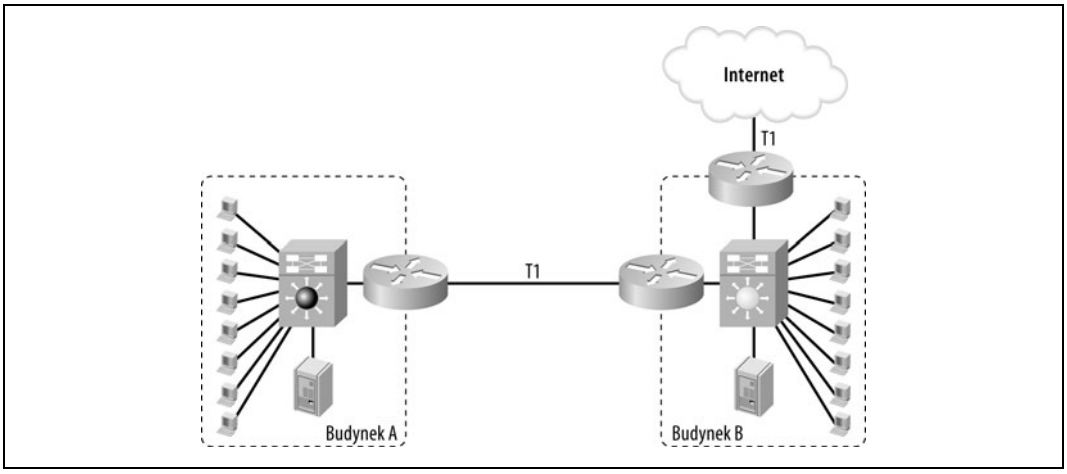
Na przykład w routerach Cisco każde łącze o przepustowości niższej niż 2 Mb/s ma domyślnie włączony mechanizm równomiernego kolejkowania ważonego (WFQ — *Weighted Fair Queuing*). Oto fragment dokumentacji Cisco na jego temat:

> WFQ jest jedną z najważniejszych technik kolejkowania firmy Cisco. Wykorzystuje algorytm analizy przepływu danych, który wykonuje jednocześnie dwie operacje — przenosi pakiety usług interaktywnych na początek kolejki i równomiernie dzieli pozostałe pasmo pomiędzy strumienie szerokopasmowe.

Według mnie to jest właśnie QoS! Co więcej, mechanizm ten jest prawdopodobnie wykorzystywany w sieci, w której pracujesz. Wyłączenie funkcji WFQ w mocno obciążonym łączu mogłoby się okazać tragiczne w skutkach.

Niektóre formy techniki QoS są bardzo łatwe do zaimplementowania, ale oddziałują jedynie na interfejsy, w których zostały uruchomione. Inne rodzaje mechanizmów można zastosować zarówno w odniesieniu do pojedynczego interfejsu, jak i do połączeń między punktami końcowymi w sieciach korporacyjnych. Jednym z najczęstszych powodów uruchamiania mechanizmów QoS na wielką skalę jest wdrażanie technologii VoIP.

Przeanalizujmy rysunek 31.4, który prezentuje dwubudynkową sieć z systemem VoIP. W obydwu budynkach działają komputery PC i telefony IP. W klasycznej konfiguracji wszystkie zasoby sieciowe są wykorzystywane równomiernie. Taki sposób działania nie jest jednak pożądany w przypadku rozwiązań VoIP, które wymagają szybkiego dostarczania danych z zachowaniem odpowiedniej kolejności i bez ryzyka ich utraty. Strumień głosu jest transportowany za pośrednictwem protokołu UDP, więc nie można zagwarantować niezawodności transmisji, którą zapewnia protokół TCP. Poza tym pakiety VoIP muszą rywalizować o pasmo z ruchem wynikającym ze stosowania także innych protokołów.



Rysunek 31.4. Sieć VoIP obejmująca dwa budynki

Załóżmy, że w danej chwili dwie osoby przebywające w budynkach A i B prowadzą rozmowę VoIP. W pewnej chwili inny pracownik budynku A zaczyna pobierać duży plik z serwera znajdującego w budynku B. Używa do tego celu protokołu FTP. W tym samym momencie

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

połączenie głosowe zaczyna szwankować. Znając moje szczęście, prawdopodobnie byłaby to rozmowa między prezesem i księgowością na temat podwyżki dla mnie. Z uwagi na niedostateczną jakość połączenia rozmowa zostaje przerwana.

Przeanalizujmy ten problem i spróbujmy znaleźć jakieś rozwiązanie. W realizacji zadania pomocne będą informacje na temat zasad działania mechanizmów QoS oraz rodzajów wdrażanych systemów QoS.

Działanie mechanizmów QoS

Działanie mechanizmu QoS polega na kolejnym oznaczaniu pakietów, nakładaniu polityki i harmonogramowaniu transmisji. **Oznaczanie** pakietów wiąże się z przypisywaniem pakietom odpowiedniego priorytetu. Na przykład strumień RTP powinien otrzymać najwyższy priorytet. Termin **nakładanie polityki** odnosi się do działań podejmowanych przez router na podstawie etykiet przypisanych pakietom na etapie oznaczania. Można by na przykład ustalić, że pakiety oznaczone jako należące do ruchu o najwyższym priorytecie powinny mieć zagwarantowane co najmniej 10 procent dostępnego pasma. **Harmonogramowanie** odpowiada za ustalenie kolejności obsługi pakietów, które zostały uprzednio oznaczone i poddane działaniu określonej polityki. W przypadku strumienia RTP (strumienia o najwyższym priorytecie) pakiety połączeń głosowych będą przekazywane w pierwszej kolejności, a o kolejności dostarczania pakietów innych form ruchu zdecydują ich priorytety.

Wymienione czynności mogą być realizowane przez jedno urządzenie lub przez kilka odrębnych urządzeń. Na przykład telefony IP firmy Cisco automatycznie przypisują pakietom RTP najwyższy priorytet. Wynika z tego, że sprawdzanie typu pakietów po stronie routera nie jest konieczne. Oznaczanie pakietów wnosi dodatkowe obciążenie, więc przekazanie zadania do urządzenia, które z założenia obsługuje określoną formę ruchu, wydaje się właściwym pomysłem.

Priorytety

Każdy pakiet IP zawiera pole odzwierciedlające rodzaj usługi (TOS — *Type Of Service*). Jego ośmiobitowa wartość jest podzielona na kilka sekcji, których znaczenie może być nieco mylące. Wyjątkowość rozwiązania tkwi w jego logice, którą postaram się wyjaśnić.

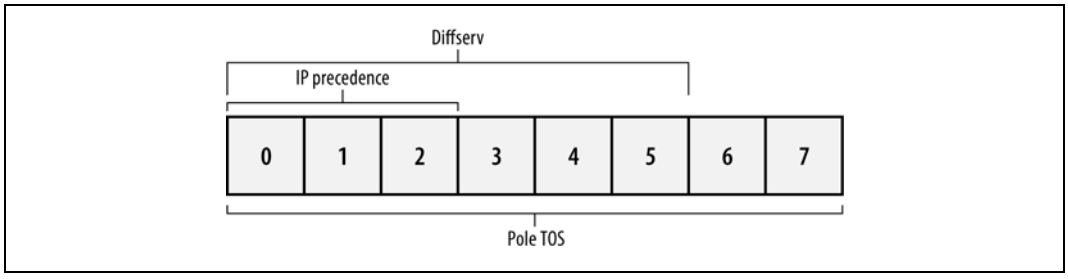
Na poziome warstwy trzeciej wykorzystywane są dwa modele wyznaczania priorytetów: **IP precedence** (pierwszeństwo IP) oraz **diffserv** (rozróżnianie usług).

Model IP precedence został opracowany w początkowej fazie rozwoju internetu. Technika diffserv jest znacznie nowsza. Wartości wyznaczane przez mechanizm diffserv są nazywane **kodami rozróżnianych usług** (DSCP — **Differentiated Services Code Point**).

Struktura pola TOS została przedstawiona na rysunku 31.5.

Pole składa się z ośmiu bitów, z których trzy pierwsze są wykorzystywane przez mechanizm IP precedence. Wartości IP precedence są nazywane **odwzorowaniami usług** (ang. *service mappings*), a ich definicje znajdują się w dokumencie RFC 795. Oto odpowiedni fragment tego dokumentu:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 31.5. Pole TOS protokołu IP

```
Network Working Group                                        J. Postel
Request for Comments:  795                                        ISI
                                                       September 1981
                            ODWZOROWANIA USŁUG
                            ------------------

Ta notatka opisuje zależności między rodzajem usługi protokołu internetowego (IP) [1] a parametrami
↪usług w sieciach.
Wartość TOS w protokole IP składa się z następujących pól:

   Bity 0-2:  Pierwszeństwo IP.
   Bit    3:  0 = Normalne opóźnienie,      1 = Małe opóźnienie.
   Bit    4:  0 = Normalna przepustowość,   1 = Wysoka przepustowość.
   Bit    5:  0 = Normalna niezawodność,    1 = Wysoka niezawodność.
   Bity 6-7:  Zarezerwowane do przyszłego wykorzystania.

      0     1     2     3     4     5     6     7
   +-----+-----+-----+-----+-----+-----+-----+-----+
   |                 |     |     |     |     |     |
   |   PRECEDENCE    |  D  |  T  |  R  |  0  |  0  |
   |                 |     |     |     |     |     |
   +-----+-----+-----+-----+-----+-----+-----+-----+

   111 - Zarządzanie siecią
   110 - Zarządzanie siecią internetową
   101 - CRITIC/ECP
   100 - Ruch błyskawiczny z zastępowaniem
   011 - Ruch błyskawiczny
   010 - Ruch natychmiastowy
   001 - Ruch priorytetowy
   000 - Ruch standardowy
```

Struktura internetu w początkowych fazach jego rozwoju (dokument RFC został opracowany 1981 roku) była znacznie mniej skomplikowana niż dzisiejsze sieci, więc trzy bity modelu IP precedence zostały uznane za wystarczające. Po pewnym czasie dostrzeżono jednak problem niedostatecznej skalowalności. Użytkownicy domagali się bowiem więcej niż ośmiu poziomów priorytetów. W reakcji na te postulaty opracowano dokument RFC 2474, który zmienił definicję pola TOS:

```
3.  Definicja pola rozróżnianych usług

   Opisywane pole nagłówka, nazywane polem DS, zastępuje istniejące
   definicje oktetu TOS w protokole IPv4 [RFC791] oraz oktetu
   klasy ruchu w protokole IPv6 [IPv6].

   Sześć bitów pola DS jest wykorzytywanych jako kody (DSCP) służące do wybierania
   określonego sposobu przetwarzania pakietu w każdym z węzłów. Nieużywane
   obecnie dwubitowe pole (CU) jest zarezerwowane, a jego definicja i
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

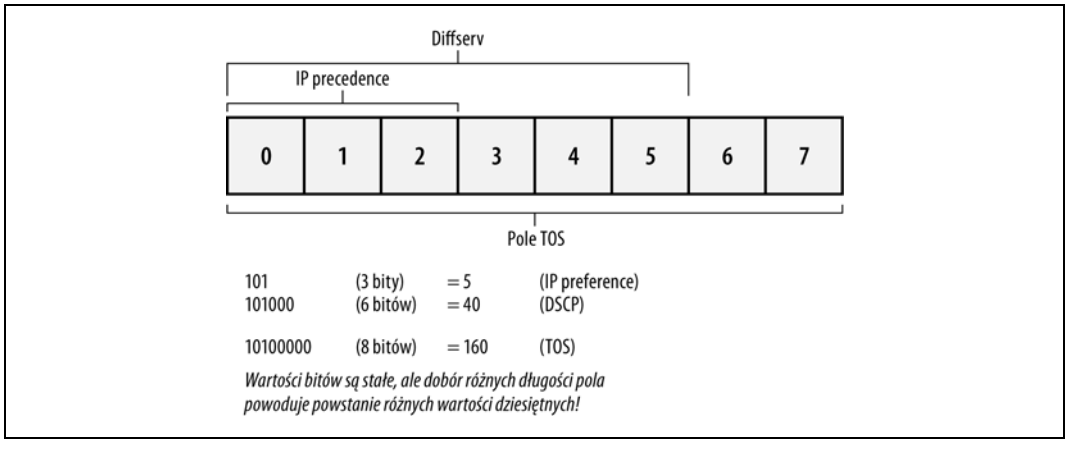
```
interpretacja wykraczają poza zakres tego dokumentu. Wartości bitów CU
są ignorowane przez węzły pracujące zgodnie z mechanizmem rozróżniania usług i
nie są brane pod uwagę podczas określania sposobu przetwarzania odebranego
pakietu.

Struktura pola DS została przedstawiona poniżej:

    0   1   2   3   4   5   6   7
  +---+---+---+---+---+---+---+---+
  |         DSCP          |  CU   |
  +---+---+---+---+---+---+---+---+

    DSCP: kod mechanizmu rozróżniania usług
    CU:   aktualnie nieużywane
```

Na czym polega wyjątkowość tego rozwiązania? Załóżmy, że pole TOS zawiera bity 10100000. Długość pola TOS zawsze odpowiada ośmiu bitom (oktetowi). Analizując je zgodnie z zasadami modelu IP precedence, musielibyśmy uwzględnić tylko trzy pierwsze bity, które reprezentują wartość 5. Jeśli spojrzymy na nie w sposób odpowiadający zasadom diffserv, sześć pierwszych bitów da nam wartość 40. Jeżeli jednak zamienimy do postaci dziesiętnej wszystkie osiem bitów, uzyskamy wartość 160. Zależności między wartościami IP precedence, diffserv i TOS zostały przedstawione na rysunku 31.6.



Rysunek 31.6. Różne wartości wynikające ze zmiany liczby bitów znaczących w polu TOS

Wiedza o tym, że wartość 160 w polu TOS odpowiada wartości 5 w mechanizmie IP precedence, może się okazać użyteczna, gdy będziesz analizował przechwycone pakiety. Na poziomie protokołu IP opisywany oktet jest po prostu polem TOS, zatem narzędzie przeznaczone do rejestracji pakietów przedstawi jego wartość jako wartość całego oktetu.

Mechanizm CoS jest formą QoS na poziomie warstwy drugiej. Podlega tym samym zasadom, które obowiązują w modelu IP precedence. Wartości CoS są zapisywane w trzybitowym polu, więc liczba ich kombinacji bitowych jest ograniczona do ośmiu. Różnica między rozwiązaniami polega na tym, że bity CoS są zapisywane w nagłówku ramki 802.1p, a nie w nagłówku pakietu IP. Wartości CoS obowiązują jedynie w Ethernecie, więc są usuwane, gdy ramka wychodzi poza sieć Ethernet. Jednak dzięki dokładnemu odwzorowaniu wartości IP precedence możliwe jest utworzenie reguł, które będą odczytywały zawartość pola CoS i przekształcały ją na wartości IP precedence.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W przełącznikach Catalyst firmy Cisco z każdym interfejsem ethernetowym związanych jest kilka kolejek. Wartości CoS można więc odwzorować na wspomniane kolejki i uruchomić w ten sposób system priorytetowania. Priorytety ethernetowe różnią się od priorytetów ustawianych w łączach sieci WAN, ale w praktyce bardzo często stosuje się odwzorowanie jeden-do-jednego.

W tabeli 31.1 przedstawione zostały zależności między wartościami CoS, IP precedence i DSCP. Z zestawienia wynika, że poziom 5 w modelu IP precedence odpowiada wartościom DSCP z przedziału od 40 do 47. Liczbie 40 odpowiada w zapisie binarnym ciąg 101000. Natomiast liczba 47 to binarna wartość 101111. Jeśli jednak przekształcimy do postaci binarnej liczbę 48, otrzymamy ciąg 110000, w którym trzy pierwsze bity zmienią się z 101 na 110. Tym samym zmieni się wartość IP precedence pola TOS z 5 na 6.

Tabela 31.1. Priorytety różnych mechanizmów QoS

CoS	IP precedence	DSCP	Nazwa
0	0	0 – 7	Ruch standardowy (best effort)
1	1	8 – 15	Ruch priorytetowy
2	2	16 – 23	Ruch natychmiastowy
3	3	24 – 31	Ruch błyskawiczny
4	4	32 – 39	Ruch błyskawiczny z zastępowaniem
5	5	40 – 47	Ruch krytyczny
6	6	48 – 55	Zarządzanie siecią internetową
7	7	56 – 63	Zarządzanie siecią

Ponieważ wartości IP precedence i DSCP są tak ściśle związane, można je stosować zamiennie. Na przykład jeśli korzystasz z rozwiązania diffserv (bazującego na kodach DSCP) w sieci, w której działają telefony IP oznaczające pakiety zgodnie z regułami IP precedence, nie musisz się martwić o współdziałanie systemów, ponieważ kod 40 w modelu diffserv odpowiada bitowo poziomowi 5 w mechanizmie IP precedence.

Odmiany mechanizmów QoS

Poniżej wymienione zostały najczęściej wykorzystywane odmiany mechanizmów QoS. W sieciach VoIP zazwyczaj stosuje się rozwiązania CBWFQ (w routerach) oraz CoS (w przełącznikach).

WFQ

WFQ jest domyślnym mechanizmem kolejkowania w połączeniach szeregowych o przepustowości 2 Mb/s i mniejszej. W większości przypadków domyślna konfiguracja dobrze spełnia swoje zadania. Sam mechanizm można bardzo precyzyjnie parametryzować, ale zazwyczaj jego ustawienia nie są w ogóle zmieniane. W sieciach VoIP zamiast rozwiązania WFQ wykorzystuje się najczęściej kolejkowanie o małych opóźnieniach.

Bazujące na klasach ważone kolejkowanie równomierne

Bazujące na klasach ważone kolejkowanie równomierne (CBWFQ — *Class-Based Weighted Fair Queuing*) umożliwia zdefiniowanie klas ruchu, a następnie powiązanie z nimi priorytetów i kolejek. Rozwiązanie CBWFQ jest podstawą kolejkowania o małym opóźnieniu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kolejkowanie priorytetowe

Sposób działania mechanizmu kolejkowania priorytetowego wynika z jego nazwy. Administrator tworzy kolejki i przypisuje do nich poszczególne klasy pakietów (zgodnie z ustalonymi wcześniej priorytetami).

Własne kolejkowanie

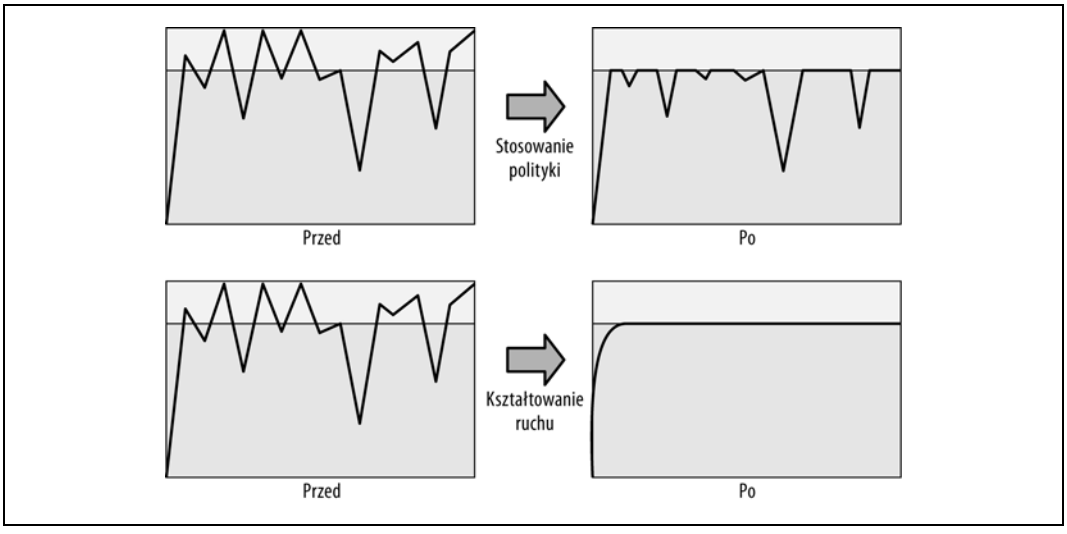
Własne kolejkowanie jest jedną z tych funkcji, których niemal nigdy nie wykorzystuje się w praktyce. Można się z nim spotkać na egzaminie Cisco lub podczas rozwiązywania bardzo nietypowego problemu w sieci, w której transmisja głosu działa poprawnie. Nie oznacza to, że własne kolejkowanie nie jest odpowiednie do obsługi strumieni głosowych, ale w sieciach VoIP zazwyczaj stosuje się kolejkowanie o małym opóźnieniu.

Kolejkowanie o małym opóźnieniu

Kolejkowanie o małym opóźnieniu (LLQ — *Low Latency Queuing*) jest odmianą bazującego na klasach algorytmu WFQ z *kolejką uprzywilejowaną*. W rozwiązaniu tym do przesyłania pakietów wymagających najmniejszego opóźnienia stosuje się układy sprzętowe. Mechanizm jest szczególnie użyteczny w transmisji głosu i sekwencji wizyjnych, w których opóźnienia są szczególnie uciążliwe. Algorytm LLQ jest preferowaną techniką QoS w sieciach telefonii IP i został szczegółowo opisany w dalszej części książki.

Kształtowanie ruchu

Kształtowanie ruchu jest nieco inną formą kolejkowania pakietów. Mechanizmy kształtowania ruchu monitorują przepływ danych, a gdy natężenie ruchu przekroczy pewien poziom progowy, rozpoczynają kolejkowanie pakietów. Operacja kolejkowania jest kontynuowana do czasu, aż zostaną spełnione warunki do wysłania pakietów. Zastosowanie tego rozwiązania pozwala na lepsze wykorzystanie pasma lub dobranie odpowiedniego tempa słania pakietów w łączach o niskiej przepustowości. Głównymi jego zaletami są równomierne wykorzystanie pasma (zobacz rysunek 31.7) oraz buforowanie, a nie odrzucanie pakietów. Do wad trzeba zaliczyć większe zapotrzebowanie na pamięć wynikające z konieczności gromadzenia pakietów. Poza tym przepełnienie bufora oznacza usunięcie pakietów.



Rysunek 31.7. Stosowanie polityki i kształtowanie ruchu

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

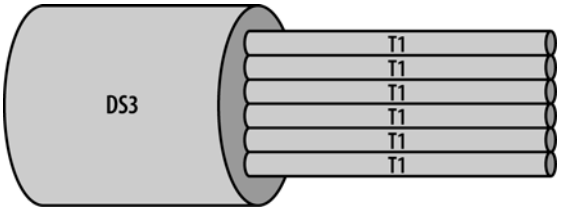
W pierwszym wydaniu książki napisałem, że kształtowanie ruchu nie sprawdza się w transmisji głosu i nie jest często stosowane w nowoczesnych sieciach. Teraz wiem, że się myliłem. Uwielbiam wytykać błędy, szczególnie sobie. Dla jednego z moich klientów kształtowanie ruchu stało się tak istotnym aspektem funkcjonowania sieci, że postanowiłem poświęcić mu specjalny podrozdział.

Zasady kształtowania ruchu

Wykonywałem kiedyś zlecenie dla dużej sieci handlu detalicznego, która posiadała ponad 300 sklepów połączonych z siedzibą główną firmy. Zmienialiśmy technologię połączeń z Frame Relay (o przepustowości 256 kb/s) na MPLS (o pełnej przepustowości T1). W siedzibie firmy łącze DS3 (45 Mb/s) zostało zastąpione łączem OC3 (155 Mb/s). Przygotowywaliśmy jednocześnie sieć do wdrożenia technologii VoIP.

Wkrótce okazało się, że mamy kłopot, ponieważ w chmurze MPLS zaczęły ginąć pakiety. Nie było w tym winy operatora. Problem wynikał z przepustowości łączy na każdym końcu połączenia. Dowiedzieliśmy się, że takie same niedogodności przez lata występowały w sieci Frame Relay. Klient uznał wówczas, że obniżenie wydajności wynika z wyjątkowo niskiej przepustowości łączy. Gdy przeanalizowałem transmisję pakietów, doszedłem do wniosku, że wykorzystywane protokoły nie zajmują nawet połowy pojemności łącza. Dlaczego więc występował problem z wydajnością?

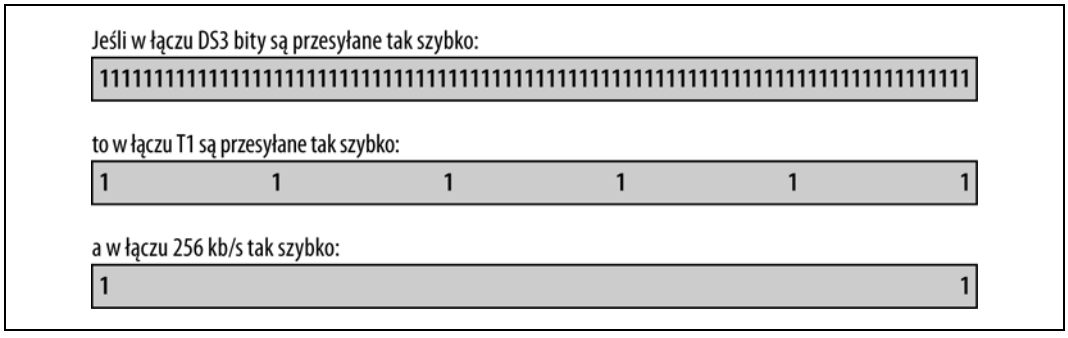
Większość niedogodności w transmisji danych z kształtowaniem ruchu wynika z nieznajomości zasad działania łączy szeregowych. Porównajmy linie T1 i DS3. Łącze DS3 nie jest **grubszą rurą**, jak wielu z nas sądzi. Jest **szybszym** łączem. Nieustannie popełniamy ten sam błąd logiczny, twierdząc, że DS3 jest łączem „większym" od T1, „ponieważ składa się z 28 łączy T1". To jest oczywiście prawda, ale wiele osób, szczególnie menedżerów, sądzi, że rysunek 31.8 jest prawdziwy (gdy taki nie jest). To prowadzi do kolejnego spostrzeżenia, że często błędnie używamy określenia **szerokość pasma**. Tak, nawet ja w swojej książce nadużywam tego słowa. Dlaczego? Ponieważ przyzwyczajenie to druga natura człowieka. Przez ponad 25 lat twierdzimy, że potrzebujemy szerszego pasma, gdy tymczasem potrzebujemy większej **przepustowości**. Jeśli powiem dyrektorowi, że potrzebna jest większa przepustowość, niż ta, jaką zapewnia linia T1, poprosi o wyjaśnienia. Jeśli powiem, że potrzebujemy szerszego pasma, zapyta, ile to będzie kosztowało, i będę mógł szybciej wyjść z biura.



Rysunek 31.8. Szybsze łącza to nie większe łącza

Strumień DS3 może „pomieścić" wiele strumieni T1 tylko dzięki multipleksacji. Łącze DS3 przekazuje dane bit po bicie tak samo jak OC3 i OC192. Szybsze łącza przesyłają bity… szybciej. Różnica między transmisją w łączach DS3, T1 i 256 kb/s (w częściowym łączu T1) została zobrazowana na rysunku 31.9 (bez zachowania skali). Wszystkie połączenia są połączeniami szeregowymi. Wszystkie przesyłają po jednym bicie w danym czasie. DS3 po prostu przesyła je szybciej.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

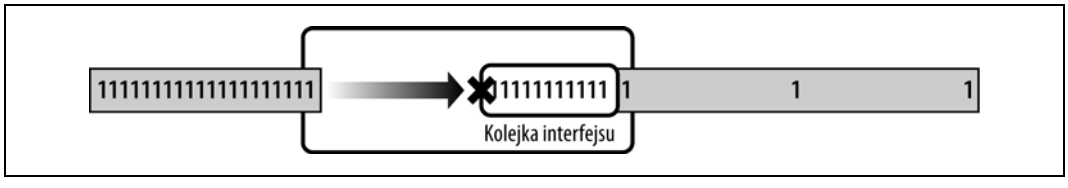


Rysunek 31.9. Szybkość przesyłania bitów w różnych łączach

Problem, który występował w opisywanym przypadku, wiązał się z tym, że oddziały mogły odbierać dane z szybkością 256 **tysięcy** bitów na sekundę, podczas gdy routery w siedzibie głównej wysyłały je z szybkością 45 **milionów** bitów na sekundę. Sieć operatora była tak wydajna, że dostarczała bity z szybkością 45 Mb/s aż do odgałęzienia prowadzącego do oddziału, które miało przepustowość 256 kb/s i przepełniające się bufory (zobacz rysunek 31.10).

Rysunek 31.10. Różnica szybkości transmisji między dwoma łączami

Przyczyna problemu wynikała z pojemności buforów. Urządzenie obsługujące łącze o częściowej przepustowości T1 (działające po stronie operatora) buforowało pakiety, które były wysyłane z siedziby głównej. Jednak dane wychodzące z siedziby głównej były emitowane z taką szybkością, że router dostarczający je do oddziału nie mógł ich wszystkich zbuforować (zobacz rysunek 31.11).

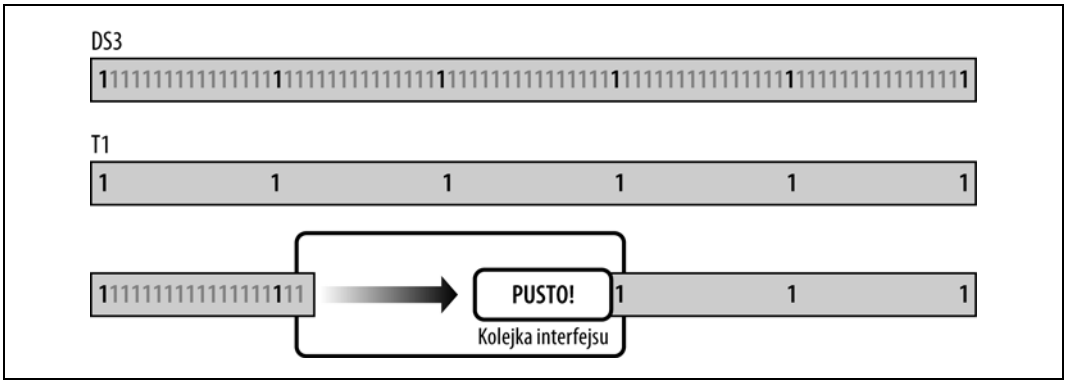


Rysunek 31.11. Zbyt wiele bitów w buforze — odrzucenie bitów

Nie mamy w tym przypadku do czynienia z nadsubskrypcją łącza. Jego wykorzystanie ledwo przekraczało 50 procent, a mimo to liczba utraconych pakietów była dość duża. Nazwaliśmy tę sytuację **problemem mikrozbitek**, ponieważ problem rozpoczynał się i kończył w tak krótkim czasie, że raporty odnoszące się do szerokości pasma go nie wykazywały. Niedogodność ta nie jest tylko cechą protokołu Frame Relay. Może się uwidocznić w każdym systemie, w którym istnieje duża rozbieżność między szybkościami transmisji w dwóch łączach. Zazwyczaj obserwuje się ją w topologiach wielopunktowych lub siatkowych, takich jak Frame Relay lub MPLS (w których wiele ośrodków zdalnych jest przyłączonych do jednego punktu koncentrującego wszystkie połączenia). Jak będziesz się mógł przekonać, czytając kolejny rozdział, rozwiązanie problemu w sieci Frame Relay było łatwe dzięki wykorzystaniu obwodów wirtualnych.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Na szczęście możliwe okazało się również usunięcie naszego problemu. Wystarczyło zastosować kształtowanie ruchu. Skonfigurowaliśmy router w siedzibie głównej tak, aby kolejkował bity wysyłane do poszczególnych oddziałów. Dane były następnie wysyłane z taką szybkością, jaka pozwoliła urządzeniom odległym na ich prawidłowe odbieranie. Wykonaliśmy zadanie, wstawiając bity do strumienia DS3 w predefiniowanych interwałach odpowiadających dokładnie szybkości odbierania danych po drugiej stronie połączenia (zobacz rysunek 31.12).



Rysunek 31.12. Kształtowanie ruchu opóźniające wysyłanie bitów

Kształtowanie ruchu bywa trudne do zrozumienia i zaimplementowania, ale zaznajomienie się z zasadami jego działania pozwala na rozwiązywanie podobnych problemów. Wiem, że jesteś ciekaw, jak to działa w praktyce, ale musisz być cierpliwy. Szczegółowe informacje na temat konfiguracji zamieściłem w kolejnym rozdziale.

Typowe przypadki niezrozumienia mechanizmu QoS

Istnieje wiele nieporozumień na temat możliwości mechanizmu QoS. Jest to rozwiązanie dość skomplikowane, a większość osób tylko pobieżnie przegląda dokumentację i wyciąga wnioski na podstawie przeczytanych informacji. Część zapoznaje się jedynie ze streszczeniem problemu lub, co gorsza, buduje opinię na podstawie opowieści usłyszanych od błędnie poinformowanych osób.

QoS jest bardzo użyteczną technologią, ale nie magią. Nie sprawi, że łącze T1 zacznie pracować jak DS3, i nie spowoduje, że pakiety w ogóle nie będą usuwane. Możliwości mechanizmu QoS zależą od budowy sieci i formułowanych wobec niego wymagań. Jeśli chcesz uzyskać przepustowość kanału DS3, a dysponujesz łączem T1, żadne formy kolejkowania lub rezerwacji nie spełnią Twoich oczekiwań. Celem tego podrozdziału jest więc zaprezentowanie kilku typowych błędów w myśleniu o rozwiązaniach QoS.

QoS dzieli łącze na łącza logiczne o mniejszej przepustowości

Nie zliczę, ile razy słyszałem podobne stwierdzenia na seminariach i spotkaniach dotyczących mechanizmów QoS. Mimo że taka analogia jest użyteczna w niektórych sytuacjach, bywa również groźna, ponieważ sugeruje, że poszczególne „fragmenty” łącza nie mogą zostać wykorzystane

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

do przenoszenia różnego rodzaju ruchu. A to nie jest prawdą. Trzeba pamiętać, że choć harmonogramowanie pakietów zawsze jest stosowane, ograniczenia są nakładane jedynie w czasie przeciążenia sieci.

Przeanalizujmy zagadnienie na przykładzie kolejki LLQ. Pakiety są dostarczane do interfejsu, gdzie są poddawane harmonogramowaniu zgodnie z określoną konfiguracją. Załóżmy, że zdefiniowałeś następujący podział pasma:

- głos: 10 procent,
- sterowanie głosem: 10 procent,
- FTP: 10 procent,
- poczta: 10 procent,
- telnet (SSH): 10 procent,
- pozostały ruch (domyślna kolejka): pozostałe 50 procent.

Kuszące wydaje się założenie, że łącze zostało podzielone na sześć kawałków i że połowa przepustowości jest równomiernie podzielona między głos, sterowanie głosem, FTP, pocztę oraz telnet (SSH), a połowa pozostaje na inne formy komunikacji. Sądzę, że taki sposób przedstawiania problemu wynika z tego, że zagadnienie staje się wówczas łatwiejsze do zrozumienia.

Rzeczywistość wygląda inaczej. Jeśli ruch pocztowy jest jedynym przenoszonym przez łącze w danej chwili, zajmuje 100 procent dostępnego pasma. Ta sama zasada obowiązuje w odniesieniu do innych rodzajów danych. Gdy nie występuje przeciążenie, każdy protokół korzysta z całej dostępnej przepustowości. Załóżmy, że użytkownik rozpoczął pobieranie dużego pliku, używając do tego celu protokołu HTTP. Jeśli żadne inne dane nie są przekazywane w tym samym czasie, operacja pobierania może wysycić całe łącze.

Przyjmijmy jednak, że jednocześnie inny użytkownik zaczął pobieranie pliku za pomocą protokołu FTP. Ponieważ protokołowi FTP przysługuje 10 procent przepustowości, mechanizm harmonogramowania pakietów będzie przekazywał pakiety FTP przed pakietami HTTP. Kolejka FTP będzie obsługiwana przed kolejką domyślną, więc transmisja FTP wpłynie na operację pobierania pliku w protokole HTTP. Bez zagłębiania się w szczegóły działania mechanizmu harmonogramowania odpowiedzmy sobie na typowe pytanie: który z protokołów (FTP czy HTTP) zostanie lepiej obsłużony? Protokołowi FTP zagwarantowano 10 procent pasma. Czy to oznacza, że HTTP otrzyma resztę? Czy nie będzie wówczas lepiej obsługiwany niż FTP?

Odpowiedzi wynikają ze sposobu działania mechanizmu harmonogramowania pakietów. Ponieważ ruch FTP ma wyższy priorytet niż HTTP (który trafia do domyślnej kolejki), pakiety FTP będą przetwarzane przed pakietami HTTP. Różnica w sposobie obsługi strumieni prawdopodobnie pozostanie niezauważona przez użytkowników aż **do chwili wystąpienia przeciążenia**. Jest to bardzo ważne rozróżnienie. Pakiety FTP będą wówczas przetwarzane jako pierwsze, więc istnieje większe prawdopodobieństwo, że pakiety HTTP zostaną odrzucone. Odrzucanie pakietów jest wykonywane tylko w czasie przeciążenia, więc działanie obydwu protokołów będzie standardowe aż do czasu wystąpienia wspomnianego przeciążenia.

Podsumowując, mechanizm QoS umożliwia wskazanie pakietów, które powinny zostać usunięte w czasie przeciążenia łącza. Odnosząc się do wcześniejszego przykładu, można stwierdzić, że po wystąpieniu przeciążenia pakiety HTTP prawdopodobnie będą odrzucane. W normalnych

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

warunkach pakiety FTP również będą przekazywane jako pierwsze, ale w najgorszym przypadku jedyną konsekwencją takiego sposobu działania będzie nieznacznie dłuższe przechowywanie pakietów HTTP w kolejce.

QoS wyznacza szerokość pasma

Bądźmy rzeczowi. Mechanizm QoS nie **wyznacza** szerokości pasma, ale je **gwarantuje**, a to nie to samo. Jeśli dysponujesz łączem T1 o przepustowości 1,5 Mb/s i chcesz przesłać strumień FTP o przepływności 1 Mb/s, konfiguracja gwarantująca 10 procent pasma nie ograniczy szybkości transmisji danych FTP do 10 procent dostępnej przepustowości.

Ponownie przypominam, że gdy przeciążenie nie występuje, każdy protokół może zająć taką część pasma, jaka jest mu potrzebna. Jeśli zdefiniuję 20 protokołów w łączu DS3 i przypiszę każdemu z nich przepływność 1 Mb/s, każdy z protokołów i tak będzie mógł wykorzystać całą szerokość pasma, o ile oczywiście nie będzie przeciążenia.

QoS rozwiązuje problem niedostatecznej szerokości pasma

Zgodnie z wcześniejszym stwierdzeniem QoS nie sprawi, że łącze T1 stanie się „tak dobre" jak DS3. Jeśli wystąpił problem nadsubskrypcji łącza, najlepszym rezultatem działania mechanizmu QoS może być określenie, które pakiety będą przekazywane w pierwszej kolejności, co jest tożsame ze wskazaniem pakietów do odrzucenia! Tak, to prawda. Ułatwienie przesyłania pakietów określonych protokołów powoduje utrudnienie transmisji innych. Pamiętaj jednak o zastrzeżeniu: **tylko w czasie przeciążenia**.

Jeśli korzystasz z łącza T1 i starasz się przesłać nim dane z szybkością 20 Mb/s, nie ma znaczenia, w jaki sposób zdefiniujesz priorytety i kolejki, pakiety wszystkich rodzajów danych i tak będą odrzucane.

Oczywiście to, że QoS nie rozwiązuje problemu niedostatecznej szerokości pasma, nie oznacza, iż „zwiększenie szerokości pasma zawsze rozwiązuje problem". Zrozumienie tej zasady czyni administratora wartościowym dla firmy inżynierem. Jeśli wykorzystujesz łącze T1, w którym połowa pakietów jest odrzucana, to przy założeniu, że nie ma problemów w warstwie fizycznej, najprawdopodobniej zwiększenie przepustowości łącza jest konieczne.

Jeśli wykorzystujesz łącze T1 do przenoszeniu ruchu VoIP, a prowadzone rozmowy mają niską jakość, zwiększenie przepustowości nie musi rozwiązać problemu, ponieważ nie jest on zależny jedynie od szerokości pasma. Znowu najważniejszym elementem jest w tym przypadku przeciążenie. Jeśli rozmowy mają niską jakość również w czasie, gdy łącze nie jest przeciążone, konieczne jest uruchomienie mechanizmu QoS, który zapewni przekazywanie pakietów głosowych w pierwszej kolejności.

QoS zapobiega usuwaniu pakietów

Jeśli masz problem z masowym usuwaniem pakietów, techniki QoS go nie rozwiążą. Jedyne, na co mogą mieć wpływ, to wyróżnienie ruchu, który będzie przenoszony bez strat, co jednak oznacza, że mniej istotne pakiety nadal będą tracone.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kształtowanie ruchu **może** zapobiec usuwaniu pakietów (zakładając, że bufory tego mechanizmu nie są przepełnione), choć nie jest to rozwiązanie proste. Poza tym działanie algorytmu polega na buforowaniu pakietów przez pewien czas. Jeśli więc łącze jest przeciążone przez cały dzień, kształtowanie ruchu będzie bezużyteczne.

QoS zwiększa atrakcyjność dla płci przeciwnej

Niestety obawiam się, że nie jest to prawda. Choć teoretycznie jest prawdopodobne, że Twoja przyszła „druga połowa" właśnie czyta tę książkę!

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 32.

Projektowanie QoS

Zaprojektowanie schematu QoS może być prostym zadaniem lub też dużym projektem. Skala rozwiązania zależy od złożoności sieci oraz potrzeb biznesowych. Moją podstawową zasadą jest planowanie działań wtedy, kiedy w myślach widzę gotowe rozwiązanie. Nawet jeśli planuję schemat QoS dla firmy, która obecnie nie posiada VoIP, staram się zaprojektować rozwiązanie pozwalające na zastosowanie tej technologii w przyszłości.

W tym rozdziale przedstawię projektowanie schematu QoS dla sieci wykorzystującej VoIP. Jest to proces składający się z dwóch etapów. W pierwszym musisz określić wymagania dla systemu, w drugim skonfigurować routery. Skorzystamy z rekomendowanego przez Cisco mechanizmu LLQ (ang. *low latency queuing*). Przedstawię również dwa scenariusze dla kształtowania ruchu i pokażę kroki związane z implementacją.

Scenariusz LLQ

Pierwsza część procesu projektowania QoS wymagać będzie od Ciebie małego śledztwa. Powinieneś przeprowadzić wywiady ze wszystkimi jednostkami biznesowymi korzystającymi z sieci i określić, które protokoły są dla nich istotne. Załóż od razu, że każda z jednostek powie Ci, iż jej protokoły są najważniejsze i powinny zajmować większość dostępnego pasma. Musisz więc zasymilować zebrane informacje i podjąć decyzję o przydziale pasma. Z pomocą zdrowego rozsądku powinieneś być w stanie stworzyć listę wymagań, z którą będziesz mógł rozpocząć pracę.

Protokoły

Stwórz listę protokołów używanych w Twojej sieci. Przydatne przy tym mogą się okazać narzędzia analizujące pakiety, jak Ethereal, technologia NetFlow w routerach Cisco lub moduły NAM (*Network Analysis Module*) w przełącznikach 6500. Notuj wszystko skrupulatnie. Weź pod uwagę, że przedsiębiorstwa często zmieniają domyślne porty popularnych protokołów. Przykładowo ruch generowany przez przeglądarkę może wykorzystywać porty 80, 443 oraz 8080.

Zanotuj, kto używa jakiego protokołu i jak się ten protokół zachowuje. Niektóre z nich bowiem mogą być podstępne. Na przykład RTP wykorzystywany jest do przesyłania strumienia głosowego w technologii VoIP firmy Cisco. Jeśli założysz, że wszystkie pakiety RTP powinny mieć

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

najwyższy priorytet, tak aby zapewnić odpowiednią jakość rozmów, mogą pojawić się nieprzewidziane konsekwencje. Inne aplikacje, jak na przykład strumieniowa transmisja wideo, mogą również korzystać z RTP. Jeśli więc zaprojektujesz schemat QoS, który będzie najwyżej priorytetyzował ruch RTP bez wprowadzania dodatkowej granulacji, konferencje wideo mogą spowodować problemy związane z połączeniami głosowymi. Jeżeli wiesz, że konferencje wideo odbywają się jedynie w sali zarządu, możesz odpowiednio priorytetyzować strumieniowanie wideo, tak aby uniknąć problemów.

Priorytety

Określenie priorytetów nie jest proste. Na tym etapie polityka może odgrywać decydującą rolę — bądź na to przygotowany.

Istnieją niezmienne zasady wskazujące na dobre praktyki dla QoS. Na przykład:

- Nie powinieneś ustawiać klasy usług (ang. *IP precedence*) na 6 lub 7, ponieważ są one zarezerwowane dla **Internetwork Control** oraz **Network Control**.
- Pakiety głosowe RTP (ang. *voice RTP*) powinny być oznaczone klasą 5 (DSCP: EF).
- Wszystkie pakiety protokołów sterujących dla połączeń głosowych powinny być oznaczone klasą 3 (DSCP: AF31).
- Ruch danych powinien mieć mniejszy priorytet aniżeli pakiety głosowe RTP oraz pakiety sterujące.

Zauważ, że jeśli będziesz przestrzegał powyższych dobrych praktyk, przebrniesz dość szybko przez poziomy klas usług. Z oryginalnych ośmiu pozostają jedynie 0, 1, 2, 3 i 4. Klasa 4 powinna być stosowana dla krytycznych aplikacji, które jednak nie są tak krytyczne jak przesyłanie głosu, a więc przykładowo dla konferencji wideo. Klasa usług 0 (zasadniczo bez priorytetu) używana jest dla ruchu, który można określić jako ruch typu *best effort*. Pozostają jedynie klasy 1 i 2. Jak widać, klasy usług nie skalują się zbyt dobrze. Generalnie lepiej jest stosować *diffserv*, tam gdzie jest to możliwe.

Nie zapomnij o sobie! Kiedy łącza są obciążone, a Ty musisz rozwiązywać problemy na routerze znajdującym się na drugim końcu Twojej sieci, nie chcesz, aby Twój ruch telnet lub SSH był odrzucany kosztem użytkowników ściągających pliki z wykorzystaniem protokołu FTP. W związku z tym, że ruch telnet oraz SSH nie wymagają dużej przepustowości, wpływ nadania im wysokiego priorytetu będzie niewielki, a znaczenie takiego posunięcia w sytuacjach kryzysowych może okazać się ogromne.

W tym momencie powinieneś już posiadać uporządkowaną listę opisującą, jak chciałbyś priorytetyzować swoje protokoły. Na jej samym dole powinien się znajdować ruch typu *best effort*. Będzie to docelowo Twoja kolejka domyślna. Dobrym pomysłem jest użycie wpisu „do ustalenia" opisującego normalny ruch sieciowy. Zakładając, że na samej górze znajdzie się ruch o najwyższym priorytecie, lista może wyglądać następująco:

- głosowe pakiety RTP;
- pakiety sterujące w połączeniach głosowych (ang. *voice control*);
- telnet/SSH (wiele protokołów może współdzielić ten sam priorytet oraz kolejkę);
- „do ustalenia";
- ruch pozostały (*best effort*).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ogólnie rzecz biorąc, powinieneś opierać się namowom związanym z umieszczeniem ruchu HTTP w kolejce na innym poziomie aniżeli ruch pozostały. HTTP przeważnie nie jest krytyczne, chyba że mamy do czynienia ze specyficznym środowiskiem. Na przykład w centrum obsługi telefonicznej, gdzie agenci odpowiedzialni za kontakty z klientami stosują przeglądarkowe aplikacje działające na zdalnym serwerze, priorytetyzowanie ruchu HTTP jest w pełni uzasadnione. Skorzystamy z tego przykładu i założymy, że każdy ruch z lub do serwera (`10.10.100.100`) na porcie 80 powinien być priorytetyzowany. W takim przypadku lista będzie wyglądać następująco:

- głosowe pakiety RTP;
- pakiety sterujące w połączeniach głosowych;
- telnet/SSH;
- HTTP z/do `10.10.100.100`;
- „do ustalenia”;
- ruch pozostały (*best effort*).

Zauważ, że ruch telnet/SSH znajduje się powyżej ruchu dla centrum obsługi telefonicznej. Pamiętaj, aby na wypadek awarii zawsze mieć takiego właśnie asa w rękawie.

Naszym kolejnym obiektem rozważań jest e-mail, który może zużyć sporą część pasma w sieci. W wielu starszych (i ciągle używanych) środowiskach Microsoft Exchange foldery prywatne poczty przechowywane są w plikach PST (rozszerzenie *.pst*). Możesz mieć użytkowników, którzy swoje pliki *.pst* przechowują na serwerach plików w zdalnych lokalizacjach na drugim końcu łącza. Osobiście spotkałem się z problemem spowodowanym otwieraniem pliku *.pst* o rozmiarze 1 GB poprzez WAN. Co więcej, użytkownicy posiadający takie same konfiguracje przychodzili do pracy mniej więcej o tej samej porze i próbowali otwierać swoje ogromne pliki *.pst* w tym samym czasie każdego ranka. Sieć działała wtedy bardzo wolno. Mimo to zarządzający nalegali na priorytetyzowanie ruchu e-mail. Zgodnie z zasadą, że ostatnie zdanie należy do zarządu, poniżej kolejka dla naszego przykładu:

- głosowe pakiety RTP;
- pakiety sterujące w połączeniach głosowych;
- telnet/SSH;
- HTTP z/do `10.10.100.100`;
- „do ustalenia”;
- e-mail (SMTP, POP, IMAP, Exchange);
- ruch pozostały (*best effort*).

Ze swojego doświadczenia wiem, że problemy związane z ruchem e-mail wynikają zazwyczaj ze słabego egzekwowania dopuszczalnych reguł (lub nawet ich braku). Na przykład jeden użytkownik, wysyłając wiadomość z załącznikiem o rozmiarze 20 MB, może w sposób znaczący wpłynąć na całą sieć.

Po usunięciu pozycji „do ustalenia” otrzymujemy finalną listę protokołów w porządku malejącego priorytetu:

- głosowe pakiety RTP;
- pakiety sterujące w połączeniach głosowych;
- telnet/SSH;

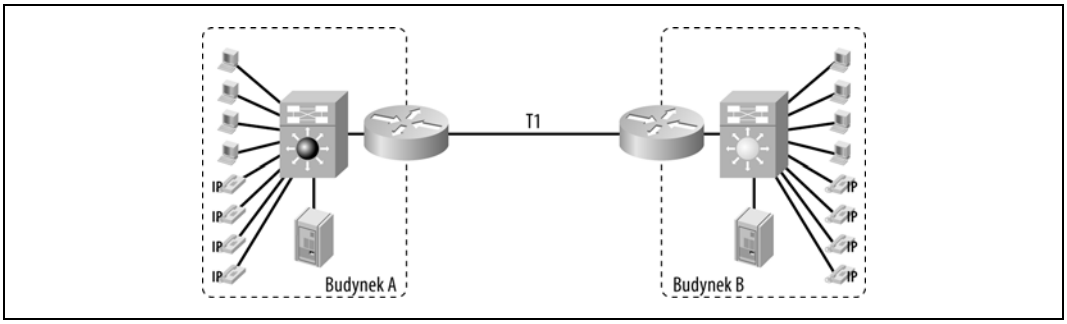
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

- HTTP z/do `10.10.100.100`;
- e-mail (SMTP, POP, IMAP, Exchange);
- ruch pozostały (*best effort*).

Powyższa lista jest dobrą reprezentacją realnych środowisk, z którymi się spotkałem. Jej implementacja wskaże, jak oznaczać na różne sposoby różne typy ruchu.

Określenie wymagań dla pasma

Aby zilustrować, jak określić wymagania pasma dla różnych protokołów, wykorzystamy przykładową sieć z poprzedniego rozdziału przedstawioną na rysunku 32.1.



Rysunek 32.1. Sieć w dwóch budynkach połączonych pojedynczym łączem T1

Dla VoIP preferowanym mechanizmem QoS jest LLQ. Stosując go, możesz stworzyć pojedynczą kolejkę dla ruchu o wysokim priorytecie, czyli takiego, który nie może ucierpieć w wyniku utraty pakietów (pakiety audio, wideo itp.), oraz kolejkować pozostały ruch na zasadzie hierarchii priorytetów.

Kolejce priorytetowej może być przypisany procent dostępnego pasma (np. 15%) lub też określony rozmiar pasma (np. 300 kb/s). Nie powinna ona konsumować więcej niż Ï dostępnego pasma. W rzeczywistości Cisco zapobiega alokowaniu kolejce priorytetowej więcej niż 75% pasma, choć ta wartość może być zmieniana za pomocą komendy interfejsu `max-reserved-ban ➥dwidth`:

```
Router(config-if)# max-reserved-bandwidth 80
```

Łącza WAN nie powinny być budowane przy założeniu planowanej utylizacji powyżej 75%. Powinieneś zawsze pozostawić 25% na narzut, protokoły routingu, administrację itp. Specyfika harmonogramowania powoduje, że jeśli przypiszesz nadzwyczaj dużą wartość kolejce priorytetowej, żadna z pozostałych kolejek nie zostanie obsłużona.

Przypiszmy pakiety głosowe RTP do kolejki priorytetowej. Na potrzeby ćwiczenia załóżmy, że musimy alokować przestrzeń dla czterech połączeń VoIP G729a CODEC (50 pakietów na sekundę) możliwych do wykonania w każdym momencie. Szybkie rozpoznanie pokazuje, że każde z połączeń nie zabierze więcej niż 37 kb/s. Cztery połączenia po 37 kb/s to 148 kb/s, które to przypiszemy kolejce priorytetowej:

Pakiety głosowe RTP — kolejka priorytetowa — 148 kb/s

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Sterowanie połączeniem (ang. *call control*) wymaga bardzo wąskiego pasma. Niektóre dokumentacje Cisco stwierdzają, że dla jednego telefonu powinniśmy zagwarantować 150 b/s (plus narzut). Dla czterech połączeń będzie to jedynie 600 b/s. Inne dokumentacje podają, że 2 – 5% całego pasma powinno być zaalokowane dla ruchu typu *call control*. Dla T1 to około 30 – 77 kb/s. Przyjmijmy pewien kompromis i załóżmy, że potrzebujemy 30 kb/s na pakiety sterujące:

Pakiety sterujące w połączeniach głosowych — 30 kb/s

Priorytetyzujemy ruch telnet/SSH wyłącznie na potrzeby nagłych wypadków, tak więc ruch ten nie potrzebuje sporego pasma. Alokowanie dla niego 64 kb/s w zupełności wystarczy, a rozmiar ten nie będzie wpływał negatywnie na inne protokoły:

Telnet/SSH — 64 kb/s

Ruch HTTP z i do `10.10.100.100` w naszej sieci jest krytyczny, dlatego przeznaczymy na jego potrzeby znaczną część pasma. Przyjmijmy, że sprawdziliśmy, iż ruch HTTP waha się w okolicach 300 kb/s. Patrzymy w przyszłość, dlatego przyznajemy mu 384 kb/s:

HTTP z/do `10.10.100.100` — 384 kb/s

Priorytetyzację poczty elektronicznej wykonujemy tylko po to, aby zadowolić zarząd, tak więc alokujemy jej niewielką część pasma: 128 kb/s. To powinno pozwolić na poprawne działanie, a tym samym pozostawić miejsce dla innych aplikacji:

E-mail (SMTP, POP, IMAP, Exchange) — 128 kb/s

Po tym, gdy wszystkie kolejki otrzymają swoje przydziały, domyślnej kolejce zostaje przyznana pozostała część pasma. Dla łącza T1 o przepustowości 1544 kb/s i kolejki, jakie sobie zdefiniowaliśmy, otrzymujemy następujący rezultat: 1544–(148+30+64+384+128) = 790 kb/s, zatem ostatnia pozycja to:

Pozostały ruch (*best effort*) — reszta (790 kb/s)

Nasza finalna lista przedstawia się następująco:

- pakiety głosowe RTP — kolejka priorytetowa — 148 kb/s;
- pakiety sterujące w połączeniach głosowych — 30 kb/s;
- telnet/SSH — 64 kb/s;
- HTTP z/do `10.10.100.100` — 384 kb/s;
- e-mail (SMTP, POP, IMAP, Exchange) — 128 kb/s;
- pozostały ruch (*best effort*) — reszta (790 kb/s).

Konfiguracja routerów

Wykonaliśmy już całą pracę związaną z rozpoznaniem i planowaniem. Czas na praktykę, czyli konfigurację routerów. W tej implementacji skonfigurujemy QoS na routerach brzegowych łączących budynki łączem T1. Na tym samym urządzeniu wykonamy znakowanie (ang. *marking*), klasyfikację (ang. *classification*) oraz ograniczanie (ang. *policing*). Pamiętaj, że można to zrobić na osobnych urządzeniach, jeśli zajdzie taka potrzeba. Telefony VoIP z reguły znakują swoje pakiety, zatem założymy, że tak właśnie postępują też w naszej sieci.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Aby skonfigurować LLQ, musimy skonfigurować trzy elementy: mapy klas (ang. *class maps*), mapy polityk (ang. *policy maps*), polityki usług (ang. *service policies*).

Mapy klas

Musimy zidentyfikować pakiety, które powinny być znakowane. Aby to zrobić, stworzymy **mapy klas**, które wyłapią właściwy typ ruchu. Mapy klas mają wiele możliwości dopasowywania — wg wartości *IP precedence*, DSCP, a nawet list ACL. Jedyną funkcją, jaką wykonują mapy klas, jest właśnie dopasowywanie.

Poniższa mapa, nazwana `Voice-Calls`, dopasuje DSCP EF (ang. *express forwarding*) albo klasę usług 5:

```
class-map match-any Voice-Calls
      description [---[ Rozmowy głosowe (Zestawiane na telefonach IP i koncówkach
      ↪telefonicznych) ]---]
  match ip dscp ef
  match ip precedence 5
```

Technicznie zapis powyższy jest zbędny, ponieważ (jak to opisano w rozdziale 31.) oba zapisy są równoznaczne. Zmieńmy to tak, aby ułatwić odczytanie:

```
class-map match-any Voice-Calls
      description [---[ Rozmowy głosowe (Zestawiane na telefonach IP i koncówkach
      ↪telefonicznych) ]---]
  match ip precedence 5
```

Równie dobrze moglibyśmy użyć DSCP EF, ale aktualna konfiguracja jest prostsza, więc będziemy się trzymać *IP precedence*.

To, co teraz musimy zrobić, to stworzyć mapy klas dla pozostałych pozycji z naszej listy. Pakiety sterujące w połączeniach głosowych:

```
class-map match-any Voice-Control
      description [---[ Pakiety sterujące w połączeniach głosowych (Zestawiane na menedżerze wywołań
      ↪oraz telefonach IP) ]---]
  match ip precedence 3
```

Call Manager, bramki VoIP oraz telefony IP są w stanie za nas ustawić swoją klasę usług. Skorzystamy z tego, zakładając, że wszystkie kontrolne pakiety głosowe zostały oznaczone klasą usług 3.

Kolejna pozycja na naszej liście to telnet oraz SSH. W tym przykładzie w mapie klasy dopasujemy protokoły po nazwach:

```
class-map match-any Telnet-Traffic
  description [---[ telnet & SSH ]---]
  match protocol telnet
  match protocol ssh
```

W przypadku HTTP nie wystarczy nazwa protokołu. Nie możemy po prostu napisać `match protocol http`, ponieważ wtedy dopasowywalibyśmy ruch HTTP z każdego źródła do każdej lokalizacji. Nasze wymagania wskazują, że powinniśmy dopasowywać ruch HTTP (tylko port 80) skierowany do lub z adresu `10.10.100.100`. Z pomocą przychodzą listy kontroli dostępu:

```
class-map match-any HTTP-Call-Center
  description [---[ Call Center HTTP ]--]
  match access-group 121
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Aby dopasować wyrażenie `access-group`, musimy stworzyć listę ACL:

```
access-list 121 remark [---[ Call Center HTTP ]---]
  access-list 121 permit tcp any 10.10.100.100 0.0.0.0 eq www
  access-list 121 permit tcp any eq www 10.10.100.100 0.0.0.0
```

Listy ACL pozwalają na dużo bardziej elastyczne stosowanie map klas. Zapewniają m.in.:

- Tworzenie czasowych list kontroli dostępu (ang. *time-based ACL*) (różne QoS dla różnych pór dnia).
- Tworzenie list ACL opartych na adresach IP (ang. *IP-based ACL*) (wskazany adres IP ma priorytet).
- Możliwość usunięcia priorytetu dla jednego użytkownika z grupy, która ma priorytet (wstawienie `deny` na samym szczycie, a poniżej `permit` dla grupy).
- Tworzenie dynamicznych list kontroli dostępu („Jeśli używam SSH, nie faworyzuj pakietów SSH Bogdana").

Dzięki listom dostępu masz praktycznie nieograniczone możliwości, bądź jednak ostrożny — im bardziej skomplikowane są Twoje reguły, tym trudniej nimi zarządzać i je modyfikować.

Dla pozycji e-mail na naszej liście ponownie możemy skorzystać z nazwy protokołu. W związku z tym, że e-mail wykorzystuje wiele protokołów, wypiszemy je wszystkie. Router jest na tyle inteligentny, że będzie w stanie przypisać właściwy typ ruchu do wyrażenia `exchange`:

```
class-map match-any Mail-Traffic
  description [---[ Ruch email (Zestawiany na ruterze WAN) ]---]
  match protocol smtp
  match protocol pop3
  match protocol imap
  match protocol exchange
```

Mapy polityk

Teraz, kiedy mamy już zdefiniowane mapy klas, musimy utworzyć mapy reguł. Podczas gdy mapy klas wskazują, **na które** pakiety nałożymy pewne polityki, mapy polityk określą, **jak** to zrobimy. Krok ten jest niezwykle prosty, jeśli odpowiednio wcześniej wszystko zostało zaplanowane tak, jak to zrobiliśmy w naszym przypadku.

Jako że dysponujemy pojedynczym łączem T1, potrzebna będzie nam wyłącznie jedna mapa polityk. Nazwijmy ją `WAN-Link`:

```
policy-map WAN-Link
  description [---[ Apply Treatment for Classes ]---]
```

Jak zapewne zauważyłeś, bardzo często używam polecenia `description`. Uważam, że jest to jedna z najlepszych rzeczy, jakie przytrafiły się oprogramowaniu IOS firmy Cisco. Jak już mówiłem, co proste, to dobre, a na dobrze udokumentowanym kodzie możemy jedynie zyskać. Zachęcam Cię do opisywania swoich konfiguracji zawsze, kiedy masz taką możliwość.

Mapa polityk to po prostu lista wszystkich map klas, gdzie każda z map ma przypisaną komendę. Najpierw stwórzmy mapę polityk dla klasy `Voice-Calls`:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
policy-map WAN-Link
  description [---[ Alokacja pasma dla klas ]---]
  class Voice-Calls
    priority 148
```

Użycie komendy `priority` poinformuje router, aby przypisał klasę do kolejki priorytetowej. Liczba po komendzie `priority` stanowi przydzielony rozmiar pasma dla kolejki. W tym przypadku alokujemy 148 kb/s na potrzeby kolejki priorytetowej. Istnieją jeszcze dodatkowe podkomendy, zachęcam do ich przestudiowania, choć w omawianym przypadku są one zbędne.

Niektóre nowe wersję IOS pozwalają Ci przydzielać łącze z wykorzystaniem procentów. Aby to zrobić, należy skorzystać z komendy `priority percent 10` (w tym przypadku 10 procent pasma z całego łącza zostanie przydzielone kolejce priorytetowej).

Jeśli już skonfigurowaliśmy kolejkę priorytetową, zajmijmy się pozostałymi kolejkami. Nie musimy już korzystać z polecenia `priority`. Od tego momentu będziemy przydzielać pasmo, używając komendy `bandwidth` w każdej z klas. Poniższa komenda przyznaje 30 kb/s pasma klasie `Voice-Control`:

```
class Voice-Control
  bandwidth 30
```

W nowszych wersjach IOS możesz określić procent pozostałego pasma dla każdej z kolejek niemających priorytetu. Na przykład jeśli przyznasz 10% kolejce priorytetowej i zechcesz przeznaczyć 5% pozostałego pasma (czyli tego, które pozostało po wykorzystaniu 10%) klasie `Voice-Control`, możesz skorzystać z poniższych poleceń:

```
class Voice-Control
  bandwidth remaining percent 5
```

Każda z pozostałych klas zostanie skonfigurowana w ten sam sposób, oczywiście przy uwzględnieniu odpowiedniego przydziału pasma:

```
class Telnet-Traffic
  bandwidth 64
class HTTP-Call-Center
  bandwidth 384
class Mail-Traffic
  bandwidth 128
```

Na końcu należy skonfigurować kolejkę domyślną. Aby być pewnym działania mechanizmu kolejki ważonej WFQ dla pozostałego ruchu, należy skonfigurować przypisanie do klasy `class-default`:

```
class class-default
  fair-queue
```

`class-default` jest nazwą zarezerwowaną, nie możesz zatem utworzyć takiej klasy, ale możesz przypisać do niej mapę reguł.

Polityki usług

Komenda `service-policy` aktywuje mapy klas oraz mapy reguł na interfejsie. Naszym zadaniem jest poprawienie wydajności łącza T1, dlatego polecenie `service-policy` wykonamy na routerach WAN:

```
int S0/0
  service-policy output WAN-Link
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Polityki usług uruchamiane są dla wyjścia interfejsu. Pamiętaj, że staramy się wpłynąć na to, jak pakiety są wysyłane na łącze. Dla tych już przesłanych, po stronie odbiornika, nie jesteśmy w stanie nic zrobić.

Musisz uważać, kiedy aplikujesz *service-policy* na interfejsie. Uruchomienie nowych mechanizmów kolejkowania wymaga wyczyszczenia interfejsu. Nie będzie to problemem, jeśli chodzi o przesyłanie „tradycyjnych" danych, lecz w przypadku połączeń głosowych może nastąpić rozłączenie, a co najmniej odczuwalna przerwa w połączeniu.

Po aktywowaniu polityki usług wszystkie stworzone dotąd reguły zaczną działać.

Finalna konfiguracja przedstawia się następująco:

```
class-map match-any Mail-Traffic
  description [---[ Ruch email (Zestawiany na routerze WAN) ]---]
  match protocol smtp
  match protocol pop3
  match protocol imap
  match protocol exchange
!
policy-map WAN-Link
  description [---[ Alokacja pasma dla klas ]---]
  class Voice-Calls
    priority 148
  class Voice-Control
    bandwidth 30
  class Telnet-Traffic
    bandwidth 64
  class HTTP-Call-Center
    bandwidth 384
  class Mail-Traffic
    bandwidth 128
  class class-default
    fair-queue
!
int S0/0
  service-policy output WAN-Edge
```

Scenariusze kształtowania ruchu

Istnieją dwa scenariusze, w których odnalazłem zalety zastosowania kształtowania ruchu (ang. *traffic shaping*). Pierwszy uwzględnia zastosowanie zakończenia ethernetowego z bardzo szybką transmisją, jak na przykład OC3, a drugi użycie Frame Relay z dużą liczbą węzłów. Zacznijmy od pierwszego.

Scenariusz 1.: Zakończenie ethernetowe

Aktualnie operatorzy telekomunikacyjni dostarczają już usługę zakończenia ethernetowego dla tradycyjnych łączy. Ma to znaczący wpływ na koszty, jakie ponosi użytkownik końcowy. Weźmy na przykład łącze OC3. Aby terminować OC3, będziesz potrzebował routera wspierającego karty OC3. Takie routery oraz wymagane karty są nieco droższe niż typowe routery używane w zdalnych lokalizacjach. Jeżeli możesz otrzymać takie same łącze OC3 przekonwertowane przez dostawcę na łącze typu *Gigabit Ethernet*, będziesz w stanie podpiąć się pod dowolny ze swoich dotychczas używanych przełączników.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Istnieją pewne wady takiego rozwiązania, które nie są widoczne na pierwszy rzut oka. Pierwszą jest QoS. Jeśli na Twoim łączu wymagane jest LLQ, możesz nie być w stanie zapewnić go w sposób, jaki przedstawiłem. Pamiętaj, przełączniki zaprojektowane są na potrzeby sieci LAN — wykorzystują nie QoS, lecz CoS, dlatego właśnie komendy, które opisałem w poprzednim rozdziale, nie będą w stanie zaimplementować schematu QoS. Jeśli jest on wymagany, od samego początku miej to na uwadze.

Popełniłem pewien błąd, tak abyś Ty nie musiał tego robić. Planowaliśmy wdrożenie MPLS o zasięgu krajowym i zamówiliśmy końcówkę ethernet z założeniem, że jedyne, co będziemy musieli zrobić, to podłączyć się do naszych przełączników. Kiedy rozpoczęliśmy implementowanie komend QoS, zwyczajnie nie chciały one działać. Szybki telefon do Cisco potwierdził, że byłem idiotą, sądząc, iż wykonam te komendy na przełączniku; potrzebowałem routera. Interfejs nie miał znaczenia. Sprawę zakończyliśmy, kupując router i wykonując komendy na interfejsie Ethernet. Problemem było pójście do dyrektora IT i stwierdzenie, że zakładane oszczędności były niemożliwe do zrealizowania i że teraz potrzebujemy routerów. Dyrektorzy IT uwielbiają, kiedy mówi się im, że zespół sieciowców potrzebuje więcej pieniędzy.

Kiedy już zdobyliśmy odpowiedni sprzęt, pojawiła się jeszcze jedna rzecz, o której nie pomyśleliśmy. Był to znowu problem związany z rozbieżnością prędkości.

Dostawcy często zapewniają zakończenie ethernet poprzez umieszczenie w Twojej lokalizacji urządzenia z OC3 (jak w tym przypadku) z jednej strony, a z *Gigabit Ethernet* z drugiej. Problemem jest to, że będziesz wysyłał pakiety do tego urządzenia z prędkością **miliarda** bitów na sekundę, podczas gdy OC3 może je wysyłać z prędkością 155 **milionów** bitów na sekundę. Podobnie jak w scenariuszu z Frame Relay przedstawionym w rozdziale 31. interfejs OC3 będzie kolejkował tylko tyle pakietów, ile będzie potrafił wysłać, pozostałe odrzuci. Wskaźnik nadmiaru wynosi zatem 10:1,5, co skutkować będzie prawdopodobnie sporą liczbą porzuconych pakietów.

Rozwiązaniem było kształtowanie ruchu dla ruchu przychodzącego z naszej strony na *Gigabit Ethernet*, co spowodowało znaczne spowolnienie. Taki ruch OC3 było już w stanie odebrać (i przesłać dalej).

Pierwszym krokiem jest stworzenie mapy polityk. W związku z tym, że nie stosujemy poleceń CoS oraz QoS, jest to bardzo prosta mapa polityk. Istnieje wiele możliwości skomplikowania sposobu zapisu, ale to, co właściwie chcemy powiedzieć interfejsowi, to: „wykonaj polecenie `shape average` dla żądanej szybkości i zastosuj je dla domyślnej klasy (ang. *default class*)”:

```
policy-map Provider-Outbound
    description [--[ Dla zakończenia ethernetowego 150M ]--]
  class class-default
    shape average 150000000
```

Jeśli chcesz dodać inne komendy CoS, musisz umieścić je po komendzie `shape average`. Jeśli umieścisz je przed nią, CoS zostanie zastosowany, zanim pojawi się kształtowanie ruchu, a tym samym pakiety będą gubione, kiedy bufory o rozmiarze 155 Mb/s zostaną zalane ruchem 1 Gb/s. Na szczęście mapy polityk możemy umieszczać kaskadowo, więc w naszym przypadku pod klasą `default-class`. Coś podobnego zrobiłem w rozdziale 30., gdzie zastosowałem QoS dla podinterfejsu:

```
policy-map Provider-Outbound
  description [--[ Dla zakończenia ethernetowego 150M ]--]
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
class class-default
  shape average 150000000
  service-policy Police-CoS
```

Oczywiście dla powyższego przykładu musiałbyś mieć gotową mapę reguł `Police-CoS`. Takie rozwiązanie miałoby zastosowanie, jeśli posiadałbyś zakończenie ethernet do chmury MPLS z gotowymi poziomami QoS. Pójdźmy jeszcze dalej i przyjrzyjmy się temu, jak może wyglądać mapa polityk dla CoS. Zapamiętaj, że jest to część rozwiązania używana w rzeczywistych zastosowaniach, mimo iż nie ma ona nic wspólnego z kształtowaniem ruchu. Oto prosta mapa polityk odpowiadająca poziomom usług dostarczonym przez jednego z dostawców:

- kolejka priorytetowa: 60%,
- druga kolejka: 20%,
- trzecia kolejka: 10%,
- *best effort*: pozostałe 10%.

Poniższy kod przedstawia polityki oraz mapy klas niezbędne do implementacji tych poziomów:

```
class-map match-any CoS1
  match dscp ef
class-map match-any CoS2
  match dscp af31
  match dscp af32
class-map match-any CoS3
  match dscp af21
  match dscp af22
!
policy-map Police-CoS
  class CoS1
   set dscp ef
   priority percent 60
  class CoS2
    bandwidth remaining percent 60
  class CoS3
    bandwidth remaining percent 30
  class class-default
    fair-queue
```

Teraz jedyne, co należy zrobić, to zastosować mapę polityk do interfejsu, korzystając z polecenia `service-policy`:

```
interface GigabitEthernet0/0
  description *** Provider MPLS Ethernet 150M ***
  bandwidth 150000
  ip address 10.10.10.1 255.255.255.252
  duplex auto
  speed auto
  media-type sfp
  no negotiation auto
  service-policy output Provider-Outbound
```

Gotowy kod pozwoli nam na kształtowanie ruchu; spójrzmy na wynik polecenia `show policy-map interface` *`nazwa_interfejsu`*:

```
Router# sho policy-map interface g0/0
  GigabitEthernet0/0
Service-policy output: Provider-Outbound
  Class-map: class-default (match-any)
    2989747764 packets, 2763665319444 bytes
    5 minute offered rate 83672000 bps, drop rate 0 bps
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Match: any
Traffic Shaping
  Target/Average          Byte      Sustain     Excess    Interval Increment
      Rate                Limit     bits/int    bits/int    (ms)    (bytes)
  150000000/150000000   937500     3750000      3750000     25      468750
  Adapt   Queue Packets      Bytes      Packets  Bytes   Shaping
  Active  Depth                         Delayed  Delayed  Active
     -     0    2989747911   2001535028  217      144908    no
```

Wynik zastosowania polecenia przedstawia statystyki dla parametrów kształtowania ruchu. Widzimy, że średnia z pięciu minut wynosi 85 Mb/s, przy czym żaden bit nie został odrzucony. W większości miejsc, gdzie pracowałem, byłoby to akceptowalne. Zauważ, że w wyniku znajduje się zapis `Shaping Active: no`. Nie przejmuj się tym. Status ten odczytywany jest w milisekundzie naciśnięcia przez Ciebie klawisza *Enter*. Jeśli wykonasz komendę ponownie, prawdopodobnie odczyt wskaże status o wartości `yes`. Pole `Bytes Delayed` świadczy o tym, że mechanizm kształtowania ruchu działa. Wskazuje ono bajty, które dzięki kształtowaniu ruchu nie przepadły w efekcie przepełnienia bufora.

Pozostała część wyniku użycia polecenia przedstawia efekt działania mapy polityk CoS zastosowanej po uruchomieniu mechanizmu kształtowania ruchu:

```
Service-policy : Police-CoS
    Class-map: CoS1 (match-any)
      292144914 packets, 21611629756 bytes
      5 minute offered rate 1189000 bps, drop rate 0 bps
      Match: dscp ef (46)
        292144912 packets, 21611629608 bytes
        5 minute rate 1189000 bps
      QoS Set
        dscp ef
          Packets marked 292144916
      Queueing
          Strict Priority
          Output Queue: Conversation 264
          Bandwidth 60 (%)
          Bandwidth 90000 (kbps) Burst 2250000 (Bytes)
         (pkts matched/bytes matched) 48/3552
         (total drops/bytes drops) 0/0
    Class-map: CoS2 (match-any)
      17228250 packets, 6205496922 bytes
      5 minute offered rate 264000 bps, drop rate 0 bps
      Match: dscp af31 (26)
        17228250 packets, 6205496922 bytes
        5 minute rate 264000 bps
      Match: dscp af32 (28)
                      0 packets, 0 bytes
        5 minute rate 0 bps
      Queueing
        Output Queue: Conversation 265
        Bandwidth remaining 60 (%)Max Threshold 64 (packets)
        (pkts matched/bytes matched) 3/214
    (depth/total drops/no-buffer drops) 0/0/0
    Class-map: CoS3 (match-any)
      2189761162 packets, 2514045625716 bytes
      5 minute offered rate 71825000 bps, drop rate 0 bps
      Match: dscp af21 (18)
       44743755 packets, 28337090377 bytes
       5 minute rate 835000 bps
      Match: dscp af22 (20)
        2145017432 packets, 2485708545019 bytes
```

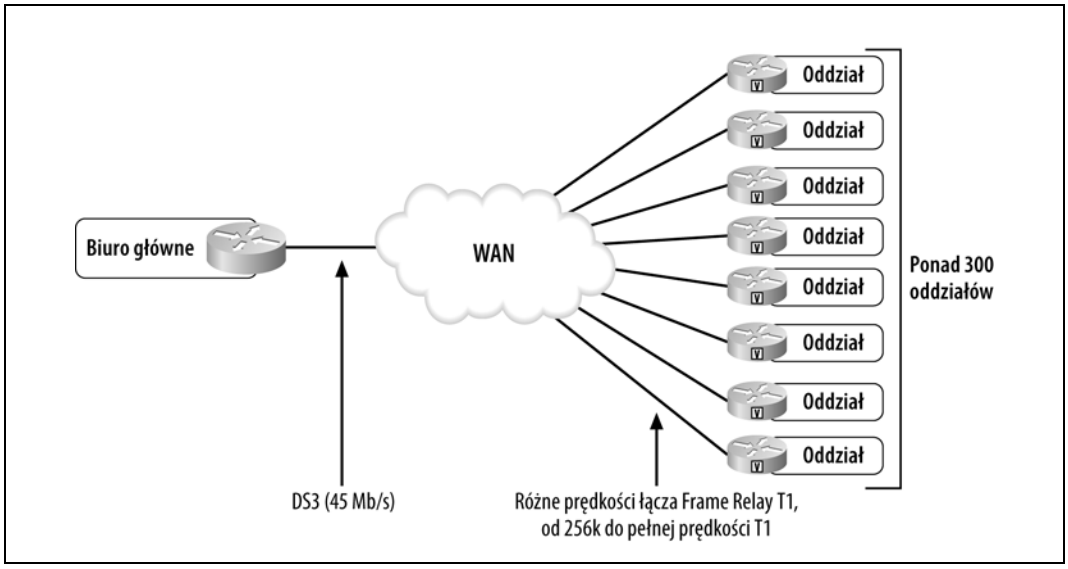

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
    5 minute rate 70984000 bps
  Queueing
    Output Queue: Conversation 266
    Bandwidth remaining 30 (%)Max Threshold 64 (packets)
    (pkts matched/bytes matched) 121/120844
(depth/total drops/no-buffer drops) 0/12/0
Class-map: class-default (match-any)
  490613498 packets, 221802649874 bytes
  5 minute offered rate 10374000 bps, drop rate 0 bps
  Match: any
  Queueing
    Flow Based Fair Queueing
    Maximum Number of Hashed Queues 256
(total queued/total drops/no-buffer drops) 0/0/0
```

Pamiętaj, że jest wiele sposobów na skonfigurowanie QoS i przedstawione rozwiązanie może się u Ciebie nie sprawdzić. Jednak ze swojego doświadczenia muszę powiedzieć, że ten typ aplikacji QoS służył mi bez zarzutów.

Scenariusz 2.: Frame Relay — niedopasowanie prędkości

Przyjrzyjmy się temu, jak rozwiązaliśmy problem z rozdziału 31. związany z niedopasowaniem prędkości we Frame Relay. Sieć, której to dotyczy, przedstawia rysunek 32.2.

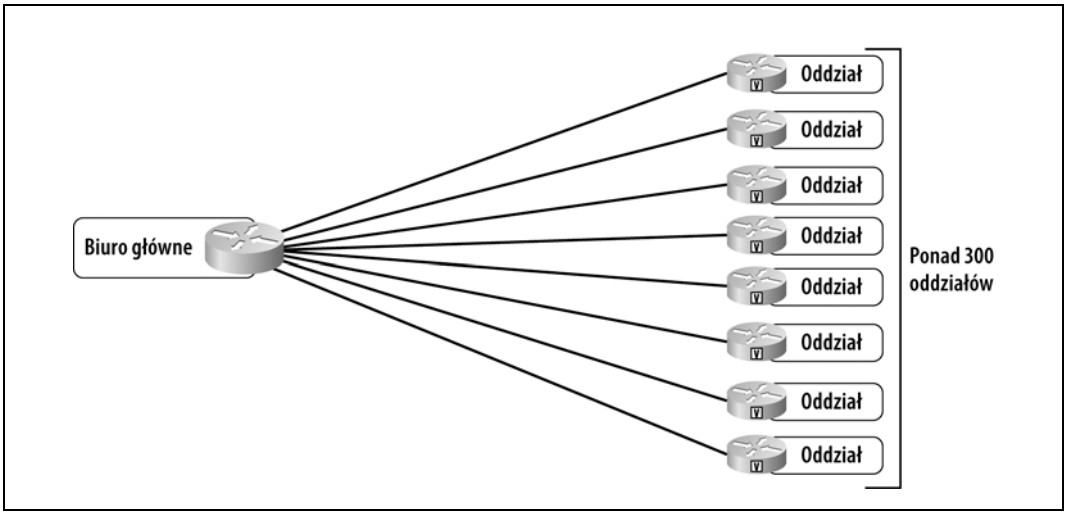


Rysunek 32.2. Frame Relay — niedopasowanie prędkości

Działa ona prawidłowo dzięki temu, że Frame Relay łączy obwody PVC (ang. *private virtual circuits*) pomiędzy każdym z oddziałów a lokalizacją główną. Tak więc w warstwie logicznej sieć wygląda jak na rysunku 32.3.

Jest to istotne, ponieważ zamierzamy użyć podinterfejsów stworzonych dla każdego oddziału oraz głównego biura, aby zastosować na nich mechanizmy kształtowania ruchu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 32.3. Warstwa logiczna Frame Relay z PVC

Ponieważ MPLS, inaczej niż Frame Relay, nie ma możliwości dokonania podziału łącza w warstwie logicznej (w tym przypadku dotyczyć to będzie urządzenia w lokalizacji głównej), z tego rozwiązania nie skorzystamy. Mimo wszystko uważam, że znajomość takiego podejścia jest bardzo cenna, bo nie wiadomo, czy w przyszłości nie nadarzy się okazja, aby je wykorzystać.

Terminologia kształtowania ruchu

Następujące terminy pojawiają się w kontekście zagadnień związanych z mechanizmami kształtowania ruchu. Kiedy pierwszy raz się z nimi spotkasz, mogą być dość niejasne, jednak dla nas okażą się raczej proste:

CIR (ang. *commited information rate*)
: Gwarantowana przepływność wyrażana w bitach na sekundę.

B_c
: Maksymalny rozmiar paczki (ang. *burst size*) wyrażany w bitach (nie bitach na sekundę!); Zazwyczaj $B_c = T_c \cdot CIR$ (zakładając 2× ilość danych transmitowanych).

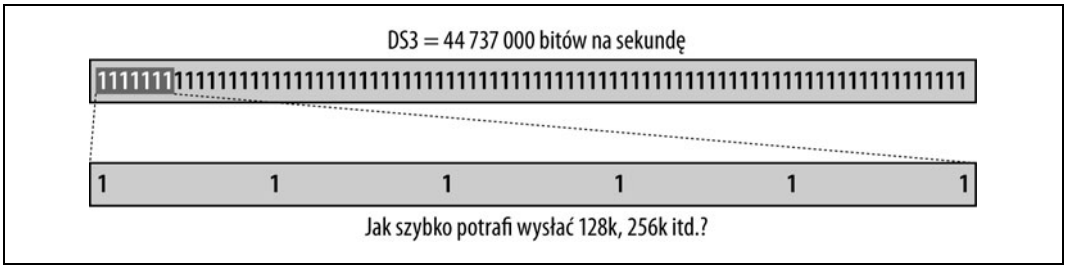
T_c
: Ilość czasu, jaką zajmie wysłanie B_c;
Zazwyczaj $T_c = B_c$ / CIR (zakładając 2× ilość danych transmitowanych).

B_e
: Nadmiar transmitowanej ilości danych wyrażany w bitach (nie w bitach na sekundę!) (zazwyczaj niewymagany).

Aby skonfigurować kształtowanie ruchu, powinniśmy spowolnić wysyłane do każdego oddziału bity (rysunek 32.4). Aby to zrobić, musimy wyliczyć, bazując na prędkości połączeń w każdym oddziale, jak szybko DS3 wysyła bity. Trzymajcie się, dzieci, bo teraz troszkę matematyki.

DS3 wysyła dane z prędkością 44 736 000 bitów na sekundę. Zatem 128 bitów (128 kb/s) zostanie przesłane przez DS3 w 0,0029 sekundy albo 2,9 milisekundy (128 000/44 736 000 = 0,0028612).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 32.4. Prędkość DS3 vs 128k

Odnieśmy to do przykładowego łącza. Dla łącza 256 kb/s, uwzględniając dwukrotną ilość transmitowanych danych (512 kb/s), wartości dla kształtowania ruchu, przy założeniu, że łącze czołowe DS3 znajduje się w naszej sieci, wynosić będą:

CIR: 256 000

B_c: 1465 (256000·0,0057)

T_c: 0,0057

B_e: Niewymagane

Wyliczyłem powyższe wartości dla wszystkich prędkości T1 i mniejszych oraz dla DS3 przyjmującego typowe ilości transmitowanych danych (ang. *burst*). Wartości te zawiera tabela 32.1.

Tabela 32.1. Wartości kształtowania ruchu dla typowych prędkości T1

CIR	Ilość danych	Port	Tc	Bc
128k	128k (2×)	256k	0,0029 s [2,9 ms]	366
256k	256k (2×)	512k	0,0057 s [5,7 ms]	1465
384k	384k (2×)	768k	0,0086 s [8,6 ms]	3296
512k	512k (2×)	1024k	0,0114 s [11,4 ms]	5860
640k	640k (2×)	1280k	0,0143 s [14,3 ms]	9156
768k	768k (2×)	1536k	0,0172 s [17,2 ms]	13185
896k	640k	1536k	0,0200 s [20,0 ms]	9156
1024k	512k	1536k	0,0229 s [22,9 ms]	5860
1152k	384k	1536k	0,0258 s [25,8 ms]	3296
1280k	256k	1536k	0,0286 s [28,6 ms]	1465
1408k	128k	1536k	0,0315 s [31,5 ms]	366
1536k	0k	1536k	0,0343 s [34,3 ms]	0

Zrozumienie powyższej tabeli jest ważne, choć dla naszego przykładu nie będzie ona potrzebna. Na szczęście nasze wymagania są dość proste. Router zakłada 2× ilość transmitowanych danych. Jest to już ustalone, wobec tego konfiguracja jest zbędna. Oto mapa polityk dla czterech różnych prędkości łącza, wszędzie 2× ilość danych (128 kb/s, 256 kb/s, 384 kb/s i 512 kb/s). W skrócie, tworzymy mapę polityk dla każdej prędkości, którą napotkamy, a następnie przekazujemy do `default-class`, jaką prędkość należy ustawić na każdym podinterfejsie, za pomocą komendy `shape peak` *`bps`*:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
policy-map Frame-128
    description [--[ Dla 128k CIR - ilość danych 256k ]--]
  class class-default
    shape peak 128000

policy-map Frame-256
    description [--[ Dla 256k CIR - ilość danych 512k ]--]
  class class-default
    shape peak 256000

policy-map Frame-384
    description [--[ Dla 384k CIR – ilość danych 768k ]--]
  class class-default
    shape peak 384000

policy-map Frame-512
    description [--[ Dla 512k CIR – ilość danych 1024k  ]--]
  class class-default
    shape peak 512000
```

Jak pisałem wcześniej, aby powyższa konfiguracja zadziałała, musisz posiadać podinterfejs dla każdego obwodu PVC w każdym z oddziałów. Oznacza to mnóstwo dodatkowej pracy podczas konfiguracji routera, no, ale cóż, takie życie.

W przykładzie z końcówką ethernet skorzystałem z komendy `shape average`, lecz tutaj używam komendy `shape peak`. Powodem jest to, że aktualna konfiguracja wykonywana jest dla Frame Relay, który pozwala na określenie prędkości transmisji (tzw. *bursting speed*). W tym przypadku należy użyć komendy `peak`, w innym `average`.

Aby zastosować te konfiguracje, używamy komendy `service-policy` na podinterfejsach. Oto przykład konfiguracji setnego podinterfejsu:

```
interface Serial0/0.100 point-to-point
  description [< Oddział 100 >]
  bandwidth 256
  ip address 10.10.20.1 255.255.255.252
  frame-relay interface-dlci 100
  service-policy output Frame-128
end
```

Uważni czytelnicy zapewne zauważyli, że pasmo wynosi 256 kb/s, z kolei zastosowana mapa reguł na interfejsie ma 128 kb/s. Wynika to z cech charakterystycznych transmisji Frame Relay. Na potrzeby wykresów i statystyk pasmo musi być równe maksymalnemu rozmiarowi paczki. Kształtowanie ruchu konfigurowane jest jednak na podstawie CIR, a CIR dla tego łącza wynosi 128 kb/s.

Komenda `show policy-map interface` *`nazwa_interfejsu`* umożliwia Ci wyświetlenie statusu dla kształtowania ruchu:

```
Main-Office-Router#sho policy-map interface s0/0.100
  Serial0/0.100
    Service-policy output: Frame-128
      Class-map: class-default (match-any)
        1147178 packets, 583516684 bytes
        5 minute offered rate 23000 bps, drop rate 0 bps
        Match: any
        Traffic Shaping
          Target/Average      Byte    Sustain     Excess   Interval Increment
              Rate            Limit   bits/int    bits/int  (ms)     (bytes)
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
          256000/128000       1984    7936         7936      62        1984
Adapt   Queue   Packets   Bytes   Packets  Bytes  Shaping
Active  Depth                     Delayed  Delayed Active
-         0     1146896 583184180  411456 434751961  no
```

Podobnie jak wcześniej nie należy przejmować się pozycją `Shaping Active = no`. Wskazuje ona jedynie status aktywności mechanizmów kształtowania ruchu w momencie naciśnięcia klawisza *Enter*. Zmienia stale swoją wartość z `on` na `off`. Interesujące jest natomiast to, że pakiety zostały opóźnione, a mimo tego wskaźnik pakietów odrzuconych wynosi 0 b/s. Na podstawie tego, co zwróciła komenda w polu `Target/Average Rate`, można zaobserwować, że mimo skonfigurowania CIR na 128 kb/s przyjęty został dwukrotnie większy rozmiar ilości danych.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

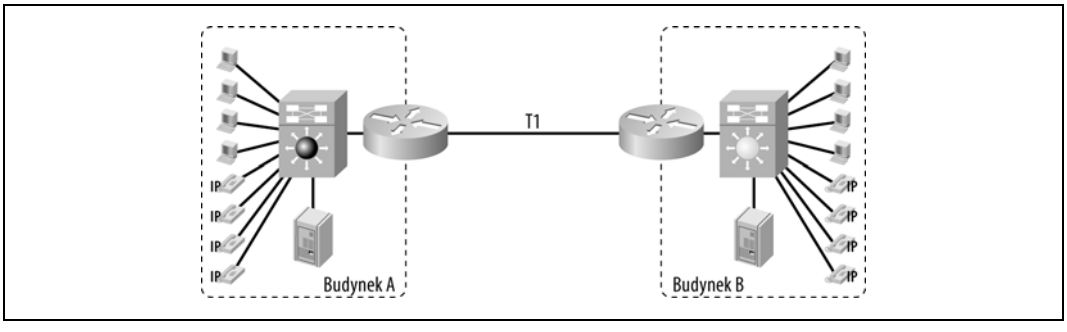
ROZDZIAŁ 33.

Sieci przeciążone

Sieć przeciążona (ang. *congested network*) to taka, w której jest zbyt dużo danych, a dostępne pasmo nie jest wystarczające, aby sprawnie obsłużyć ruch. W przypadku sieci przeciążonych mogą pomóc mechanizmy QoS, jednak nie zlikwidują one źródła problemu. Jedynym sposobem, aby pozbyć się przeciążenia w sieci, jest zwiększenie dostępnego pasma lub też ograniczenie ruchu. Mimo to przyjrzyjmy się temu, jak QoS może ulżyć naszym problemom.

Rozpoznawanie sieci przeciążonych

Skąd wiesz, że Twoja sieć jest przeciążona? Spójrzmy ponownie na naszą ulubioną dwubudynkową firmę:



Rysunek 33.1. Typowa sieć łącząca dwa budynki

Użytkownicy obu siedzib skarżą się, że dostęp do budynku obok jest powolny. Spójrzmy na interfejs jednego z routerów podłączonych do obwodu T1. Oto rezultat użycia komendy `show interface` dla interfejsu szeregowego routera w budynku B:

```
Bldg-B-Rtr#sho int s0/0
Serial0/0 is up, line protocol is up
  Hardware is PowerQUICC Serial
  Description: <[ Łącze WAN T1 ]> ❶
  Internet address is 10.10.10.2/30
  MTU 1500 bytes, BW 1544 Kbit, DLY 20000 usec, ❷
    reliability 255/255, txload 42/255, rxload 249/255 ❸
  Encapsulation PPP, loopback not set
  Keepalive set (10 sec)
  LCP Open
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Open: IPCP, CDPCP
Last input 00:00:00, output 00:00:00, output hang never
Last clearing of "show interface" counters 3w4d
Queueing strategy: fifo
Output queue 0/40, 548941 drops; input queue 0/75, 3717 drops
5 minute input rate 1509000 bits/sec, 258 packets/sec ❹
5 minute output rate 259000 bits/sec, 241 packets/sec ❺
  287554125 packets input, 3659652949 bytes, 0 no buffer
  Received 0 broadcasts, 0 runts, 0 giants, 0 throttles
  819 input errors, 559 CRC, 227 frame, 0 overrun, 0 ignored, 33 abort
  282883104 packets output, 3613739796 bytes, 0 underruns
  0 output errors, 0 collisions, 1 interface resets
  0 output buffer failures, 0 output buffers swapped out
  0 carrier transitions
  DCD=up DSR=up DTR=up RTS=up CTS=up
```

Patrząc na wynik użycia polecenia, powinieneś natychmiast zauważyć pewne rzeczy. Po pierwsze, jest to T1, po drugie, przepływność wynosi 1544 kb/s (zobacz ❶ i ❷).

Pamiętaj, że pasmo, jakie przedstawia rezultat wykonania komendy `show interface`, niekoniecznie prezentuje prawdziwy rozmiar pasma łącza. Wartość ta jest domyślnie prawidłowa, jednak może zostać zmieniona za pomocą komendy `bandwith`. Jest tak w związku z tym, że EIGRP oraz inne protokoły używają jej do swoich kalkulacji.

Spoglądając nieco niżej (❹), możesz zauważyć, że wskaźnik z 5 minut wskazuje 1 509 000 b/s przy 258 pakietach na sekundę (p/s). W rezultacie otrzymujemy 1509 kb/s na łączu 1544 kb/s. Mógłbyś powiedzieć: „Łącze ma jeszcze 33 kb/s", pamiętaj jednak, że nasz wynik jest **średnią** z 5 minut.

Może lepszym sposobem na ukazanie charakterystyki wykorzystania łącza jest zapis ❸, mianowicie:

```
reliability 255/255, txload 42/255, rxload 249/255
```

Patrząc na te liczby, widzisz, że obciążenie transmisji to jedynie 42 z 255 (`txload`), ale już obciążenie przy odbiorze (`rxload`) to 254 z 255. Prosta operacja matematyczna (249/255·100%) pozwoli Ci wyliczyć, że łącze jest utylizowane na poziomie 97,65%.

Zastanawiałeś się kiedyś, dlaczego wartości `txload` oraz `rxload` wyrażane są w odniesieniu do 255, a nie do 100? Liczba 100 w notacji szesnastkowej to 64. Jeśli już musimy zapisać dwie cyfry szesnastkowe w pamięci routera, dlaczego by nie wykorzystać ich pełnych możliwości. Największa liczba szesnastkowa składająca się z dwóch cyfr to `FF`, czyli 255. Wykorzystując 255 zamiast 100, otrzymujemy ponad dwa razy większą granularność pomiaru.

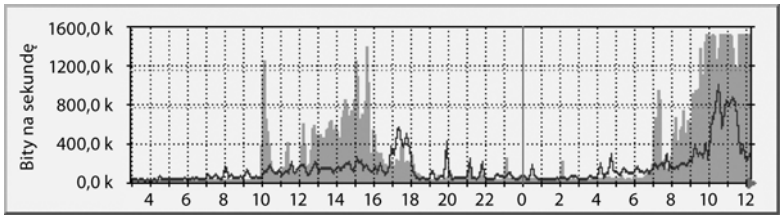
Łącze jest wysycone tylko w jednym kierunku, `rxload` wynosi 249/255, a `txload` tylko 42/255. Zapisy ❹ i ❺ pokazują owo wysycenie w kontekście wykorzystania pasma. Otrzymujemy 1509 kb/s, a transmitujemy tylko 259 kb/s. Pamiętaj, że T1 korzysta z pełnego dupleksu. To, co otrzymujemy, z reguły nie ma wpływu na to, co wysyłamy.

T1, podobnie jak większość łączy telekomunikacyjnych, działa w trybie pełnego dupleksu. Prędkość przesyłu wynosi 1,54 Mb/s, a transmisja z taką prędkością może odbywać się w obu kierunkach jednocześnie.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Prostszym nawet sposobem na sprawdzenie, czy łącze jest wysycone, jest zainstalowanie dedykowanego narzędzia, jak na przykład *Multi Router Traffic Grapher* (MRTG), które dostarczy Ci graficzną reprezentację ruchu na Twoim łączu.

Rysunek 33.2 przedstawia wykres MRTG dla omawianego łącza. Oś pozioma wskazuje godziny. Mały trójkąt na prawym brzegu wykresu pokazuje aktualny czas (w przykładzie jest to godzina 12:12). Linia pionowa umieszczona na wartości 0 wskazuje północ. Poczynając od około 9:30 – 10:00, aż do końca obserwowanego przedziału łącze jest całkowicie wysycone w jednym kierunku. Pokrywa się to znakomicie z tym, co widzieliśmy na routerze.



Rysunek 33.2. Wykres przedstawiający utylizację łącza T1

Zatem łącze jest wysycone. Co z tego? Nasze pieniądze zostały dobrze wykorzystane. Dlaczego więc użytkownicy skarżą się na spowolnienie sieci? Spójrzmy nieco głębiej. Pierwsza rzecz, której powinieneś szukać, to problemy w warstwie fizycznej. Poszukaj błędów w rezultacie wykonania komendy `show interface`. Pogrubione linie blisko końca wyniku, jaki zwraca komenda, wskazują pewne błędy:

```
819 input errors, 559 CRC, 227 frame, 0 overrun, 0 ignored, 33 abort
  0 output errors, 0 collisions, 1 interface resets
```

Powyższe dwie linie wskazują na błędy wejścia (ang. *input errors*), błędy CRC, odrzucane pakiety (ang. *abort errors*), błędy ramek (ang. *framing errors*), restarty interfejsu. Nie jest więc za dobrze, ale nie bądźmy jeszcze tym tak podekscytowani. Pierwsze, co należy zrobić w przypadku wystąpienia takich błędów, szczególnie ich niewielkiej liczby, to zaczekać kilka minut i wykonać komendę `show interface` ponownie. A oto dość ciekawy rezultat użycia tej komendy 10 minut później:

```
 819 input errors, 559 CRC, 227 frame, 0 overrun, 0 ignored, 33 abort
  0 output errors, 0 collisions, 1 interface resets
```

Oba wyniki są identyczne, żaden z liczników się nie zwiększył. Według mnie wskazuje to na problem w przeszłości, który zmusił kogoś do wyciągnięcia kabla T1 i zrestartowania interfejsu. Mogło być też tak, że dostawca telekomunikacyjny wygenerował przerwę na linii T1. Odejdźmy jednak od rozwiązywania tej zagadki, ponieważ nie jest to nam do niczego potrzebne. Skoro już ustaliliśmy, że liczniki się nie zmieniają, wyczyśćmy je tak, aby kiedy po raz kolejny zalogujemy się na router, nie uganiać się za duchami. Choć nadal potrzebujemy tych danych, nie chcemy robić tego teraz.

Kolejną rzeczą, którą powinniśmy sprawdzić, jest stan kolejek wejścia i wyjścia. Poprzedni wynik komendy `show interface` zawiera następujący zapis:

```
Queueing strategy: fifo
  Output queue 0/40, 548941 drops; input queue 0/75, 3717 drops
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Pierwszą rzeczą, jaką powinieneś zauważyć, jest to, że **strategia kolejkowania** (ang. *strategy queueing*) to FIFO (*first-in, first-out*). Łącze o przepustowości 2 Mb/s oraz niższej powinno korzystać z WFQ, zatem jest to już pewien punkt zaczepienia. Szukając dalej, zauważysz, że w kolejce wyjścia wystąpiło 548 941 odrzuceń pakietów, a w kolejce wejścia 3717. Oznacza to, że od momentu, kiedy wyzerowano liczniki (teraz wiesz, dlaczego ich wcześniej nie wyczyściliśmy), 548 941 transmitowanych pakietów zostało odrzuconych.

Linia 13 wskazuje, ile czasu minęło od momentu, kiedy liczniki zostały wyzerowane:

```
Last clearing of "show interface" counters 3w4d
```

Linia 22 wskazuje, ile pakietów zostało wysłanych:

```
282883104 packets output, 3613739796 bytes, 0 underruns
```

W ciągu 3 tygodni i 4 dni przetransmitowaliśmy 282 883 104 pakiety, a 548 941 pakietów zostało odrzuconych. Proste obliczenia pokażą, że liczba pakietów odrzuconych to jedynie 1/10 z jednego procenta wszystkich pakietów. Zatem, dlaczego mamy problem?

No dobrze, poprowadziłem Cię fałszywym tropem, ale zrobiłem to po to, abyś nabrał nieco doświadczenia. Zatrzymajmy się przez chwilę i pomyślmy. Liczby, które otrzymaliśmy, nie wyglądają przecież źle. Rzeczywiście sieć jest wysycona, jednak pakietów nie gubimy aż tak wielu, prawda? Skąd więc biorą się te narzekania użytkowników?

Odpowiedź jest zaskakująco prosta: szukamy po złej stronie łącza. Spójrzmy na zapisy ❹ i ❺:

```
5 minute input rate 1509000 bits/sec, 258 packets/sec
5 minute output rate 259000 bits/sec, 241 packets/sec
```

Nie ruch wyjściowy, a wejściowy sięga maksimum! Kiedy gubimy pakiety w związku z przeciążeniem, gubimy je na wyjściu. Pamiętasz przykład z lejkiem z rozdziału 31.? Szukamy na ujściu lejka. Jedyne pakiety, jakie się tam przedostaną, to te, które przejdą przez lejek! Ze swojej perspektywy nie zobaczymy wszystkich tych pakietów, które zostały zgubione.

Spróbujmy zatem poszukać po drugiej stronie łącza, korzystając z tej samej komendy `show interface`:

```
Bldg-A-Rtr#sho int s0/0
Serial0/0 is up, line protocol is up
  Hardware is PowerQUICC Serial
  Description: [-- Łącze WAN T1 --]
  Internet address is 10.10.10.1/30
  MTU 1500 bytes, BW 1544 Kbit, DLY 20000 usec,
    reliability 255/255, txload 250/255, rxload 47/255
  Encapsulation PPP, loopback not set
  Keepalive set (10 sec)
  LCP Open
  Open: IPCP, CDPCP
  Last input 00:00:00, output 00:00:00, output hang never
  Last clearing of "show interface" counters 3w4d
Input queue: 0/75/0 (size/max/drops); Total output drops: 152195125
  Queueing strategy: weighted fair
  Output queue: 63/1000/64/152195113 (size/max total/threshold/drops)
      Conversations 7/223/256 (active/max active/max total)
    Reserved Conversations 0/0 (allocated/max allocated)
  5 minute input rate 261000 bits/sec, 238 packets/sec
  5 minute output rate 1511000 bits/sec, 249 packets/sec
    282883104 packets input, 3613739796 bytes, 0 no buffer
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
     Received 0 broadcasts, 0 runts, 0 giants, 0 throttles
     872 input errors, 472 CRC, 357 frame, 0 overrun, 0 ignored, 12 abort
     1326931662 packets output, 2869922208 bytes, 0 underruns
     0 output errors, 0 collisions, 1 interface resets
     0 output buffer failures, 0 output buffers swapped out
     0 carrier transitions
     DCD=up DSR=up DTR=up RTS=up CTS=up
```

Krótki rzut oka na statystyki obciążenia pokazuje, że tym razem obserwujemy właściwą stronę łącza:

```
reliability 255/255, txload 250/255, rxload 47/255
```

Błędy wyglądają podobnie, nie są naszym zmartwieniem:

```
Received 0 broadcasts, 0 runts, 0 giants, 0 throttles
872 input errors, 472 CRC, 357 frame, 0 overrun, 0 ignored, 12 abort
```

Spójrzmy na kolejki:

```
Input queue: 0/75/0 (size/max/drops); Total output drops: 152195125
  Queueing strategy: weighted fair
  Output queue: 63/1000/64/152195113 (size/max total/threshold/drops)
    Conversations 7/223/256 (active/max active/max total)
    Reserved Conversations 0/0 (allocated/max allocated)
```

Pierwszą rzeczą, która zwraca moją uwagę, jest to, że po tej stronie łącza aktywny jest mechanizm WFQ (tak jak powinno być). Pamiętasz zapewne, że po przeciwnej stronie działało FIFO. Zgodności być nie musi, ale prawdopodobnie działanie mechanizmu FIFO po drugiej stronie nie ma swojego uzasadnienia.

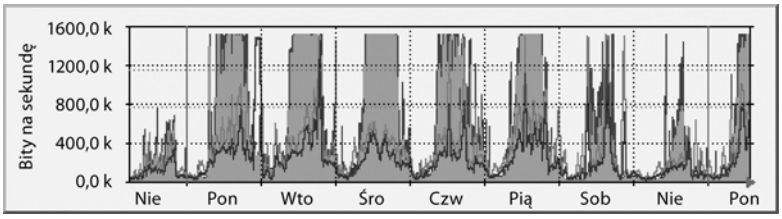
No, proszę! 152 195 113 odrzuconych pakietów na wyjściu kolejki! Jeśli podzielimy to przez ich całkowitą liczbę (152 195 113/1 326 931 662·100%), otrzymamy informację, że 11,47% wszystkich pakietów zostało zgubione. Powiedziałbym, że znaleźliśmy problem. Łącze jest tak wysycone, że ponad 1/10 pakietów zostaje odrzucana. Trudno się dziwić skarżącym się użytkownikom!

11,47% może nie wydawać się dużą wartością, ale pamiętaj, że np. VoIP nie może funkcjonować, kiedy pakiety są gubione. W przypadku VoIP nawet 1% może być problemem. Dla innych protokołów, które potrafią odzyskiwać gubione pakiety, jest to dość niski procent. Zatem, jak powinny wyglądać Twoje statystyki gubionych pakietów? W świecie idealnym powinny wynosić 0. Bardziej praktycznie jednak: nie powinny przyrastać, gdy je obserwujesz, ich wartości powinny być jak najbardziej zbliżone do 0% wszystkich wysłanych pakietów. Jeżeli gubisz nawet 1% pakietów transmitowanych łączem szeregowym, jest ono przeciążone.

Kiedy już wiemy, gdzie leży problem, musimy przez jakiś czas poobserwować łącze. Chcemy się upewnić, że te wszystkie odrzucone pakiety to nie kwestia ostatnich 10 minut (choć jest to mało prawdopodobne, biorąc pod uwagę wysoką liczbę zgubionych pakietów). Bezpiecznie jest uzyskać pewność. Jeśli skonfigurowaliśmy MRTG do monitorowania wykorzystania pasma, historyczne dane tam zawarte powinny nam pomóc, a ładne wykresy wyjaśnić zarządzającym wiele rzeczy dużo łatwiej.

Rysunek 33.3 pokazuje, że łącze wysycone jest w dość dużym stopniu w ciągu całego dnia. Problem pojawia się w przybliżeniu wtedy, kiedy pracownicy przychodzą do pracy, a znika wtedy, gdy wychodzą do domów. Zwróć uwagę na piki na wykresie i utrzymywanie się obciążenia na poziomie 1600 kb/s codziennie i w ciągu całego dnia. To łącze jest przeciążone.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 33.3. Tygodniowy wykres MRTG dla naszego łącza

Rozwiązanie problemu

Co zatem zrobimy? Czy QoS może nas ocalić? No więc… tak i nie. QoS w zasadzie działa na łączu (WFQ) i ma konkretne założenia, zgodnie z którymi ruch powinien być priorytetyzowany. WFQ również działa tylko w jednym kierunku (na szczęście w kierunku, który nas interesuje i w którym przesyłana jest większość danych).

Mówiliśmy wcześniej o tym, czego nie może QoS. Jedną z istotnych rzeczy, jakiej nie potrafi, jest dostarczenie większej przepływności pasmu. Łącze potrzebuje zdecydowanie więcej pasma, ponieważ więcej niż 10% pakietów przesyłanych z budynku A do B jest gubione. W tym przypadku jedynym rozwiązaniem jest zwiększenie przepływności łącza lub zaprzestanie przesyłania tak dużej ilości danych.

To, co **może** QoS, to określić, które pakiety nie powinny być gubione, a które mogą być odrzucane (zobacz rozdział 31. i 32. zawierające informacje o projektowaniu schematów QoS).

Wyczyśćmy teraz liczniki, aby przy kolejnym odczycie otrzymać świeże dane:

```
Bldg-A-Rtr#clear counters s0/0
Clear "show interface" counters on this interface [confirm] // wyczyść liczniki polecenia „show interface"
                                                            // na tym interfejsie [potwierdź]
```

Z drugiej strony również:

```
Bldg-B-Rtr#clear counters s0/0
Clear "show interface" counters on this interface [confirm] // wyczyść liczniki polecenia „show interface"
                                                            // na tym interfejsie [potwierdź]
```

Jeśli chcesz doprowadzić do zakupu szybszego łącza, musisz zgromadzić dane, które Ci pomogą. Jeżeli przyjdziesz z problemem do zarządu, posiadanie informacji podobnych do tych, które przedstawiłem w tym rozdziale, pomoże Ci uzasadnić wydatek na nowe łącze. Monitoring z wykorzystaniem SNMP jest jednym z najbardziej użytecznych narzędzi, które możesz zastosować na potrzeby swojej sieci. Nie potrzebujesz fantazyjnych narzędzi w rodzaju CiscoWorks albo OpenView, choć jeśli Cię na to stać, są one godne polecenia. Darmowe narzędzia, jak na przykład MRTG, są bardzo potężne i użyteczne, zwłaszcza w przypadku analizy dużych łączy.

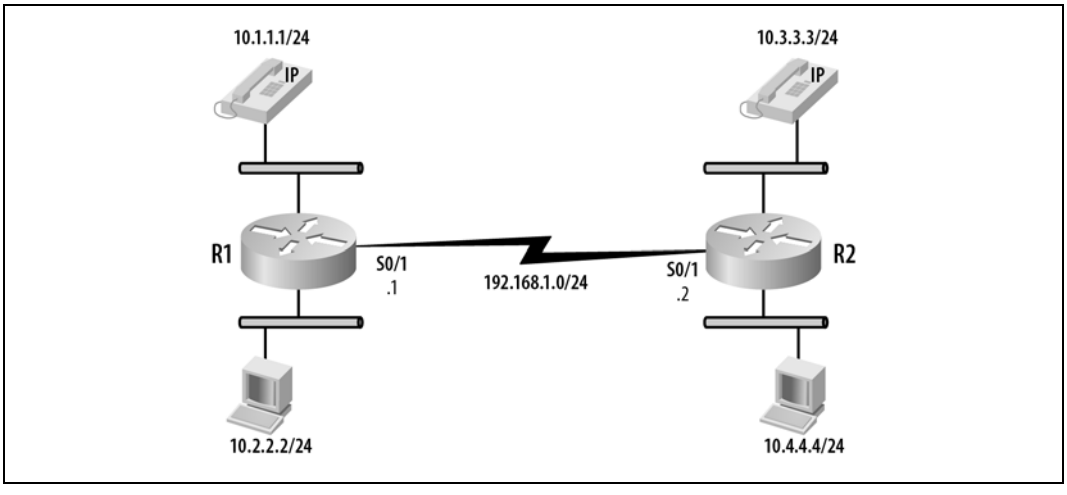
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 34.

Sieci konwergentne

W tym rozdziale zobaczysz sieć konwergentną w akcji. Mimo że sieć będzie bardzo prosta, przedstawione zasady odnosić się będą do sieci i łączy prawie każdych rozmiarów.

Rysunek 34.1 prezentuje przykładową sieć, którą wykorzystamy na potrzeby tego rozdziału. R1 i R2 mają po dwie karty sieciowe, do jednej podłączony jest telefon IP, a do drugiej komputer osobisty. Routery połączone są ze sobą łączem T1 terminowanym w obu na `S0/1`.



Rysunek 34.1. Prosta sieć konwergentna

Konfiguracja

Interfejs, na którym się skupimy, to `S0/1` na urządzeniu R1. Oto konfiguracja:

```
interface Serial0/1
  ip address 192.168.1.1 255.255.255.0
  service-policy output WAN
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Wyrażenie `service-policy` przypisuje mapę polityk `WAN` do interfejsu. Oto konfiguracja mapy polityk:

```
policy-map WAN
  class Voice-RTP
    priority 128
  class Voice-Control
    bandwidth percent 5
  class HTTP
    bandwidth percent 10
  class class-default
    fair-queue
```

Mapa polityk odnosi się do czterech klas: `Voice-RTP`, `Voice-Control`, `HTTP` oraz specjalnej klasy `class-default`. Pamiętaj, że `class-default` jest klasą specjalną używaną na potrzeby konfiguracji klasy domyślnej. Klasa domyślna to taka, do której trafiają pakiety niepasujące do żadnej innej klasy. Kolejki zostały zaprojektowane jak poniżej:

`Voice-RTP`
: Pakiety z tej klasy trafią do kolejki priorytetowej. Kolejka otrzyma pasmo 128 kb/s. Pasmo stanowi w przybliżeniu tyle, ile dwa połączenia VoIP 64 kb/s kodowane G.711. Pakiety z *IP precedence* równym 5 zostaną umieszczone w tej kolejce. Zdefiniowane jest to w mapie klasy `Voice-RTP`.

`Voice Control`
: Pakiety kontrolne audio nie powinny szczególnie zajmować pasma. Na ich potrzeby zagwarantowałem 5% dostępnej przepustowości. Pakiety oznaczone *IP precedence* równym 3 zostaną przyporządkowane do tej kolejki. Zdefiniowane jest to w mapie klasy `Voice-Control`.

`HTTP`
: Tej klasie zagwarantowałem 10% całego pasma. Dzięki mapie klas `HTTP` oraz liście dostępu `101` każdy ruch, którego przeznaczeniem będzie port 80, będzie pasował do tej klasy.

`class-default`
: Klasa, do której trafią pakiety niepasujące do żadnej innej klasy. Jedynym sposobem na definiowanie tej klasy jest polecenie `fair-queued`. Będzie ona traktowana jak pozostałe na łączu T1, jednak dopiero wtedy, kiedy inne kolejki zostaną przeprocesowane.

Konfiguracja dla map klas jest przedstawiona poniżej. Klasa `class-default` jest klasą domyślną, która w IOS nie wymaga konfiguracji.

```
class-map match-any HTTP
  match access-group 101
class-map match-any Voice-RTP
  match ip precedence 5
class-map match-all Voice-Control
  match ip precedence 3
```

Klasa `HTTP` w celu dopasowania pakietów odnosi się do listy dostępu `101`, która wykrywa obecność pakietów HTTP:

```
access-list 101 permit tcp any any eq www
access-list 101 permit tcp any any eq 443
```

Mógłbym użyć zapisu `match protocol http` w mapie klasy, zamiast wywoływać listę dostępu, jednak dopasowany zostanie wtedy jedynie port 80. Zastosowanie listy dostępu pozwala na elastyczne dodanie HTTPS (port 443) oraz HTTP na niestandardowych portach (8080).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Monitorowanie QoS

Monitorowanie kolejek odbywa się głównie z zastosowaniem kilku komend. Rezultat wykonania tych komend zawiera jednak wiele informacji, dlatego wyjaśnię, czego należy szukać. Na początku wywołajmy komendę `show interface`:

```
R1#sho int s0/1
Serial0/1 is up, line protocol is up
  Hardware is PQUICC with Fractional T1 CSU/DSU
  Internet address is 192.168.1.1/24
  MTU 1500 bytes, BW 1544 Kbit, DLY 20000 usec,
    reliability 255/255, txload 21/255, rxload 22/255
  Encapsulation HDLC, loopback not set
  Keepalive set (10 sec)
  Last input 00:00:00, output 00:00:00, output hang never
  Last clearing of "show interface" counters 5d04h
  Input queue: 2/75/0/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: weighted fair
  Output queue: 0/1000/64/0 (size/max total/threshold/drops)
    Conversations 0/4/256 (active/max active/max total)
    Reserved Conversations 2/2 (allocated/max allocated)
    Available Bandwidth 1030 kilobits/sec
5 minute input rate 134000 bits/sec, 53 packets/sec
5 minute output rate 131000 bits/sec, 53 packets/sec
  22874566 packets input, 2838065420 bytes, 0 no buffer
  Received 52348 broadcasts, 0 runts, 0 giants, 0 throttles
  0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored, 0 abort
  22964396 packets output, 2678320888 bytes, 0 underruns
  0 output errors, 0 collisions, 0 interface resets
  0 output buffer failures, 0 output buffers swapped out
  0 carrier transitions
  DCD=up DSR=up DTR=up RTS=up CTS=up
```

Wynik wykonania komendy wskazuje, że strategia kolejkowania to WFQ (`weighted fair`). Używamy LLQ, więc dlaczego komenda `show interface` wskazuje inaczej? LLQ to w praktyce CBWFQ (ang. *class-based weighted fair queuing*) mające kolejkę priorytetową. Jako że CBWFQ jest typem WFQ, router wskazuje, iż strategią kolejkowania jest `weighted fair`.

Zauważ, że polecenie `show interface` nie pokazuje poszczególnych kolejek, które określiliśmy w mapach reguł. Aby zobaczyć ich statusy, musimy skorzystać z komendy `show policy-map inter ↪face` *`nazwa_interfejsu`*. Komenda ta generuje sporo linii, a im więcej kolejek skonfigurowałeś, tym dłuższy będzie rezultat jej wykonania:

```
R1#sho policy-map interface s0/1
  Serial0/1
    Service-policy output: WAN
      Class-map: Voice-RTP (match-any)
        19211963 packets, 3918472220 bytes
        5 minute offered rate 81000 bps, drop rate 0 bps
        Match: ip precedence 5
          19211963 packets, 3918472220 bytes
          5 minute rate 81000 bps
        Queueing
          Strict Priority
          Output Queue: Conversation 264
          Bandwidth 128 (kbps) Burst 3200 (Bytes)
          (pkts matched/bytes matched) 19211963/3918472220
           (total drops/bytes drops) 0/0
      Class-map: Voice-Control (match-all)
        793462 packets, 161097064 bytes
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: ip precedence 3
      Queueing
        Output Queue: Conversation 265
        Bandwidth 5 (%) Max Threshold 64 (packets)
         (pkts matched/bytes matched) 793462/161097064
         (depth/total drops/no-buffer drops) 0/0/0
    Class-map: HTTP (match-any)
      717928 packets, 42537234 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: access-group 101
        717928 packets, 42537234 bytes
        5 minute rate 0 bps
      Queueing
        Output Queue: Conversation 266
        Bandwidth 10 (%) Max Threshold 64 (packets)
         (pkts matched/bytes matched) 717934/42537620
         (depth/total drops/no-buffer drops) 0/0/0
    Class-map: class-default (match-any)
      2243454 packets, 2851898036 bytes
      5 minute offered rate 52000 bps, drop rate 0 bps
      Match: any
      Queueing
        Flow Based Fair Queueing
        Maximum Number of Hashed Queues 256
         (total queued/total drops/no-buffer drops) 0/0/0
```

Każda kolejka została przedstawiona dość szczegółowo, widać różnicę, jeśli chodzi o to, co komenda raportuje dla każdej z kolejek. Kolejka priorytetowa została użyta w mapie klasy `Voice-RTP`. Oto wynik dla tej kolejki:

```
Class-map: Voice-RTP (match-any)
  19211963 packets, 3918472220 bytes
  5 minute offered rate 81000 bps, drop rate 0 bps
  Match: ip precedence 5
    19211963 packets, 3918472220 bytes
    5 minute rate 81000 bps
  Queueing
    Strict Priority
    Output Queue: Conversation 264
    Bandwidth 128 (kbps) Burst 3200 (Bytes)
    (pkts matched/bytes matched) 19211963/3918472220
    (total drops/bytes drops) 0/0
```

Interesujące fragmenty pogrubiłem. Pierwsza pogrubiona linia wskazuje na zapis `5 minute offe ↪red rate`, a po nim `drop rate`. Parametr `drop rate` powinien mieć zawsze wartość 0, szczególnie w przypadku kolejki priorytetowej. Jeśli pakiety są odrzucane, oznacza to, że kolejka jest za mała. `5 minute offered rate` wskazuje, jaką część pasma zajmuje kolejka. Ta informacja zostanie również przedstawiona dalej jako `5 minute rate`.

W akapicie rozpoczynającym się od `Queueing` widać, czy mapa wykorzystuje kolejkę priorytetową. Poniżej znajduje się informacja o rozmiarze kolejki, a na końcu liczba odrzuconych pakietów i bajtów. Zapis ten powinien zawsze wynosić `0/0`. Jeśli Twój ruch będzie przekraczał zakładany poziom, mogą pojawić się odrzucane pakiety. W przypadku transmisji głosu odrzucanie pakietów może skutkować zniekształceniami oraz spadkiem jakości głosu, a tym samym niezadowoleniem użytkowników.

Informacje zwrócone dla następnej kolejki są nieco inne. Nie jest to kolejka priorytetowa i wynik jest zgodny z tym, co zaprojektowaliśmy:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
Class-map: Voice-Control (match-all)
  793462 packets, 161097064 bytes
  5 minute offered rate 0 bps, drop rate 0 bps
  Match: ip precedence 3
  Queueing
    Output Queue: Conversation 265
    Bandwidth 5 (%) Max Threshold 64 (packets)
    (pkts matched/bytes matched) 793462/161097064
    (depth/total drops/no-buffer drops) 0/0/0
```

W tym przykładzie parametr `5 minute offered rate` ma wartość 0. Pakiety kontrolne audio wysyłane są jedynie w przypadku nawiązywania oraz zrywania połączeń. Wartość ta jest prawidłowa, jeśli nie nawiązano bądź nie zerwano w ciągu ostatnich 5 minut żadnego połączenia.

Poniżej, w wyniku wykonania komendy widzimy, że na potrzeby kolejki alokowano 5% pasma. Ostatnia linia pokazuje po raz kolejny, że żadne pakiety tej kolejki nie zostały odrzucone. `Depth` oznacza liczbę pakietów przebywających aktualnie w kolejce. Na mocno utylizowanym łączu wartość ta może przybierać dość zmienne wartości. Dopóki jednak nie obserwujemy odrzucania pakietów, wszystko jest w porządku.

Przeskakując do ostatniej kolejki, widzimy jeszcze inny rodzaj zapisu:

```
Class-map: class-default (match-any)
  2243454 packets, 2851898036 bytes
  5 minute offered rate 52000 bps, drop rate 0 bps
  Monitoring QoS | 619
  Match: any
  Queueing
    Flow Based Fair Queueing
    Maximum Number of Hashed Queues 256
    (total queued/total drops/no-buffer drops) 0/0/0
```

Pierwsza z pogrubionych linii powinna być szczególnie uważnie monitorowana. Jeśli wszystkie pozostałe kolejki zachowują się poprawnie, ta linia powinna Ci pokazać, ile pasma potrzebne jest kolejce domyślnej. Informuje również o wartości parametru `drop rate`, która powinna wynosić 0. Jeśli jest jakaś kolejka, której pozwolimy gubić pakiety, to właśnie ta. Trafia tutaj cały ruch uznany za niekrytyczny. Jeżeli w związku z przeciążeniem łącza chcesz, aby jakieś pakiety były odrzucane, powinien to być właśnie ruch niekrytyczny, a nie przykładowo transmisja głosu.

Rozwiązywanie problemów w sieciach konwergentnych

Gubienie pakietów w ramach poszczególnych kolejek może być trudne do zdiagnozowania. Jeśli jednak spędzisz trochę czasu, pracując z sieciami konwergentnymi, poczujesz, jak powinny się zachowywać poszczególne kolejki. Chociaż każda sieć jest inna, istnieją pewne typowe symptomy, na które należy zwrócić uwagę.

Błędna konfiguracja kolejki

Jeśli wszystkie kolejki są skonfigurowane poprawnie, żadna z nich nie powinna gubić pakietów, gdy sieć nie jest przeciążona. Jeśli przeciążenie wystąpi, pakiety powinny być gubione w pierwszej kolejności w najmniej ważnej kolejce. Najmniej istotną kolejką jest kolejka domyślna, która zawiera pakiety bez priorytetu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Skierowałem na łącze T1 nadmiernie duży ruch. Kiedy większość kolejek jest zajęta, tylko kolejka domyślna gubi pakiety:

```
R1#sho policy-map interface s0/1
  Serial0/1
    Service-policy output: WAN
      Class-map: Voice-RTP (match-any)
        11091627 packets, 2262021588 bytes
        5 minute offered rate 324000 bps, drop rate 0 bps
        Match: ip precedence 5
          11091627 packets, 2262021588 bytes
          5 minute rate 324000 bps
        Queueing
          Strict Priority
          Output Queue: Conversation 264
          Bandwidth 500 (kbps) Burst 12500 (Bytes)
          (pkts matched/bytes matched) 11091627/2262021588
          (total drops/bytes drops) 0/0
      Class-map: Voice-Control (match-all)
        7312459 packets, 1477128588 bytes
        5 minute offered rate 255000 bps, drop rate 0 bps
        Match: ip precedence 3
        Queueing
          Output Queue: Conversation 265
          Bandwidth 5 (%) Max Threshold 64 (packets)
          (pkts matched/bytes matched) 7312784/1477194238
         (depth/total drops/no-buffer drops) 0/0/0
      Class-map: HTTP (match-any)
        79255 packets, 4690398 bytes
        5 minute offered rate 1000 bps, drop rate 0 bps
        Match: access-group 101
          79255 packets, 4690398 bytes
          5 minute rate 1000 bps
        Queueing
          Output Queue: Conversation 266
          Bandwidth 10 (%) Max Threshold 64 (packets)
          (pkts matched/bytes matched) 78903/4669112
         (depth/total drops/no-buffer drops) 0/0/0
      Class-map: class-default (match-any)
        23750755 packets, 2030244008 bytes
        5 minute offered rate 887000 bps, drop rate 4000 bps
        Match: any
        Queueing
          Flow Based Fair Queueing
          Maximum Number of Hashed Queues 256
         (total queued/total drops/no-buffer drops) 12/16501/0
```

Gdy na tym etapie użytkownicy się skarżą, można rozważyć kilka przyczyn:

- Łącze może być wysycone i w takim przypadku należy rozważyć wymianę na większe.
- Pewien ruch w kolejce domyślnej powinien być priorytetyzowany.
- Jedna lub więcej kolejek może być błędnie skonfigurowana. Spójrz na kolejkę `HTTP`. Kiedy działa prawidłowo, `5 minut offered rate` wynosi tylko 1000 b/s. Kolejce przyznano 10% pasma. Jeśli zmniejszymy dla niej alokację pasma, kolejka domyślna może przestać odrzucać pakiety. Aktualnie, biorąc pod uwagę ilość ruchu na łączu, jest to mało prawdopodobne. `5 minut offered rate` dla kolejki domyślnej wynosi 887 000 b/s. To prawie dwie trzecie całego łącza T1.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zbyt niski priorytet kolejki

Szybkość dla kolejki `Voice-RTP` wydaje się dość stabilna i utrzymuje się na poziomie 80 kb/s. Dla zabawy zmniejszę rozmiar kolejki priorytetowej do 64 kb/s:

```
R1#conf t
R1(config)# policy-map WAN
R1(config-pmap)# class Voice-RTP
R1(config-pmap-c)# priority 64
```

A teraz spójrzmy na kolejki, korzystając z polecenia `show policy-map interface` *`nazwa_in`*
↪*`terfejsu`*:

```
R1#sho policy-map interface s0/1
  Serial0/1
    Service-policy output: WAN
      Class-map: Voice-RTP (match-any)
        19682025 packets, 4014346028 bytes
        5 minute offered rate 79000 bps, drop rate 17000 bps
        Match: ip precedence 5
          19682025 packets, 4014346028 bytes
          5 minute rate 79000 bps
        Queueing
          Strict Priority
          Output Queue: Conversation 264
          Bandwidth 64 (kbps) Burst 1600 (Bytes)
          (pkts matched/bytes matched) 19682025/4014346028
          (total drops/bytes drops) 1538/313752
```

Spójrz na statystyki odrzuconych pakietów. Pierwsza z pogrubionych linii wskazuje `drop rate` na poziomie 17 000 b/s. Wskazuje to na spory problem związany z tym, że rozmiar kolejki jest za mały. Zapis mówi nam, że kolejka zaoferowała do 79 000 b/s, ale musiała odrzucić 17 000 b/s. Aby zapobiec gubieniu pakietów, kolejka musi zostać ustawiona na co najmniej 79 000+17 000 = 90 000 b/s. Jeśli ujrzysz wartość inną niż 0 w parametrze `drop rate`, oznacza to problem.

Ostatnia zwrócona linia pokazuje całkowitą liczbę odrzuconych pakietów i bajtów. Ponieważ jest to licznik, będzie rósł do momentu, aż zostanie wyczyszczony, lub do momentu, kiedy interfejs zostanie zresetowany. Aby to zademonstrować, pozwoliłem routerowi działać z kolejką priorytetową o małym rozmiarze przez godzinę, potem ustawiłem pasmo ponownie na 128 kb/s. Odczekałem 5 minut, aby liczniki się ustabilizowały, i wykonałem komendę raz jeszcze:

```
R1#sho policy-map interface s0/1
  Serial0/1
    Service-policy output: WAN
      Class-map: Voice-RTP (match-any)
        20049682 packets, 4089333416 bytes
        5 minute offered rate 79000 bps, drop rate 0 bps
        Match: ip precedence 5
          20049682 packets, 4089333416 bytes
          5 minute rate 79000 bps
        Queueing
          Strict Priority
          Output Queue: Conversation 264
          Bandwidth 128 (kbps) Burst 3200 (Bytes)
          (pkts matched/bytes matched) 20049682/4089333416
          (total drops/bytes drops) 78648/16044192
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

`Drop rate` w pierwszej pogrubionej linii ma wartość 0, a więc tyle, ile powinien. Jednak ostatnia linia nadal pokazuje dość wysokie liczby. Jeśli zobaczyłbym coś takiego na poprawnie działającym łączu, podejrzewałbym jedną z dwóch rzeczy:

- Ktoś zmienił konfigurację tak, jak zrobiłem to ja, ale nie wyczyścił liczników po rozwiązaniu problemu.
- Chwilowo wzrósł ruch, który spowodował przepełnienie się skonfigurowanego rozmiaru kolejki.

W każdym z przypadków powinieneś określić, co się stało. Jeśli był to przypływ ruchu, może to świadczyć o tym, że wystąpiło więcej połączeń głosowych, aniżeli zakładałeś, lub też aktywne połączenia nie używają właściwego kodeka. Nawet gorzej, pakiety mogą być znakowane *IP precedence* równym 5, mimo że nie powinny.

Jeśli kolejka priorytetowa jest wysycona, mimo że dostępna będzie wolna przestrzeń pasma, pakiety będą odrzucane. Kolejka priorytetowa różni się tym od pozostałych, że nie przepełni się, wykorzystując dostępne wolne pasmo na łączu. Pakiety będą gubione, jeśli kolejka jest za mała.

Zbyt duża kolejka priorytetowa

Jeśli kolejka priorytetowa jest za duża, może to generować problemy wydajnościowe sieci. Problem taki może być bardzo trudno zdiagnozować. Poniżej rozmiar kolejki priorytetowej został ustawiony na 256 kb/s, tymczasem trafiający do niej ruch to jedynie około 80 kb/s:

```
Class-map: Voice-RTP (match-any)
  3150111 packets, 642496644 bytes
  5 minute offered rate 81000 bps, drop rate 0 bps
  Match: ip precedence 5
    3150111 packets, 642496644 bytes
    5 minute rate 81000 bps
  Queueing
    Strict Priority
    Output Queue: Conversation 264
    Bandwidth 256 (kbps) Burst 6400 (Bytes)
    (pkts matched/bytes matched) 3150111/642496644
    (total drops/bytes drops) 0/0
```

Ta kolejka wygląda na prawidłową i taka też jest. Problem jest nieco bardziej subtelny i na podstawie tych liczb nie do zdiagnozowania.

Kolejki są obsługiwane przez coś, co określa się jako **scheduler**. Obsługuje on każdą kolejkę, biorąc pod uwagę jej rozmiar oraz pozycję w konfiguracji. Kolejka priorytetowa jest zawsze obsługiwana jako pierwsza. Tylko wtedy, kiedy kolejka ta jest pusta, obsłużone zostaną pozostałe.

Widziałem problemy przy dużych łączach, jak DS3, gdzie kolejce priorytetowej ustawiono bardzo duży rozmiar. Na przykład przydzielenie takiej kolejce 75% przepustowości pasma powoduje, że brakuje go kolejkom poniżej. Innymi słowy, **scheduler** poświęca tyle czasu na obsługę kolejki priorytetowej, że inne kolejki, mimo iż ruch w nich jest niewielki, nie są obsługiwane. Skutkiem tego jest odrzucanie pakietów dla niższych kolejek, choć przepustowości pasma nie brakuje. Postaraj się utrzymywać rozmiar kolejki priorytetowej na poziomie, którego rzeczywiście wymaga.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zbyt małe kolejki bez priorytetu

W naszym przykładzie korzystamy z łącza T1 z kolejką `Voice-Control`, której przyznaliśmy 5% łącza. 5% z 1,544 Mb/s to około 77 kb/s. Zwiększyłem ruch **protokołów sterujących połączeniami** tak, aby przekroczył rozmiar kolejki. Przepycham teraz przez kolejkę 90 kb/s:

```
Class-map: Voice-Control (match-all)
  213781 packets, 40614182 bytes
  5 minute offered rate 90000 bps, drop rate 0 bps
  Match: ip precedence 3
  Queueing
    Output Queue: Conversation 265
    Bandwidth 5 (%) Max Threshold 64 (packets)
    (pkts matched/bytes matched) 213900/40638220
    (depth/total drops/no-buffer drops) 3/0/0
```

Zauważ, że kolejka nie odrzuca pakietów, mimo że jest zbyt mała. Zachowanie to jest znacząco inne od tego, co widzieliśmy wcześniej dla kolejki priorytetowej. Wysycona kolejka priorytetowa odrzuca pakiety, nawet jeśli na łączu istnieje niewykorzystane pasmo. Pozostałe kolejki nie będą gubiły pakietów do momentu, kiedy całe łącze zostanie wykorzystane. `3/0/0` w ostatniej linii pokazuje głębokość kolejki (ang. *queue depth*), całkowitą liczbę pakietów odrzuconych (ang. *total drops*) oraz liczbę pakietów odrzuconych w wyniku przepełnienia bufora (ang. *no-buffer drops*). Głębokość kolejki wskazuje, że trzy pakiety są zakolejkowane. Jest to rzadkość, nawet przy utylizacji łącza na poziomie powyżej 50%:

```
R1#sho int s0/1 | include minute
  5 minute input rate 883000 bits/sec, 830 packets/sec
  5 minute output rate 897000 bits/sec, 846 packets/sec
```

Nawet jeśli zwiększę ruch tak, że kolejka będzie miała trzy razy mniejszą pojemność, dopóki łącze nie będzie wysycone, pakiety nie będą odrzucane:

```
Class-map: Voice-Control (match-all)
  281655 packets, 54324322 bytes
  5 minute offered rate 237000 bps, drop rate 0 bps
  Match: ip precedence 3
  Queueing
    Output Queue: Conversation 265
    Bandwidth 5 (%) Max Threshold 64 (packets)
    (pkts matched/bytes matched) 281936/54381084
    (depth/total drops/no-buffer drops) 0/0/0
```

Dla nasyconego łącza kolejka będzie odrzucała pakiety, jeśli będzie przepełniona.

Zbyt duża kolejka bez priorytetu

Podobnie jak przy kolejce priorytetowej, także gdy kolejka bez priorytetu jest zbyt duża, problem może być niebanalny. **Scheduler** może spędzać zbyt wiele czasu, obsługując kolejkę zupełnie niepotrzebnie, co może skutkować gubieniem pakietów przez kolejki z mniejszym priorytetem, ale tylko wtedy, gdy obsłużona zostanie kolejka priorytetowa. Poniżej sytuacja, w której `HTTP` alokowano 40% szybkości łącza. Dla łącza T1 przekłada się to na około 600 kb/s, jednak kolejka oferuje jedynie 1 kb/s:

```
Class-map: HTTP (match-any)
  79255 packets, 4690398 bytes
  5 minute offered rate 1000 bps, drop rate 0 bps
  Match: access-group 101
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
      79255 packets, 4690398 bytes
      5 minute rate 1000 bps
    Queueing
      Output Queue: Conversation 266
      Bandwidth 40 (%) Max Threshold 64 (packets)
      (pkts matched/bytes matched) 78903/4669112
      (depth/total drops/no-buffer drops) 0/0/0
  Class-map: class-default (match-any)
    23750755 packets, 2030244008 bytes
    5 minute offered rate 887000 bps, drop rate 4000 bps
    Match: any
    Queueing
      Flow Based Fair Queueing
      Maximum Number of Hashed Queues 256
      (total queued/total drops/no-buffer drops) 12/16501/0
```

Zauważ, że kolejka domyślna gubi pakiety z szybkością 4000 b/s.

Gdy tylko zmienię w HTTP przydzielone pasmo z 40% na 5%, router prawie kompletnie przestanie odrzucać pakiety w kolejce domyślnej:

```
  Class-map: HTTP (match-any)
    86935 packets, 5150618 bytes
    5 minute offered rate 1000 bps, drop rate 0 bps
    Match: access-group 101
      86935 packets, 5150618 bytes
      5 minute rate 1000 bps
    Queueing
      Output Queue: Conversation 266
      Bandwidth 5 (%) Max Threshold 64 (packets)
      (pkts matched/bytes matched) 86204/5106232
      (depth/total drops/no-buffer drops) 0/0/0
  Class-map: class-default (match-any)
    26519010 packets, 2404610690 bytes
    5 minute offered rate 849000 bps, drop rate 0 bps
    Match: any
    Queueing
      Flow Based Fair Queueing
      Maximum Number of Hashed Queues 256
      (total queued/total drops/no-buffer drops) 0/32252/0
```

Podstęp polega na tym, że jedna kolejka wpływa bezpośrednio na zachowanie innej, dokładnie tej znajdującej się pod nią. W tym przypadku nie mogłem zrobić niczego, aby poprawić wydajność kolejki domyślnej. Zmian należało dokonać w kolejce powyżej.

Zbyt mała kolejka domyślna

Rozmiar kolejki domyślnej nie jest bezpośrednio konfigurowalny. Gdy kolejka ta jest zbyt mała, oznacza to, że albo łącze jest zbyt małe, albo inne kolejki są zbyt duże. Niezależnie od tego, co jest powodem, skończy się to tym, że pakiety w kolejce domyślnej będą odrzucane.

Zbyt duża kolejka domyślna

Uważam, że kolejka domyślna nie może być zbyt duża. Jeśli dysponujesz dużą klasą domyślną kolejki, oznacza to, że masz wystarczająco dużą pojemność, co czyni życie prostszym, przynajmniej do czasu!

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 35.

Projektowanie sieci

Istnieją setki książek, które opiszą Ci, jak zbudować trójwarstwową sieć korporacyjną. To nie jest jedna z nich. Zamierzam pokazać Ci, co powinieneś zrobić, **zanim** zbudujesz taką sieć. Następnie zaprezentuję Ci kilka zaimplementowanych już przykładów projektów sieci.

Nie jest to ten typ technicznych informacji, których potrzebujesz, aby uzyskać certyfikat. Wiedza ta pozwoli Ci raczej lepiej wykonywać Twoją pracę. Rozdział ten w większości zakłada, że będziesz projektował swoją sieć od podstaw. Mimo że nie zawsze tak jest, to informacje tu zamieszczone można zastosować do każdego projektu związanego z siecią.

Dokumentacja

Dokumentacja jest zmorą wielu inżynierów. Nie wiem, dlaczego tak jest, ale inżynier lubiący pisać dokumentację stanowi rzadkość. Pisanie jest ciężką pracą (spróbuj napisać książkę!), ale zyski z tego są ogromne.

Niektórzy inżynierowie uważają, że jeśli będą gromadzić informacje, staną się niezastąpieni. Zaufaj mi w tej kwestii — zarówno mnie, jak i Ciebie zawsze można zastąpić. Właściwie to zarabiałem na życie, pisząc dokumentację, gdy „niezastąpiony" pracownik został zwolniony.

Dobrze napisana dokumentacja oszczędza czas i pieniądze. Jeżeli ktoś potrafi naprawić sieć po lekturze Twojej dokumentacji, oznacza to, że wykonałeś dobrą robotę. Odwrotnie jest, jeśli tylko Ty jesteś w stanie naprawić swoją sieć.

Kiedy to tylko możliwe, nawet w małych sieciach, powinieneś dokumentować każdy detal. Nawet gdy sieć jest tak mała, że pamiętasz każdy jej adres. Zacznij od swojej sieci domowej. Moja ma lepszą dokumentację aniżeli ta w firmie, w której pracuję. Powiem jednak szczerze, że jest ona prawdopodobnie bardziej skomplikowana niż niejedna sieć firmowa. Jestem dumny ze swojej dokumentacji, częściowo przyczyniła się ona do zdobycia przeze mnie pracy. Mając do wyboru dwóch równie uzdolnionych inżynierów, zatrudnię tego, który potrafi wykonać lepszą dokumentację.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Specyfikacja wymagań

Na początku każdego projektu, niezależnie od dostępnego czasu, należy napisać specyfikację wymagań. Dokument ten powinien zawierać wszystkie wymagania projektu oraz poczynione założenia. Nawet jeśli wymaganie jest proste: „projekt nowej sieci korporacyjnej", zapisz to i załącz wszystkie założenia, które zrobisz, aby osiągnąć planowany cel.

Kiedy przygotujesz dokument specyfikacji wymagań, wyślij go do wszystkich zaangażowanych. Na liście tej powinien znaleźć się, absolutnie koniecznie, Twój szef. W zależności od kultury korporacji mogą znaleźć się na niej również kierownicy projektów, właściciele biznesowi danego projektu (sponsorzy) i każda inna osoba związana z procesem. Twój szef może chcieć sam przesłać dokument do pozostałych osób. Niezależnie od tego powinieneś opublikować coś wskazującego na to, co będziesz projektował i dlaczego.

W momencie kiedy zakres projektu ulegnie zmianie, stworzona specyfikacja wymagań będzie najlepszym sposobem zabezpieczenia własnej osoby.

Zakres projektu ulegnie zmianie. Nie przypominam sobie żadnego przypadku, w którym tak się nie stało. Zmiana zakresu projektu jest powszechna w świecie projektów.

Na przykład jeśli zamówiłeś łącze internetowe T1, a dyrektor po jakimś czasie ma do Ciebie żal, ponieważ nie sądził, że pasmo będzie aż tak szerokie, powinieneś być w stanie wskazać mu dokument, który opublikowałeś na początku projektu (na szczęście usprawiedliwiający Twoją decyzję). Zdejmuje to z Ciebie cały ciężar odpowiedzialności, jako że wcześniej wysłałeś dokument opisujący zapotrzebowanie związane z projektem. Jeżeli nie wiedział o T1, oznacza to, że nie czytał dokumentu.

Specyfikacja wymagań nie musi być rozbudowana i skomplikowana — właściwie im prostszy dokument, tym łatwiej będzie go zrozumieć. Każdy z punktów powinien być tak prosty, jak to tylko możliwe (w szczególności założenia). Oto przykład specyfikacji wymagań:

Wymaganie:

- Sieć musi obsługiwać 300 użytkowników.

Założenia:

- Każdy użytkownik będzie dysponował jedną stację roboczą.
- Każda stacja robocza będzie wyposażona w jeden interfejs Ethernet.
- Do wszystkich interfejsów podłączony zostanie kabel 1 Gb Ethernet.
- Sieć nie musi wspierać przepustowości 1 Gb/s dla każdego użytkownika w tym samym czasie.
- Każdy użytkownik będzie miał do dyspozycji jeden telefon z jedną linią telefoniczną.
- Żaden z telefonów nie będzie telefonem IP.
- Wdrożenie VoIP nie jest rozważane w możliwej do przewidzenia przyszłości.
- Każdy użytkownik równa się jednemu stanowisku lub biuru.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

- Każde stanowisko lub biuro będzie wyposażone w jedno gniazdko internetowe oraz jedno gniazdko telefoniczne.
- Okablowanie zostanie zaterminowane w małej serwerowni na piętrze.

Kiedy dyrektor mówi: „Potrzebujemy sieci obsługującej 300 użytkowników", nie rozumie często implikacji swojego prostego stwierdzenia. Twoja rola polega częściowo na tym, aby się upewnić, że inni zaangażowani ludzie posiadają wszystkie niezbędne informacje umożliwiające zrozumienie, co się dzieje, a częściowo na tym, by chronić swój własny tyłek, na wypadek gdyby ktoś czegoś nie zrozumiał. Dyrektor będzie w stanie przeczytać opublikowane przez Ciebie założenia i zadać pytania, jeśli będzie je miał. Może istnieje plan na zmianę systemu telefonii na VoIP w przyszłym roku i musisz zostać poinstruowany, aby zamówić kompatybilny z tą wizją sprzęt.

Widziałem wiele nieudanych projektów, gdzie przyczyną porażki był brak specyfikacji wymagań. Byłem również świadkiem (oraz stroną) wymiany wielu gorących argumentów związanych z tym, co było powiedziane pół roku temu, ale nie zostało zapisane.

Tylko dlatego, że ludzie mówią, iż coś rozumieją, nie zakładaj, że tak jest — w szczególności wtedy, kiedy są to osoby „nietechniczne". Zaprojektowałem kiedyś opracowanie mechanizmu przełączenia na zapasowe łącze internetowe dla pewnej firmy. Kiedy dostarczono DS3 (po 60-dniowym czasie realizacji), dyrektor firmy, który podpisał zamówienie, zaczął na mnie krzyczeć, ponieważ nie wiedział, po co nam to jest potrzebne. Nigdy nie pomyślałbym, że nie rozumie projektu. Założyłem, że jesteśmy „po tej samej stronie", ale jako że nie spisałem odpowiedniego dokumentu, dosłownie nie istniała strona, po której moglibyśmy być.

Arkusz układu portów

Po spisaniu i zatwierdzeniu wymagań można rozpocząć fazę projektowania. Pierwszym krokiem, jaki zawsze wykonuję, rozpoczynając projektowanie sieci, jest spisanie wszystkich urządzeń, które będą miały do niej dostęp. W przypadku sieci LAN jest kilka rzeczy, na które warto zwrócić uwagę:

- Ilu użytkowników będzie obsługiwała sieć?
- Ile serwerów będzie się znajdowało w sieci?
- Ile drukarek będzie dołączonych do sieci i gdzie będą zlokalizowane?
- Jakie aplikacje będzie obsługiwała sieć? Jak użytkownicy będą z tych aplikacji korzystali (http, oprogramowanie klienckie, terminale, Citrix)?
- Jakiego typu ochrony potrzeba?
- Czy istnieje wymóg wdrażania rozwiązań wysokiej dostępności i czy jest to opłacalne?
- Jaki jest przewidywany trend rozwojowy (w procentach)?
- Czy wszystkie interfejsy powinny być gigabitowe?
- Czy potrzeba interfejsów 10-gigabitowych?
- Czy sieć powinna wspierać VoIP?
- Czy wspierana będzie jedna fizyczna lokalizacja, czy wiele (wliczając w to wiele pięter w obrębie pojedynczego budynku)?

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Twoim celem jest uzyskanie docelowego numeru każdego interfejsu sieciowego w każdej z lokalizacji. Kiedy będziesz miał takie zestawienie, możesz zdecydować o typie potrzebnego sprzętu. Przełączniki Gigabit Ethernet składają się z wielokrotności 48 portów. Aby określić, ile przełączników lub modułów potrzebujesz, podziel wymaganą sumaryczną liczbę interfejsów gigabitowych przez 48. Na przykład jeżeli w lokalizacji znajduje się 340 interfejsów gigabitowych, będziesz potrzebował 340/48 = 7,08 (czyli 8 modułów).

Wyznaczając liczbę potrzebnych portów, nie zapomnij uwzględnić serwerów zapewniających wysoką dostępność. Wiele współczesnych serwerów pozwoli na podłączenie swoich interfejsów do dwóch przełączników, tak aby zapewnić mechanizm *failover*.

Powinieneś również planować, uwzględniając określony wzrost. W planowaniu pojemności dobrą praktyką jest przyjęcie 15% wzrostu: 340·1,15 = 391. Podziel tę liczbę przez 48, aby zobaczyć, ilu modułów będziesz potrzebował, jeśli weźmiesz pod uwagę rozwój: 391/48 = 8,15. Oznacza to liczbę dziewięciu modułów zapewniających 432 porty. Dziewięć nie jest dobrą liczbą, ponieważ oznacza, że jeden przełącznik będzie miał więcej modułów. Może to nie mieć żadnego znaczenia w Twoim środowisku, choć ja lubię, aby wszystko do siebie idealnie pasowało. Lepiej jest zawsze mieć więcej portów aniżeli mniej, dopóki oczywiście budżet na to pozwala. Zaokrąglenie liczby modułów do 5 w każdym *chassis* spowoduje, że będziesz dysponował 10 modułami i 480 portami.

Masz zatem 480 dostępnych portów oraz aktualne wymaganie na poziomie 340. Współczynnik wzrostu wynosi wtedy niemal 30%. Istnieje szansa, że coś pominąłeś (zawsze to robię), więc możliwości rozwoju na poziomie powyżej 15% nie są niczym nagannym. W przypadkach jak ten upewniam się, czy zarezerwowałem dodatkowe miejsce na cele rozbudowy. Później zawsze się to przydaje.

Kolejnym wymaganiem, o którym inżynierowie często zapominają, są połączenia między przełącznikami. Switche będące w parze *failover* powinny być ze sobą połączone, zazwyczaj wielogigabitowymi łączami. Jeżeli korzystasz z modułów firewall, powinny mieć one dedykowane linki. Moduły ACE powinny również mieć dedykowane łącza, tak jak i RSPAN, jeśli go używasz. Przyjmując 2 Gb/s *EtherChannel* dla każdego z opisanych połączeń, alokujesz aż 16 portów (8 po każdej ze stron).

Pracowałem z sieciami o wysokiej przepustowości, gdzie każde z tych połączeń wymagało konfiguracji 4 Gb/s *EtherChannel*. Należało zarezerwować 32 porty (16 po każdej stronie). Wyłącznie na potrzeby komunikacji między przełącznikami wymagany był prawie cały moduł z 48 portami!

Jeżeli potrzebujesz portów 10 gigabit, pamiętaj, że porty te mogą być bardzo drogie i jeśli nie posiadasz przełączników Nexus, ich liczba może być problemem. Uważaj na porty 10 Gb. Nowoczesne klatki serwerowe mają po dwa porty 10 Gb, przez co mogą się one bardzo szybko wysycić.

Kiedy już określiłeś liczbę niezbędnych portów, zaplanuj rodzaj sprzętu, jakiego będziesz potrzebował, i dowiedz się, które urządzenia spełniają Twoje wymagania. Ja zdecydowałem, że w rdzeniu sieci zastosowane zostaną przełączniki 6509-V-E, łączność z internetem oraz dostęp do strony korporacyjnej zapewnią routery 2921 zlokalizowane w serwerowni. Rysunek 35.1 przedstawia arkusz zwierający powyższe informacje:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

					Interfejsy			
Nazwa	Urządzenie	Zastosowanie	Lokalizacja	Gniazda	T1	DS3	1G	10G
Rdzeń-1	6509-V-E	Przełącznik/router rdzenia	Szafa #2	9	0	0	240	0
Rdzeń-2	6509-V-E	Przełącznik/router rdzenia	Szafa #3	9	0	0	240	0
Internet	2921	Router internetowy	Szafa #2	1	2	0	2	0
Serwerownia	2921	Połączenie z serwerownią	Szafa #3	1	2	0	2	0

Rysunek 35.1. Przykładowa lista sprzętu

Kiedy już zdecydowałeś, jakiego sprzętu użyć, powinieneś zapoznać się z każdym z urządzeń, aby wiedzieć dokładnie, jaki sprzęt zamówić. Krok ten znacznie ułatwi Ci proces projektowania i budowania sieci. W momencie gdy otrzymasz sprzęt, możesz poszerzyć swój arkusz o numery seryjne każdej z części. Rysunek 35.2 przedstawia przykład planu dla routera 6509, który wybrałem. Zauważ, że uwzględniłem nie tylko moduły, ale również dodatkową pamięć. Nie zaprzątaj sobie głowy skomplikowanymi nazwami — Twój sprzedawca wyjaśni Ci wszystko podczas kompletowania zamówienia, wraz z opisaniem wszystkich dodatkowych elementów.

				Interfejsy			
Nazwa	Urządzenie	Zastosowanie	Gniazda	10/100/1G	GBIC	SF-GBIC	10G
Rdzeń-1	**WS-C6509-VE**	Core Switch/Router	**9**	**244**	**0**	**4**	**12**
Slot1	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade		48			
Slot2	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade		48			
Slot3	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade		48			
Slot4	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade		48			
Slot5	VS-S720-10G-3C	Supervisor - 720 10G		2		2	2
Slot6	VS-S720-10G-3C	Supervisor - 720 10G		2		2	2
Slot7	WS-SVC-FWM-1-K9	Firewall Switch Module					
Slot8	WS-X6708-10GE	8-port 10G Module					8
Slot9	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade		48			
MSFC Mem	MEM-MSFC2-512MB	MSFC - 512M DRAM					
Sup Mem	MEM-S2-512MB	Sup - 512M DRAM					
Flash Mem	MEM-C6K-CPTFL256M	1G Compact Flash Upgrade					
Fan Tray	WS-C6509-E-FAN	Fan Tray					
PS #1	WS-CAC-3000W-US	6000W Power Supply					
PS #2	WS-CAC-3000W-US/2	6000W Power Supply					

Rysunek 35.2. Przełącznik rdzenia sieci

Kiedy już opracowałeś listę urządzeń, samo zamówienie pozostaje w gestii kadry zarządzającej. Będziesz prawdopodobnie musiał pomóc przy zamawianiu, wskazując numery części lub też odpowiednie ilości. Po wysłaniu zamówienia czas dostawy sprzętu wyniesie kilka tygodni.

W kolejnym kroku należy przypisać każdy interfejs do urządzenia. Może to zabrzmieć nieco przesadnie, ale planowanie, gdzie każde z urządzeń zostanie podłączone, zaoszczędzi wiele nerwów w przyszłości. Posiadanie takiego planu pozwoli Ci również przekazać część czynności związanych z podłączaniem urządzeń innym osobom. Krok ten stanowi pracę nad sporym kawałkiem dokumentacji w czasie, gdy czekasz na realizację zamówienia.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Tworzę arkusz dla każdego z urządzeń, a czasem dla każdego modułu urządzenia. Nie musi to być nic skomplikowanego. Prosta lista portów, podłączonych do nich urządzeń i może IP albo VLAN (tak jak to prezentuje rysunek 35.3), aby zostały skonfigurowane poprawnie.

Port	VLAN/Trunk/IP	Urządzenie	Zdalny interfejs
1/1	VLAN 10	Router internetowy	F0/0
1/2		Zarezerwowany	
1/3			
1/4			
1/5	VLAN 777	Rdzeń-2 FWSM Failover link-1	G1/5
1/6	VLAN 777	Rdzeń-2 FWSM Failover link-2	G1/6

Rysunek 35.3. Przykład przypisania portów

Jeżeli zaplanowałeś swoją sieć do tego etapu, kiedy przybędzie sprzęt, wszystko, co będziesz musiał zrobić, to jego montaż oraz okablowanie urządzeń.

Arkusze IP i VLAN

Wraz z rozplanowaniem sprzętu i alokacją portów należy zaplanować sieć IP oraz schemat VLAN-ów. Lubię budować dość szczegółowe arkusze zawierające te informacje, podobnie jak zrobiłem to dla urządzeń, modułów i portów.

Rysunek 35.4 prezentuje alokowane przeze mnie sieci IP. Przypisuję sieć /23 dla każdego VLAN-u i rezerwuję /23 za każdym z przydziałów, uwzględniając przyszły rozwój.

Sieć	Maska	VLAN	Opis
10.1.0.0			
10.1.1.0			
10.1.2.0			
10.1.3.0			
10.1.4.0			
10.1.5.0			
10.1.6.0			
10.1.7.0			
10.1.8.0	255.255.254.0	VLAN 10	Internet DMZ
10.1.9.0			
10.1.10.0			*Zarezerwowany na potrzeby rozwoju*
10.1.11.0			*Zarezerwowany na potrzeby rozwoju*
10.1.12.0	255.255.254.0	VLAN 100	VLAN 100
10.1.13.0			
10.1.14.0			*Zarezerwowany na potrzeby rozwoju*
10.1.15.0			*Zarezerwowany na potrzeby rozwoju*

Rysunek 35.4. Układ arkusza zawierającego sieci IP

Stworzyłem arkusze dla każdej sieci IP, której będę używał, i wypełniłem je konkretnymi informacjami, uwzględniając każde z urządzeń. Stanowi to doskonałe ćwiczenie, które zmusza do ponownego przemyślenia elementów podłączanych do sieci. Rysunek 35.5 przedstawia przykładowy arkusz z układem adresów IP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Adres IP	Maska podsieci	VLAN	Opis
10.1.32.0			Sieci DC
10.1.32.1	255.255.254.0	VLAN 130	Brama domyślna (HSRP VIP)
10.1.32.2	255.255.254.0	VLAN 130	Rdzeń-1 VLAN 130 IP
10.1.32.3	255.255.254.0	VLAN 130	Rdzeń-2 VLAN 130 IP
10.1.32.4	255.255.254.0	VLAN 130	*Rezerwacja*
10.1.32.5	255.255.254.0	VLAN 130	*Rezerwacja*
10.1.32.6	255.255.254.0	VLAN 130	*Rezerwacja*
10.1.32.7	255.255.254.0	VLAN 130	*Rezerwacja*
10.1.32.8	255.255.254.0	VLAN 130	Kolorowa drukarka w biurze Andrzeja
10.1.32.9	255.255.254.0	VLAN 130	Kolorowa drukarka w biurze Ani
10.1.32.10	255.255.254.0	VLAN 130	Kolorowy skaner #1
10.1.32.11	255.255.254.0	VLAN 130	Kolorowy skaner #2

Rysunek 35.5. Arkusz przedstawiający schemat adresów IP

Udokumentowanie wszystkich tych informacji sprawi, że budowa konfiguracji odbędzie się błyskawicznie. Dodatkowo arkusze mogą posłużyć jako dokumentacja sieci. Można je wydrukować, umieścić w segregatorach i przechowywać w łatwo dostępnym miejscu.

Po wybudowaniu sieci proponuję zrobić kopię wszystkich dokumentów, które zostały wykonane. Dzięki temu, w przypadku gdy coś ulegnie zmianie, będziesz miał dokument opisujący sieć zaraz po jej wdrożeniu.

Schematy szaf dystrybucyjnych

Jeżeli nie znasz tego terminu, to schematy szaf dystrybucyjnych stanowią diagramy opisujące szczegółowo każdą z szaf (ang. *rack*). Jak dla mnie największą zaletą tego rozwiązania jest to, że jeśli stworzę takie schematy, to gdy zajdzie taka konieczność, ktoś inny może zainstalować sprzęt. Mogę poprzewracać wszystko do góry nogami, a potem, korzystając ze schematu, polecić: „Zróbcie, żeby to tak wyglądało". Przykład takiego schematu znajduje się na rysunku 35.6.

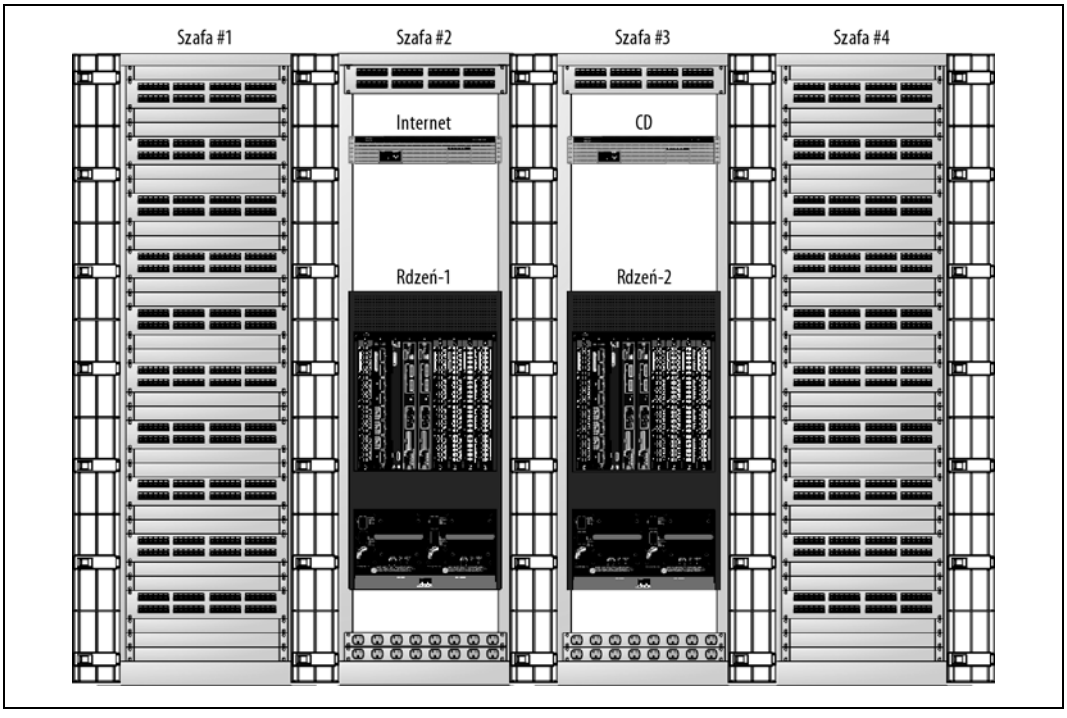
Istnieją trzy rzeczy, o których wielu inżynierów zapomina, projektując i planując zagospodarowanie przestrzeni w szafach: zasilanie, okablowanie i panele krosownicze (ang. *patch panels*). Każda z nich zajmuje miejsce.

Kosztowne może się okazać nieprawidłowe określenie wymagań związanych z zapotrzebowaniem na przestrzeń w szafach, szczególnie w przypadku dzierżawienia miejsca w centrum danych. Schematy szaf dystrybucyjnych są doskonałym sposobem na upewnienie się, że masz wystarczająco dużo szaf do obsługi wszystkich urządzeń, które zakupiłeś.

Wymagania dotyczące zasilania i chłodzenia

To jest dobry moment, aby wypracować wymagania związane z zasilaniem. Większość dostawców umieszcza pobór mocy swoich urządzeń na swoich stronach internetowych. Wystarczy to, aby dowiedzieć się, jakiego zasilania (AC/DC, napięcie, natężenie) oraz wtyczek potrzebują urządzenia, a następnie dodać te wymagania dla każdej z szaf. Porozmawiaj z osobą odpowiedzialną za infrastrukturę, w której umieszczasz szafy, i dowiedz się, czy wspierany jest instalowany przez Ciebie sprzęt.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 35.6. Schemat szaf dystrybucyjnych

To, że możesz umieścić 40 serwerów 1 U w szafie, nie oznacza, iż możesz je wszystkie włączyć! Domyślne zasilanie szaf w centrach danych to dwie listwy 20-amperowe. Jeśli Twoje serwery 1 U pobierają 2 ampery i każdy ma dwa zasilacze, możesz zainstalować jedynie 10 serwerów w każdej z szaf. Gdy chcesz dodać więcej, musisz zamówić dodatkowe zasilanie, które kosztuje. Będziesz również intensywniej korzystał z klimatyzacji, za co również należy zapłacić.

Nie zapomnij o wymaganiach związanych z chłodzeniem. Wraz ze specyfikacją zasilania urządzeń producenci dostarczają również wartości BTU (ang. *British Thermal Unit*). Osoba odpowiedzialna za środowisko, w którym sprzęt zostanie zainstalowany, powinna znać te informacje. Mimo fizycznej możliwości zainstalowania dwóch przełączników 6509 w szafie klimatyzacja może nie być w stanie obsłużyć wymagań związanych z ich chłodzeniem.

Wartości dotyczące zasilania i ciepła przełączników 6500 znajdziesz na stronie Cisco pod hasłem: *6500 power and heat numbers*. W związku z tym, że wartości te zależą w znacznym stopniu od zainstalowanych modułów, musisz zdobyć informację dla każdej z instalacji. Rysunek 35.7 zawiera dane dla przełącznika 6509-V-E. Powinieneś stworzyć podobny arkusz dla każdego z urządzeń, które będziesz instalował.

Pamiętaj o limitach związanych z liczbą urządzeń w szafie. W pierwszej edycji tej książki pokazałem, że możesz nie być w stanie umieścić w szafie więcej niż 10 – 12 urządzeń na skutek ograniczeń zasilania i chłodzenia. Od czasu pierwszego wydania książki wirtualizacja i serwery kasetowe znacząco się rozwinęły. Spowodowało to, że wiele firm zaimplementowało takie rozwiązania, co radykalnie zmieniło sposób budowania szaf.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

			Środowisko	
Nazwa	Urządzenie	Funkcja	Waty	BTU/godz.
Rdzeń-1	WS-C6509-V-E	9-Slot front-back airflow 6500 switch chassis		
Slot1	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade	367,50	1255,01
Slot2	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade	367,50	1255,01
Slot3	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade	367,50	1255,01
Slot4	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade	367,50	1255,01
Slot5	VS-S720-10G-3C	VSS Supervisor - 720 Fabric Enabled	422,63	1443,26
Slot6	VS-S720-10G-3C	VSS Supervisor - 720 Fabric Enabled	422,63	1443,26
Slot7	WS-SVC-FWM-1-K9	FWSM	214,73	733,29
Slot8	WS-X6708-10G-3C	8-port 10G Ethernet Module	555,45	1896,86
Slot9	WS-X6748-GE-TX	48 Port 10/100/1000 FE Blade	367,50	1255,01
MSFC Mem Upgrade	MEM-MSFC2-512MB	MSFC - 512M DRAM		
Sup Mem Upgrade	MEM-S2-512MB	Sup - 512M DRAM		
Flash Mem Upgrade	MEM-C6K-CPTFL512M	512M Compact Flash Upgrade	302,00	1031,00
Fan Tray	WS-C6509-V-E-FAN	Fan Tray		
Cable Management	WS-C6509-V-E-CM	Cable Management for 6509 V-E Chassis		
PS #1	WS-CAC-3000W-US	6000W Power Supply		
PS #2	WS-CAC-3000W-US/2	6000W Power Supply		
		Suma	3754,94	12822,72

Rysunek 35.7. Wartości mocy i BTU dla 6509-V-E

Podczas gdy kiedyś próbowaliśmy wepchnąć jak najwięcej serwerów do szafy, teraz staramy się umieścić jak najwięcej maszyn wirtualnych na serwerach fizycznych oraz jak najwięcej serwerów kasetowych w pojedynczej obudowie (ang. *blade enclosure*). Ma to ogromny wpływ na projektowanie szaf, głównie ze względu na spore wymagania związane z zasilaniem oraz chłodzeniem serwerów kasetowych. W wielu przypadkach możliwe jest umieszczenie wyłącznie jednej obudowy z serwerami fizycznymi w szafie właśnie ze względu na te ograniczenia. Obudowy serwerów kasetowych często zajmują jedną trzecią lub jedną czwartą szafy, pozostawiając znaczącą ilość niewykorzystanej przestrzeni. Zazwyczaj wymagają one również łączy 10 Gb w związku z obsługą znajdujących się w nich bardzo wielu serwerów wirtualnych.

Porady dotyczące diagramów sieci

Wydaje się, że inżynierowie lubią czynić swoje dokumenty tak skomplikowanymi, że nawet oni sami mają problemy z ich odczytaniem. Oto kilka porad, które powinieneś wziąć pod uwagę, tworząc dokumentację:

Zachowaj prostotę

Spójrz na dowolny rysunek zamieszczony w tej książce. Każdy z nich został zaprojektowany tak, aby stanowił prosty przekaz pojedynczej idei. Im bardziej będziesz się starał dołączyć rysunek do dokumentacji, tym trudniej będzie go zrozumieć.

Oddzielaj warstwę fizyczną od logicznej

Połączenia fizyczne są niezwykle ważne, ale postaraj się je oddzielić od kwestii VLAN-ów, routingu i innych aspektów związanych z warstwą logiczną. Lubię tworzyć dwa diagramy: jeden dla portów fizycznych, drugi dla adresów IP oraz VLAN. Przy zastosowaniu SVI staje się to tym bardziej ważne.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Nie krzyżuj linii

Za każdym razem, kiedy skrzyżujesz linie na diagramie, stanie się on trudniejszy w odbiorze. Czasami jest to nieuniknione, ale staraj się robić to jak najrzadziej.

Zorientuj linie proste

Kiedy zestawiasz linie na rysunku, na którym odstają one od pionu lub poziomu, Twój diagram wygląda niczym dzieło seryjnego mordercy. Jeśli poświęcisz czas, aby odpowiednio ułożyć linie, różnica będzie kolosalna. Podobnie linie rysowane pod kątem powinny, tam gdzie to możliwe, mieć zawsze to samo nachylenie.

Oddzielaj, jeśli możesz

Jeżeli na Twoim rysunku znajdują się dwie lokalizacje, oddziel je w jakiś sposób. Umieść każdą z nich w osobnym prostokącie (wielu ludzi używa prostokątów). Pomóc może skorzystanie z kolorów lub nawet odcieni szarości.

Uszereguj ikony

Jeśli na diagramie znajduje się rząd ikon, poświęć czas, aby je uszeregować wzdłuż jednej osi.

Konwencja nazewnicza dla urządzeń

Nazwy urządzeń powinny być tak konstruowane, aby osoba z elementarną wiedzą na temat sieci potrafiła określić funkcję danego sprzętu. Tymczasem wśród ludzi związanych z IT istnieje tendencja do nadawania nazw jak najbardziej niezrozumiałych. Oto przykład, z którym się spotkałem (aby chronić winnych, zmieniłem nazwę firmy i inne szczegóły): gadnslax1mai750901. Co to ma, do cholery, znaczyć? Czy to jest trafna nazwa? Czy bazując na niej, mogę powiedzieć, co to za system? Co ważniejsze, czy Ty możesz to zrobić?

Na jednej z sesji szkoleniowych związanych z tą siecią jeden ze studentów zapytał, co znaczy ta dziwaczna nazwa. Nikt nie potrafił odpowiedzieć bez zaglądania do schematu nazewnictwa. Nazwy urządzeń nie powinny wymagać poszukiwań! Oto rozkodowanie tej chybionej nazwy:

gadnslax1mai750901

gad — nazwa firmy (*GAD Technology*),
ns — usługi sieciowe (ang. *network services*),
lax — Los Angeles,
1 — pierwsze „coś" w Los Angeles,
mai — nazwa ulicy (Main Street),
7509 — urządzenie to Cisco 7509,
01 — pierwsze urządzenie 7509 w tej lokalizacji.

Celem nadawania nazw jest umożliwienie identyfikacji urządzenia. Nazwy powinny być łatwe do zapamiętania. Jeżeli jest je trudniej zapamiętać niż adresy IP, ich stosowanie nie ma sensu.

Gdy nazwa składa się z nazwy domeny, tworząc w ten sposób FQDN (ang. *fully qualified domain name*), czyli pełną, jednoznaczną nazwę domenową, otrzymany łańcuch powinien być prosty i zrozumiały. sw1.gad.net (*http://sw1.gad.net*) stanowi prosty i oczywisty FQDN opisujący pierwszy przełącznik w domenie **gad**.

Lubię nazwy opisujące funkcję danego urządzenia. Nazwa **WAN-Router** to doskonały przykład. W przypadku firm mających wiele lokalizacji użyteczne jest dodanie nazwy miejsca do określenia urządzenia (np. **LAX-WAN-Router**). Sprzęt zazwyczaj jest instalowany w parach, dobrym

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

pomysłem jest więc dodanie również numeru urządzenia (np. **LAX-WAN-Router-1**). Moim zdaniem nazwa ta jest za długa, a informacja, że sprzęt to router, nie ma znaczenia dla jego funkcji; **LAX-WAN-1** w zupełności wystarczy. Jeżeli chcesz użyć hierarchii DNS, **wan-1.lax-domain.com** również idealnie się nadaje.

Każda firma ma unikalne potrzeby. W przypadku dostawcy internetowego urządzenia będą rozmieszczone w lokalizacjach wielu klientów. Korzystne jest wówczas dołączenie nazwy klienta (np. **GAD-WAN-1**). Lepiej jednak oprzeć się pokusie umieszczania tej nazwy w dokumentacji. Nie powinieneś umieszczać numeru seryjnego w nazwie urządzenia (tak, spotkałem się z takimi przypadkami). Oczywiście każdy ma swoje zdanie na temat nadawania nazw urządzeniom, ale moja rada to trzymać się jak najprostszych nazw. Tak jak powiedziałem, **Agata** nie jest najlepszą nazwą dla routera. Mimo że jest to bardzo ładne imię, nie wskazuje w żaden sposób na jego funkcję!

Kiedy mam już gotowe nazwy urządzeń, dodaję do ich początku nazwy interfejsów w celu korzystania z DNS. Na przykład Serial 0/0/0:1 na **LAX-WAN-1** będzie miał nazwę w DNS **s0-0-0-1-lax-wan-1** (ponieważ w DNS można stosować jedynie myślniki, zastąpiłem nimi ukośniki i dwukropki). Dzięki temu trasy routingu są czytelne, ponieważ każdy węzeł i interfejs sieciowy jest jasno opisany:

```
[gad]$traceroute switch9.mydomain.com
traceroute to switch9.mydomain.com (10.10.10.10), 30 hops max, 40 byte packets
1 s0-0.router1.mydomain.com 9.854 ms 10.978 ms 11.368 ms
2 f0-1.switch2.mydomain.com 2.340 ms 1.475 ms 1.138 ms
3 g0-0-12.switch9.mydomain.com 1.844 ms 1.430 ms 1.833 ms
```

Projekty sieci

Nie powiem Ci, jak powinieneś projektować sieć. Mogę jednak pokazać Ci kilka popularnych projektów sieci korporacyjnych i stron *e-commerce* oraz wyjaśnić, dlaczego skupiłem się właśnie na nich.

Sieci korporacyjne

Większość sieci zaprojektowana jest zgodnie z zasadami **architektury trójwarstwowej**. Model taki składa się z **rdzenia** (ang. *core level*), **warstwy dystrybucyjnej** (ang. *distribution level*) oraz **dostępowej** (ang. *access level*). Poziomy te są jasno oddzielone od siebie i wspierane przez różne urządzenia. Tradycyjnie, routing jest najwolniejszym i najbardziej kosztownym procesem. Z tych powodów zlokalizowany jest on w rdzeniu. Pozostałe poziomy obsługują zazwyczaj przełączanie.

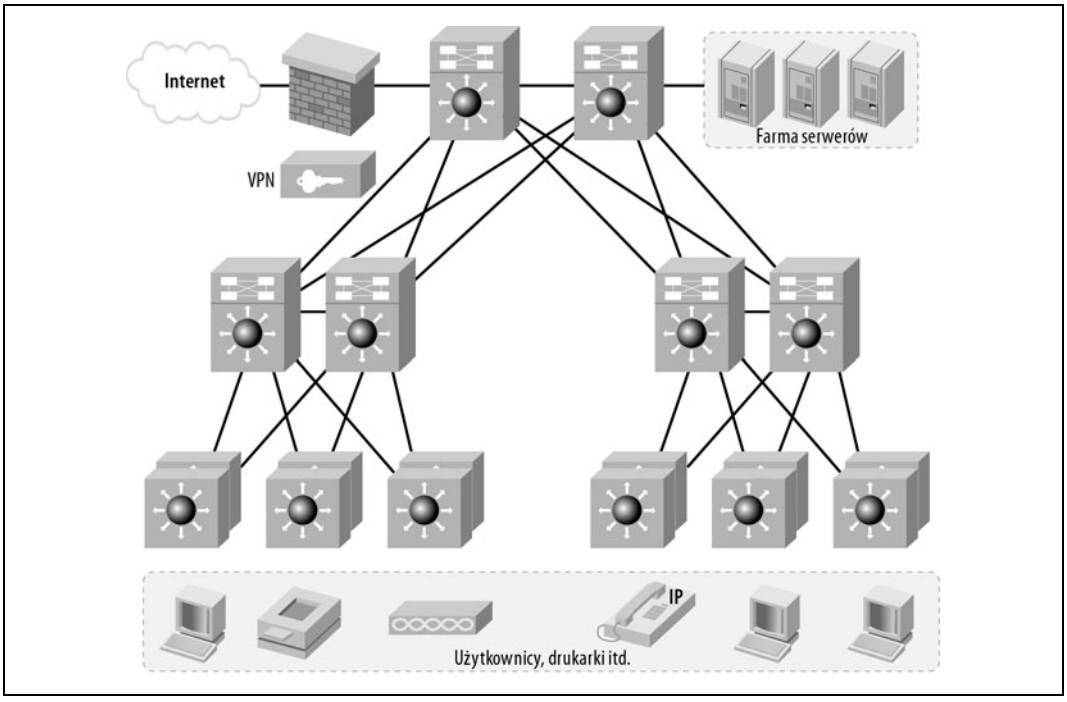
W związku z rozwojem tanich technologii przełączania w warstwie trzeciej, a tym samym łączeniem poziomów, model trójwarstwowy aktualnie zanika w sieciach korporacyjnych. Przyjrzymy się temu tradycyjnemu modelowi, ponieważ może być obecnie stosowany, ale także popularnym dzisiaj modelom sieci dwuwarstwowej i jednowarstwowej (ang. *collapsed core*).

Architektura trójwarstwowa

Architektura trójwarstwowa, najczęściej spotykana w podręcznikach, jest nadal szeroko stosowana na rynku. Fizyczna separacja warstw ma miejsce wtedy, kiedy fizycznie zachodzi taka konieczność. Doskonałym przykładem jest miasteczko studenckie: znajdują się tam przełączniki

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

rdzenia (prawdopodobnie w centralnej lokalizacji), przełączniki dystrybucyjne w każdym z budynków oraz przełączniki dostępowe jak najbliżej użytkowników. Szczegóły zależą od układu kampusu. Rysunek 35.8 przedstawia podręcznikową sieć w architekturze trójwarstwowej.



Rysunek 35.8. Typowa trójwarstwowa sieć korporacyjna

Na samym dole rysunku znajdują się użytkownicy, drukarki, telefony IP oraz bezprzewodowe punkty dostępowe (AP) podłączone do przełączników warstwy dostępowej. Przełączniki te dołączone są do warstwy dystrybucyjnej w środkowym rzędzie. Przełączniki dystrybucyjne podłączone są do dwóch przełączników na najwyższym poziomie, który stanowi rdzeń sieci. W tym przykładzie internet oraz farma serwerów łączą się z rdzeniem.

Architektura dwuwarstwowa — rdzeń połączony z warstwą dystrybucyjną

Sieci dwuwarstwowe są bardzo popularne. Pracowałem często na Manhatanie, gdzie drapacze chmur są normą. Przestrzenie biurowe zajmują tam każde z pięter. Okablowanie zazwyczaj biegnie z punktu na piętrze do centralnej lokalizacji znajdującej się na tym samym piętrze. Piętra są najczęściej połączone obwodami, które biegną między nimi. Kiedy firma znajduje się na kilku piętrach, w związku z ograniczeniami możliwości okablowania istnieje potrzeba zastosowania sieci opartej na architekturze dwuwarstwowej.

Ponieważ ilość miejsca na piętrze jest ograniczona, zazwyczaj nie ma konieczności stosowania więcej niż dwóch fizycznych warstw. Przełączniki dostępowe mogą pełnić rolę dystrybucyjnych — na jednym piętrze znajduje się rdzeń sieci, a na pozostałych przełączniki dostępowe.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Liczba portów nie stanowi zazwyczaj problemu, jako że piętra nie zajmują sporej przestrzeni. Również z perspektywy warstwy logicznej poziom przełączników dystrybucyjnych może wchodzić w skład warstwy rdzenia. Możesz w ogóle nie potrzebować logicznej warstwy dystrybucji. Raz jeszcze chciałbym podkreślić, że każde środowisko jest inne.

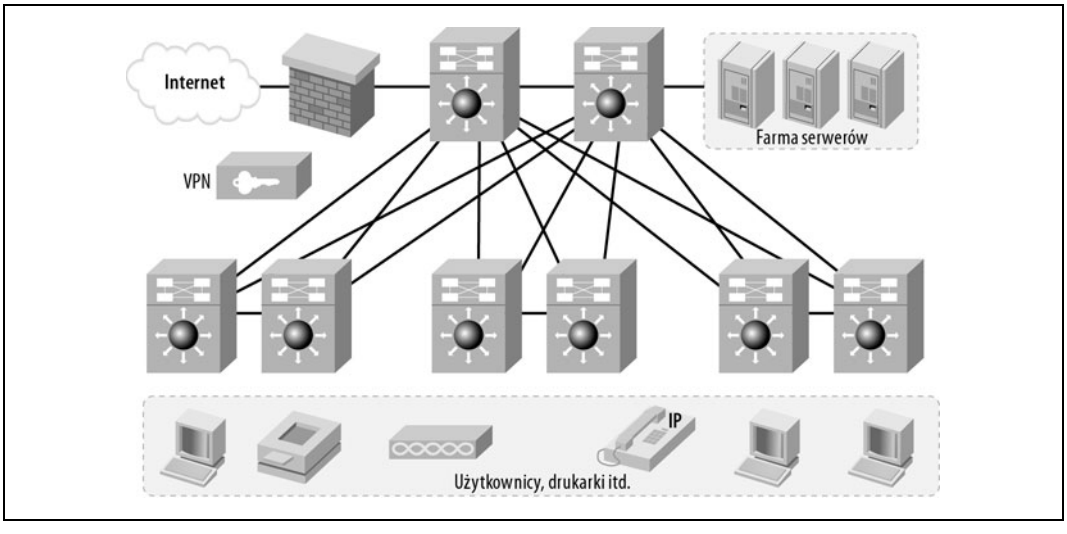
Farmy serwerów — gdzie jest ich miejsce?

Czy serwery znajdują się bliżej użytkowników, którym służą, czy raczej bliżej rdzenia sieci? To, gdzie przynależeć będą serwery, zależy od projektu sieci oraz od rodzaju danego serwera. Niektóre serwery powinny znajdować się w rdzeniu. Dla centralnego serwera poczty ma to sens. Inne serwery powinny być bliżej użytkowników, których obsługują. Na przykład w miasteczku studenckim nie ma sensu umieszczać w rdzeniu sieci serwerów księgowości, gdy cały departament zajmujący się tym znajduje się w jednym budynku. Budynki mogą też nie mieć odpowiedniej przestrzeni na serwery.

Wiele zależy również od schematu Twojej sieci. Aktualnie wiele firm ma zupełnie płaskie sieci, w których model rdzeń/dystrybucja/dostęp zawiera się w parze dużych przełączników, jak Cisco 6509. W takim przypadku wszystko podłączone jest do tych samych przełączników, kwestia serwerów zatem sama się rozwiązuje.

Innym scenariuszem, który nadaje się dla tego typu projektów, jest budynek składający się z segmentów. Pracowałem nad jednym projektem, który zakładał przeprojektowanie sieci dla dużego stadionu. Stadion ten podzielony był na cztery segmenty, każdy z nich był wyposażony we własną szafę. Przewody od użytkowników biegły do tej szafy, a same szafy połączone były ze sobą łączami światłowodowymi.

Rysunek 35.9 prezentuje przykład sieci w architekturze dwuwarstwowej pozbawionej warstwy dystrybucyjnej.

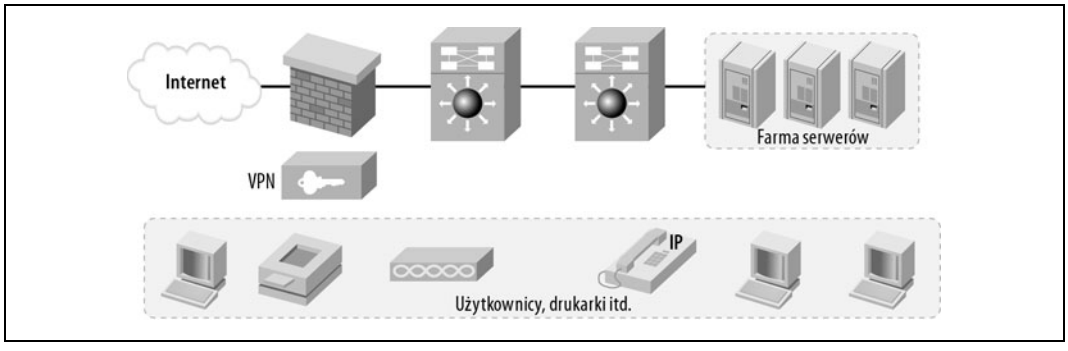


Rysunek 35.9. Sieć typu collapsed core bez warstwy dystrybucyjnej

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Architektura jednowarstwowa

Bardzo popularnym projektem sieci w firmach zajmujących wyłącznie jeden budynek jest sieć jednowarstwowa, w której znajdują się wyłącznie przełączniki rdzenia. Firmy zatrudniające setki pracowników pracujących w jednej lokalizacji często umieszczają serwerownię w centralnym punkcie całego środowiska. Dopóki nie przeszkadzają ograniczenia związane z zastosowanym okablowaniem, wszystkie połączenia mogą prowadzić bezpośrednio do przełączników rdzenia. Para wieloportowych oraz wysoko dostępnych przełączników Cisco 6509 może obsługiwać setki użytkowników. Taki projekt sieci zaprezentowany jest na rysunku 35.10.



Rysunek 35.10. Sieć typu collapsed core składająca się tylko z jednej warstwy

Kwestie konfiguracji

Nie zamierzam tutaj przechodzić przez szczegółowe kroki dotyczące konfiguracji sieci korporacyjnej. Zamiast tego wskażę pewne elementy, które powinieneś rozważyć, projektując taką sieć, oraz postaram się pokazać Ci prawidłowy kierunek. Jako przykład wykorzystam model dwuwarstwowy. Jako odniesienia możesz użyć rysunku 35.9.

Trunki. Trunki mogą być konieczne wszędzie tam, gdzie przełączniki są ze sobą połączone. Oczywiste będzie ich zastosowanie pomiędzy parą przełączników w rdzeniu oraz pomiędzy przełącznikami dostępowymi i rdzeniowymi. Nie zapomnij również o połączeniach między każdą parą przełączników dostępowych.

Możesz zdecydować się na zaprojektowanie sieci, w której połączenia między przełącznikami będą nie trunkami, a połączeniami IP. Jest to w pełni poprawne i ułatwi zrozumienie sieci innym ludziom. Jeśli zdecydujesz się na warstwę trzecią, projektując sieć, będziesz potrzebował mniejszej liczby trunków.

EtherChannels. Lubię łączyć swoje przełączniki za pomocą co najmniej dwóch połączeń w konfiguracji *EtherChannel*. Wzmacnia to nieco odporność oraz zwiększa pasmo między przełącznikami. W zależności od Twojej konfiguracji możesz chcieć zwiększyć rozmiar wiązki do 3, 4 lub więcej połączeń.

Możesz posiadać serwery, które wymagają *EtherChannel*. Zapytaj zespołu zarządzającego serwerami, czy ta funkcjonalność jest wymagana i jak powinna zostać skonfigurowana. Sprawdź w rozdziale 18. szczegóły związane z vPC (ang. *Virtual Port Channels*), które mogą łączyć wiele urządzeń za pomocą przełącznika Nexus. Podobnie system 6500 VSS może wspierać MEC (ang. *Multichassis EtherChannel*).

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli Twoje rdzeniowe przełączniki będą miały moduły firewall wymagające dodatkowych połączeń między parami *failover*, będziesz musiał przewidzieć w takim przypadku również zastosowanie *EtherChannel*. Więcej informacji o *EtherChannel* znajdziesz w rozdziale 7.

Spanning tree. Określ, które porty będą przeznaczone dla użytkowników i serwerów, oraz skonfiguruj je następująco: `spanning-tree portfast`. Ustaw jeden z przełączników rdzenia jako mostek główny STP, a drugi jako jego zapas. Więcej informacji na ten temat znajdziesz w rozdziale 8.

VTP. Nie jestem wielkim fanem VTP (ang. *Virtual Trunking Protocol*), szczególnie dla małych sieci, ale nawet jeśli go nie używasz, będziesz musiał skonfigurować nazwę domenową VTP. Jej poprawne ustawienie narzuca negocjacja trunków. Więcej informacji na ten temat znajdziesz w rozdziałach 5. i 6.

VLAN. Ilu VLAN-ów będziesz potrzebował? Upewnij się, że ich planowanie odbędzie się z wyprzedzeniem. Oto lista, jaką sporządziłem, spoglądając na rysunek 35.9:

- internet,
- internet wewnętrzny,
- farma serwerów,
- VLAN-y użytkowników.

Planowanie z wyprzedzeniem uchroni Cię przed podejmowaniem decyzji w ostatniej chwili i stratą czasu.

Strony e-commerce

Strony *e-commerce* to takie, które służą wymianie towarów i pieniędzy, dlatego właśnie nie są to proste witryny internetowe. Największym wyzwaniem jest bezpieczeństwo. Jeśli skonfigurujesz serwer Linux, a na nim prostą stronę internetową, Twoim zmartwieniem będzie jedynie bezpieczeństwo serwera WWW. Strony związane z handlem internetowym (*e-commerce*) opierać się będą prawdopodobnie na wielu takich serwerach WWW. Każdy z nich pewnie będzie też potrzebował dostępu do bazy danych, która nie powinna być dostępna z internetu. Właściwie bezpiecznie byłoby, aby taka baza nie była dostępna nawet bezpośrednio z samego serwera WWW! Powinna istnieć warstwa serwerów pośredniczących, które procesują zapytania z serwera WWW do bazy. Serwery takie nazywa się **warstwą aplikacji**. Zatem teraz musisz zatroszczyć się o bezpieczeństwo serwerów WWW, warstwy aplikacji oraz bazy danych. Ruch z niektórych albo nawet wszystkich tych serwerów powinien być balansowany, a wszystkie one powinny być zarządzane. Zarządzanie może być ciekawym wyzwaniem, ponieważ, jak napisałem, serwery bazodanowe nie powinny być dostępne z internetu.

Serwery baz danych zawierać będą dane klientów i prawdopodobnie dane kart kredytowych. Informacje te muszą być chronione. Utrzymywanie serwerów z dala od internetu jest jednym z najlepszych sposobów zapewnienia takiej ochrony.

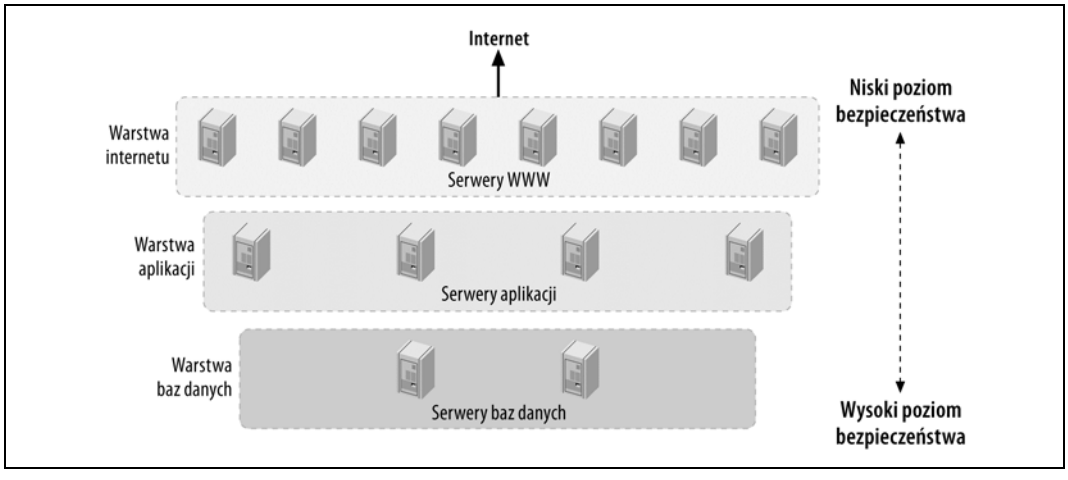
Standardowy projekt strony internetowej składa się z trzech poziomów. Pierwszy zawiera serwery WWW dostępne z internetu. Nazywa się je warstwą internetu. Drugi poziom składa się z serwerów aplikacji, które zazwyczaj nie są dostępne z internetu. Serwery te mogą komunikować się z serwerami WWW znajdującymi się ponad nimi oraz serwerami bazodanowymi pod nimi. Najniższą warstwę stanowią bazy danych, które są dostępne tylko z poziomu aplikacji.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Istnieje wiele sposobów, aby zaprojektować stronę *e-commerce*. Projekt sieci będzie zdeterminowany w dużej mierze przez zastosowane oprogramowanie i metody programistów. Pracowałem ze stronami, które miały tylko dwie warstwy, ale też z takimi, które miały cztery. Czasami programiści nalegali na zastosowanie jednej warstwy. Jeśli aplikacja nie wspierała wielu warstw, opór był marnowaniem czasu. Gdy masz styczność z serwerami przechowującymi dane kart kredytowych, upewnij się, że Twój projekt jest zgodny ze standardem bezpieczeństwa *Payment Card Industry* (PCI).

Warstwa internetu, jako bezpośrednio dostępna, jest najmniej bezpieczna. Nawet mimo tego, że chroniona jest przez zaporę sieciową, większość ludzi ma dostęp do usług, jakie oferuje. Najbardziej zabezpieczona jest warstwa baz danych. Można się do niej dostać jedynie poprzez serwery aplikacji. Serwery te mają zainstalowane dedykowane oprogramowanie, które procesuje i najczęściej buforuje informacje przesyłane z bazy do serwerów WWW.

Trójpoziomowa architektura *e-commerce* przedstawiona jest na rysunku 35.11. Warstwa internetu zawiera najwięcej serwerów. Ich liczba w kolejnych warstwach maleje, choć nie zawsze tak bywa. Na przykład warstwa baz danych może składać się z klastera Oracle korzystającego z wielu mniejszych serwerów.

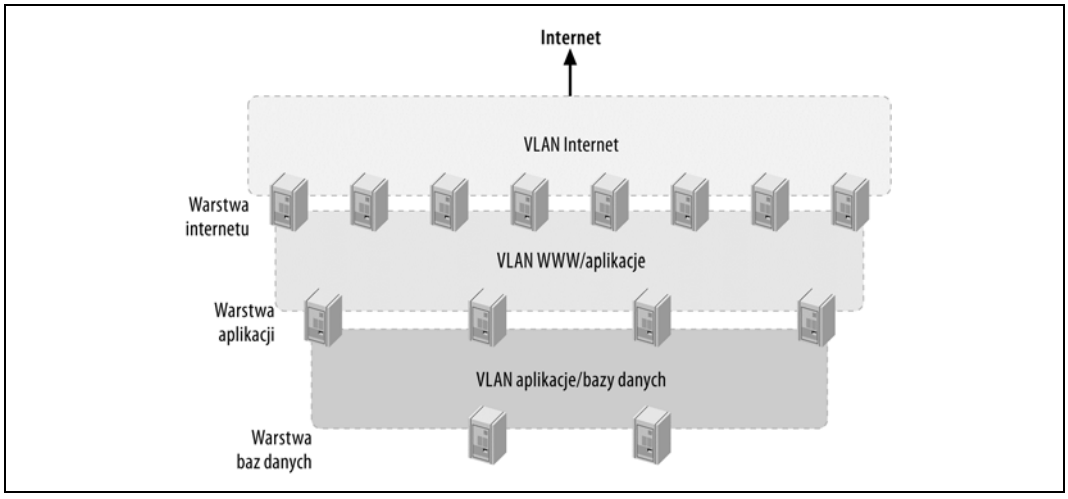


Rysunek 35.11. Typowa trójpoziomowa architektura strony e-commerce

Serwery z każdego poziomu kontaktują się z przylegającym poziomem lub poziomami. Na przykład serwery aplikacji mają interfejsy do warstwy bazodanowej oraz warstwy internetu. Wszystkie serwery muszą mieć wiele interfejsów. W przypadku sieci wysoko dostępnych serwery mogą potrzebować czterech interfejsów: dwóch dla poziomu powyżej (jeden w każdym z przełączników) i dwóch dla poziomu poniżej.

Istnieją dwie powszechnie akceptowane metody budowania takich warstw. Nazwę je **przełączana** i **rutowana**. W metodzie przełączanej niższe interfejsy wyższej warstwy są podłączone do tego samego VLAN-u co wyższe interfejsy warstwy poniżej. Rysunek 35.12 pokazuje przykład projektu strony *e-commerce* opartego na metodzie przełączanej. Zaletami takiego projektu są prostota i prędkość. Znajdują się tam tylko trzy VLAN-y, a brak jest routerów czy firewalli oddzielających warstwy. Wadą rozwiązania jest obniżone bezpieczeństwo, w związku z brakiem urządzeń separujących serwery znajdujące się w poszczególnych warstwach.

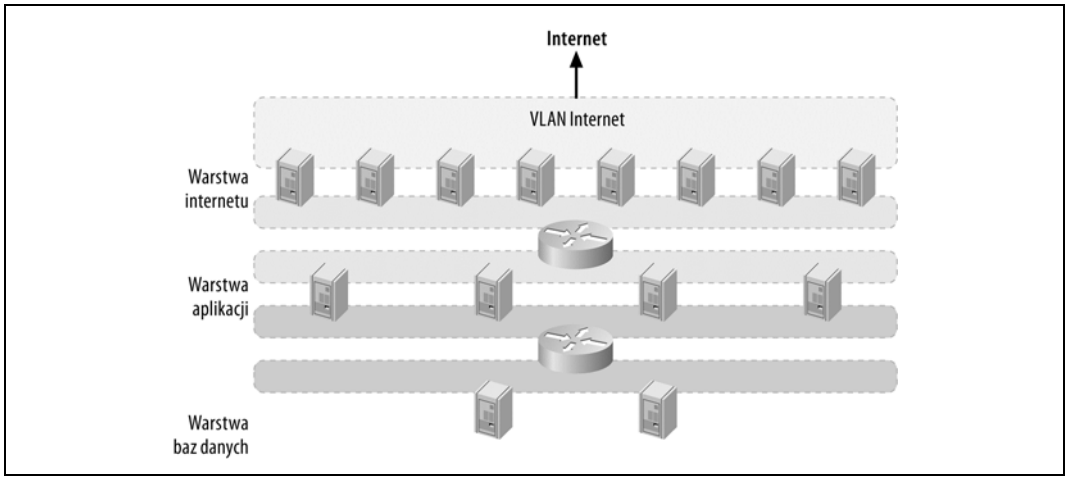
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 35.12. Trójpoziomowa sieć e-commerce typu przełączanego

Serwery w konfiguracji przełączanej powinny móc komunikować się z sieciami bezpośrednio pod nie podłączonymi. Serwery WWW są jedynymi serwerami w tym projekcie wymagającymi domyślnej bramy.

Bezpieczniejszą alternatywą dla metody przełączanej jest projekt oparty na metodzie rutowanej. Routing pomiędzy warstwami pozwala na zastosowanie tam firewalli. Rysunek 35.13 prezentuje typowy trójpoziomowy projekt rutowanej sieci *e-commerce*.



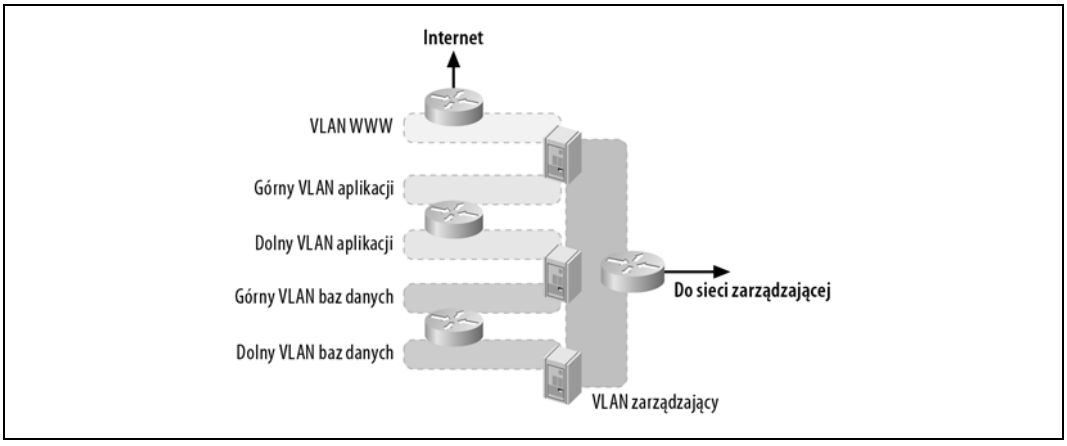
Rysunek 35.13. Trójpoziomowa sieć e-commerce typu rutowanego

Sieć rutowana jest bardziej skomplikowana, ponieważ istnieje tam prawie dwa razy więcej VLAN-ów, każdy musi mieć dowiązany SVI albo interfejs routera. W związku z fizyczną lub logiczną separacją VLAN-ów przez routery na każdym poziomie musza być zdefiniowane bramy domyślne lub wpisy konfigurujące trasy. Zaletą projektu sieci rutowanej jest bezpieczeństwo. Ponieważ cały ruch musi przechodzić przez router, może on stanowić firewall.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Problemem obu przedstawionych projektów jest trudność zdalnego zarządzania serwerami. Jeśli masz łatwy dostęp do serwerów, KVM (ang. *Keyboard Video Monitor*) umieszczony w każdej z szaf może rozwiązać problem. Witryny e-commerce zazwyczaj znajdują się jednak w dzierżawionych przestrzeniach centrum danych. W takim przypadku zdolność zdalnego zarządzania odgrywa zasadniczą rolę.

Zdalne zarządzanie jest często realizowane przy użyciu dedykowanego VLAN-u, który podłączony jest do każdego serwera oraz urządzenia sieciowego. Zarządzający VLAN-em łączy się do routera lub zapory sieciowej, co zapewnia komunikację z lokalizacją firmy. Taki projekt przedstawia rysunek 35.14.



Rysunek 35.14. Sieć zarządzająca stroną e-commerce

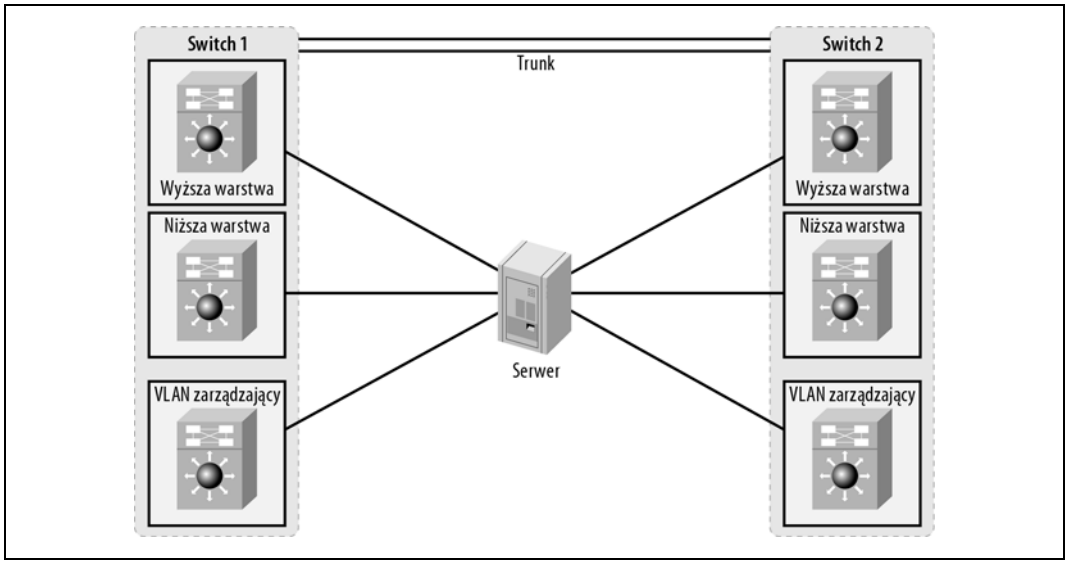
Przeszedłeś przez piekło, aby zapewnić bezpieczeństwo swoim serwerom. Odseparowałeś serwery, wykorzystując do tego warstwy, i zabroniłeś komunikacji serwerom WWW z serwerami baz danych. Tymczasem teraz dodałeś VLAN łączący wszystkie serwery. Jeśli uważasz, że to zły pomysł, masz rację. Korzyści z możliwości zdalnego zarządzania przeważają nad powstałym ryzykiem, ale tylko wtedy, gdy odpowiednio zaprojektujesz sieć.

Sieci zarządzające powinny być bardzo bezpieczne. Funkcjonalności takie jak zabezpieczenie portów (ang. *port security*) oraz prywatne VLAN-y pomagają w zapewnieniu takiego bezpieczeństwa. Żaden z serwerów nie powinien być w stanie komunikować się z innym serwerem za pomocą VLAN-u zarządzającego, wyjątkiem są serwery kopii zapasowych oraz serwery do monitoringu.

Problem związany z tymi wszystkimi VLAN-ami polega na tym, że projekt zakłada wysoko dostępną sieć. Do każdego VLAN-u, do którego musi być podłączony serwer, prowadzić powinny dwa interfejsy sieciowe: jeden do każdego z przełączników tworzących parę wysokiej dostępności. To może brzmieć jak nic wielkiego, ale już trzy VLAN-y wymagają sześciu interfejsów. Rysunek 35.15 prezentuje typowy serwer podpięty do trzech VLAN-ów na dwóch przełącznikach.

Projektując sieć jak w tym przykładzie, zadbaj o to, aby blisko współpracować z ludźmi zajmującymi się systemami. Jest wiele rzeczy, o których należy pomyśleć. Oto niektóre z nich:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

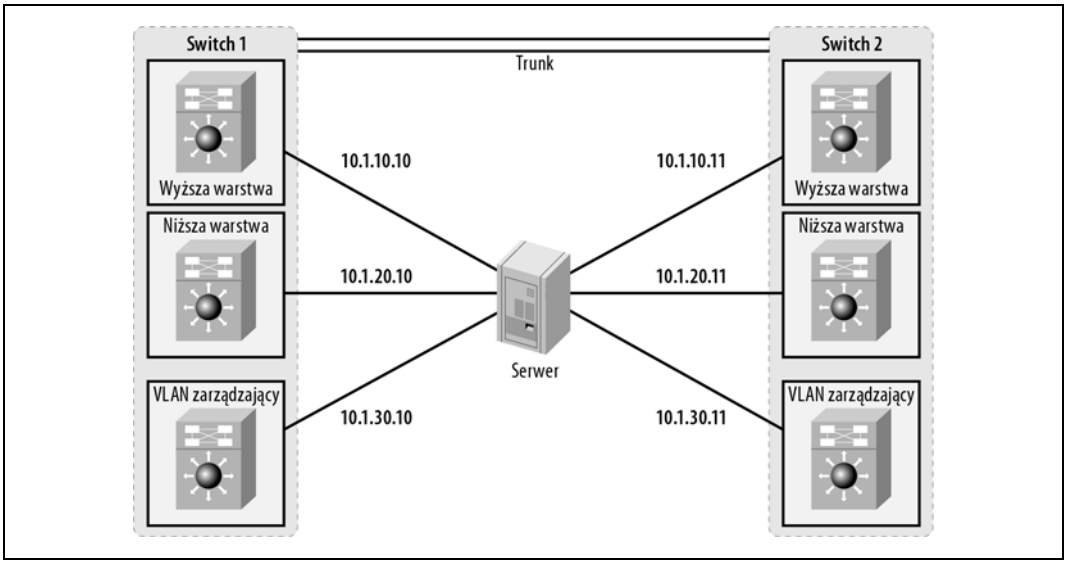


Rysunek 35.15. Interfejsy Ethernet użyte w serwerze

- Każdy interfejs będzie potrzebował adresu IP.
- W przypadku niektórych rozwiązań wysokiej dostępności będziesz potrzebował trzeciego adresu IP dla każdego z VLAN-ów. Na przykład *IP multipathing* wymaga wirtualnego adresu IP dla każdego VLAN-u na każdej z kart fizycznych.
- Każdy adres IP przypisany przez Ciebie może wymagać wpisu w DNS (włącznie z wirtualnymi adresami IP).
- Który z interfejsów jest podstawowy?
- Czy serwer potrzebuje bramy domyślnej? Jeśli tak, gdzie ma prowadzić? Czy serwer może wspierać wiele takich bram? Jak to będzie działało? Serwery WWW potrzebują bramy domyślnej wskazującej na internet. W związku z tym VLAN zarządzający musi mieć specyficzną konfigurację tras na tych serwerach.
- Jak dużo fizycznych kart musi znajdować się w serwerze, aby obsługiwał on sześć interfejsów Ethernet. Zadbaj o to, aby ich liczba była wystarczająca. Dodatkowe interfejsy są mile widziane.
- Czy serwery będą miały dwa aktywne interfejsy czy tylko jeden? Niektóre serwerowe rozwiązania wysokiej dostępności wymagają niestandardowej konfiguracji przełączników, podczas gdy inne nie. Przećwicz to w laboratorium, zanim zbudujesz swoją sieć.
- Czy Twoje serwery będą wspierały zdalne konsole Ethernet? Czy potrzebujesz dedykowanej sieci dla tego ruchu?

Kiedy pracujesz nad swoim schematem IP, dobrym pomysłem jest ustawić ostatni oktet taki sam dla każdego interfejsu na każdym z serwerów. Inaczej mówiąc, jeśli Twoja wyższa sieć to `10.1.10.0/24`, a niższa `10.1.20.0/24`, dopasuj ostatni oktet dla każdego serwera w każdej z sieci. Wtedy wyższym adresem IP będzie `10.1.10.10`, a niższym `10.1.20.10`. Pamiętaj, że musisz przypisać adres IP dla każdego z interfejsów, więc dopasuj ostatni oktet do każdego przełącznika. Rysunek 35.16 pokazuje, jak taki schemat może wyglądać.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 35.16. Dopasowanie ostatniego oktetu na wielu VLAN-ach i przełącznikach

Nowoczesne środowiska serwerów wirtualnych

Wiele centrów danych odeszło od serwerów typu *rack*, wykorzystując zamiast nich serwery kasetowe umieszczane w dedykowanych obudowach. Takie instalacje zazwyczaj pozwalają na agregację połączeń sieciowych w jednym linku 10 Gb. To może zmniejszyć potrzebną liczbę portów, jednak w zamian będziesz musiał zapewnić sprzęt sieciowy zdolny obsłużyć 10 Gb.

Dodatkowo serwery wirtualne stały się niezwykle popularne. W połączeniu z serwerami kasetowymi mogą skutkować oszałamiającymi wymaganiami związanymi z adresacją IP. Zamiana fizycznych kart sieciowych na wirtualne (te z serwerów wirtualnych) powoduje, że czasem bardzo kusząca może być dla administratora możliwość dodania nowej karty sieciowej w serwerze, co nie wiąże się wtedy z żadnymi zmianami w tradycyjnym sprzęcie. Projektując sieci dla tej nowej architektury serwerów, pamiętaj o potrzebach związanych z portami 10 Gb, zazwyczaj skonfigurowanymi jako trunki. Aby wycisnąć maksymalnie dużo z linków, rozważ agregację ruchu między obudowami — VSS lub Nexus vPC. Zwróć uwagę na QoS, ponieważ niektóre obudowy serwerów kasetowych potrafią same dzielić pasmo.

Małe sieci

Wiele małych firm nie potrzebuje skomplikowanych trójpoziomowych sieci. Mają jedno biuro, a nawet jeśli trzy, to są one małe, a sieci bardzo proste. Również niektóre większe firmy nie korzystają ze złożonych sieci. Niezależnie jednak od wielkości czy złożoności każdy element sieci powinien zostać gruntownie udokumentowany.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 36.

Projektowanie sieci IP

Zysk z dobrze zaprojektowanej sieci lub grupy sieci może być znaczący. Dokładne określenie tego zysku może być jednak dość trudne, jako że liczony jest on w godzinach w przyszłości **nieprzepracowanych**. Uwierz mi, że poprawne zaprojektowanie przestrzeni IP za pierwszym razem pozwoli Ci zaoszczędzić dosłownie tysiące godzin w kontekście czasu życia Twojej sieci.

Alokacja adresów IP jest dość rzadko wykonywana poprawnie, co kończy się bałaganem w sieciach. W wielu przypadkach małe sieci, budowane w momencie, kiedy nie było wiadomo, jak firma będzie się rozrastać w przyszłości, muszą zostać przebudowane, co generuje ogromne przedsięwzięcia. Nawet najlepsze schematy IP mogą zostać porozrywane w wyniku przejęć i fuzji.

Pomyśl, ile czasu zajmuje Ci przydzielenie serwerowi adresu IP, maski podsieci oraz bramy domyślnej. Zapewne krótko, jeśli robisz to podczas instalacji serwera. A teraz pomyśl, ile pracy wymaga zmiana tych parametrów. Serwer musi przejść w tryb *offline*, co z reguły wymaga zgłoszenia formalnej zmiany. W wielu środowiskach taka zmiana musi być wykonywana podczas dedykowanych okien czasowych, co może skutkować koniecznością przyjścia przez Ciebie do pracy o godzinie 2:00 w nocy. A teraz wyobraź sobie, że Twoja firma posiada 100, 200 albo nawet 1000 serwerów. Nie zapomnij też o tym, że serwery DNS lub/i WINS oraz Active Directory powinny również zostać zaktualizowane.

Zaprojektowanie zakresu adresów IP nie stanowi wspaniałego projektu. Nie jest to coś, za co na spotkaniu zostaniesz poklepany po plecach przez dyrektora IT. Zaprojektowanie schematu adresacji IP nie jest czymś, co doceni wielu ludzi, jeśli im tego dobrze nie wyjaśnisz. Mimo wszystko ostrożne planowanie oraz projektowanie są niezbędne, aby stworzyć schemat IP pozwalający na rozwój w nadchodzących latach.

Publiczne a prywatne zakresy IP

Jeżeli podchodziłeś do egzaminów certyfikacyjnych, powinieneś znać różnice między adresacją publiczną i prywatną. W skrócie: istnieje wybrana liczba sieci IP, które nie są rutowane w internecie. Sieci te opisane są w dokumencie RFC 1918 (*Address Allocation for Private Internets*) w poniższy sposób:

3. Prywatna przestrzeń adresowa

   Internet Assigned Numbers Authority (IANA) zarezerwowała następujące trzy bloki zakresu adresów IP na potrzeby prywatnych sieci:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

10.0.0.0 – 10.255.255.255 (10/8 prefiks)

172.16.0.0 – 172.31.255.255 (172.16/12 prefiks)

192.168.0.0 – 192.168.255.255 (192.168/16 prefiks)

Należałoby zerknąć na każdy z tych zakresów. Pierwsza sieć, `172.16.0.0`, nie ma zakresu `/16` ani też nie jest to sieć `172.16.0.0/8`. Zakres dla tej sieci to `172.16.0.0/12`. Zakres prywatnych adresów IP w tym przypadku to przedział `172.16.0.0 – 172.31.255.255`. Przedziały `172.0.0.0 – 172.15.255.255` i `172.32.0.0 – 172.255.255.255` nie wchodzą w skład adresów prywatnych i nie powinny być jako takie traktowane. Jest to typowy błąd występujący często w praktyce.

Istnieją inne zakresy również zarezerwowane, których nie powinno się używać. Popularnym przykładem jest wykorzystanie sieci `169.254.0.0/16` wtedy, kiedy DHCP nie odpowiada. Dokument RFC 3330 (*Special Use of IPv4 Addresses*) opisuje takie sieci. Zapis dla `169.254.0.0` jest następujący:

> `169.254.0.0/16` — jest to blok adresów lokalnych dla łącza (ang. *link-local*) i służy do komunikacji między hostami na pojedynczym fizycznym łączu. Hosty otrzymują ten adres automatycznie, właśnie wtedy, kiedy serwer DHCP jest niedostępny.

W ciągu dwóch lat spotkałem niejedną firmę, która cały swój schemat sieci opierała na sieci `128.0.0.0`, dlatego że łatwo było operować jej postacią binarną (`10000000.00000000.00000000.00000000`). Jest to bardzo niebezpieczna praktyka. Raz jeszcze dokument RFC 3330 dostarcza nam szczegółów:

> `128.0.0.0/16` — blok ten, odpowiadający najniższym adresom w klasie, był od początku zarezerwowany, i nadal jest, dla organizacji IANA. Biorąc pod uwagę aktualną bezklasowość w obszarze przestrzeni adresów IP, nie ma podstaw do rezerwowania tego bloku. Tematem rozmów jest przekazanie tej adresacji do RIR (ang. *Regional Internet Registry*) i przydzielanie tych adresów w normalnym trybie.

Podobnie nie powinieneś nigdy używać sieci `127.0.0.0/8`, projektując swoją sieć. Chociaż wielu ludzi wie, że adres `127.0.0.1` jest adresem *loopback*, to wielu też nie jest świadomych tego, że cała sieć jest zarezerwowana. RFC 3330 wyjaśnia:

> `127.0.0.0/8` — blok ten jest przypisany do tzw. adresów *Internet hosts loopback*. Datagram wysyłany przez protokół wyższego rzędu hosta na jakikolwiek adres z tego bloku powinien trafić zwrotnie do tego hosta i nie wyjść poza niego. Jest to zwykle zaimplementowane przy użyciu `127.0.0.1/32` dla *loopback*, ale żadne adresy z omawianego bloku nie powinny wystąpić gdziekolwiek w sieci.

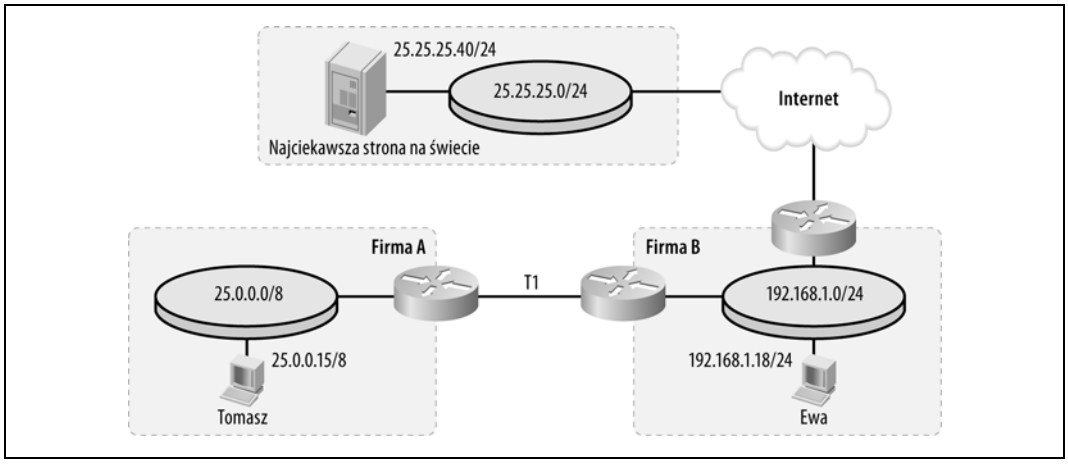
Wiele firm projektuje swoje sieci IP, kompletnie lekceważąc reguły, często używając IP odnoszących się do adresów, numerów budynków albo numerów oddziałów. Ryzyko realizowania takich projektów związane jest z łącznością i regułami routingu. Jeśli host w niewłaściwie ponumerowanej sieci `15.1.1.1/24` próbuje dostać się do strony WWW na serwerze o takiej samej sieci i masce, reguły routingu wskażą, że jest on w tej samej sieci. W takim przypadku pakiety nigdy nie zostaną skierowane do bramy domyślnej. Niektórzy twierdzą, że zastosowanie firewalla z obsługą NAT jest rozwiązaniem dla źle zaprojektowanych sieci IP, nie zważając na to, że jeśli host nigdy nie wyśle pakietu do bramy domyślnej, to nigdy nie zostanie poddany translacji NAT.

Dodatkowo jeśli firmy rozpoczynają działalność partnerską, bardzo często chcą połączyć swoje sieci. Mądra firma nie połączy swojej sieci z siecią firmy, w której reguły związane z IP zostały złamane, ponieważ odziedziczy wtedy trasy do niewłaściwie wykorzystywanych zakresów IP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Pamiętaj, że jeśli występuje trasa do sieci docelowej w tabeli routingu, brama domyślna nie pełni żadnej funkcji. Użytkownicy Twojej sieci nie będą w stanie dotrzeć do usług w sieci publicznej, która została wykorzystana wewnątrz.

Rysunek 36.1 prezentuje przykład niewłaściwie użytej sieci publicznej. W tej sieci najciekawsza na świecie strona legitymuje się adresem `25.25.25.40` z maską podsieci `255.255.255.0`. Tomasz z firmy A niewłaściwie użył adresu IP `25.0.0.15`, stosując maskę podsieci `255.0.0.0`. Ważne jest w tym przykładzie to, że żaden z pracowników firm A i B nie mógł dostać się do najciekawszej strony na świecie, mimo że jej maska była inna niż ta skonfigurowana w firmie.



Rysunek 36.1. Niewłaściwe zastosowanie sieci publicznej IP

Tomasz nie może dostać się do najbardziej interesującej strony na świecie, dopóki routery wykrywają, że znajduje się on w tej samej sieci co ona. Stos IP na jego komputerze będzie zakładał, że nie musi przesyłać pakietów do bramy domyślnej, ponieważ sieć docelowa pokrywa się ze źródłową. W tym przypadku komputer Tomasza będzie przesyłał pakiety do urządzenia lokalnego (jeśli istnieje w sieci lokalnej urządzenie z IP takim jak IP najciekawszej strony na świecie) albo wiązanie adresu IP z adresem MAC zakończy się niepowodzeniem. Niezależnie od tego Tomasz nigdy nie zobaczy najciekawszej strony na świecie.

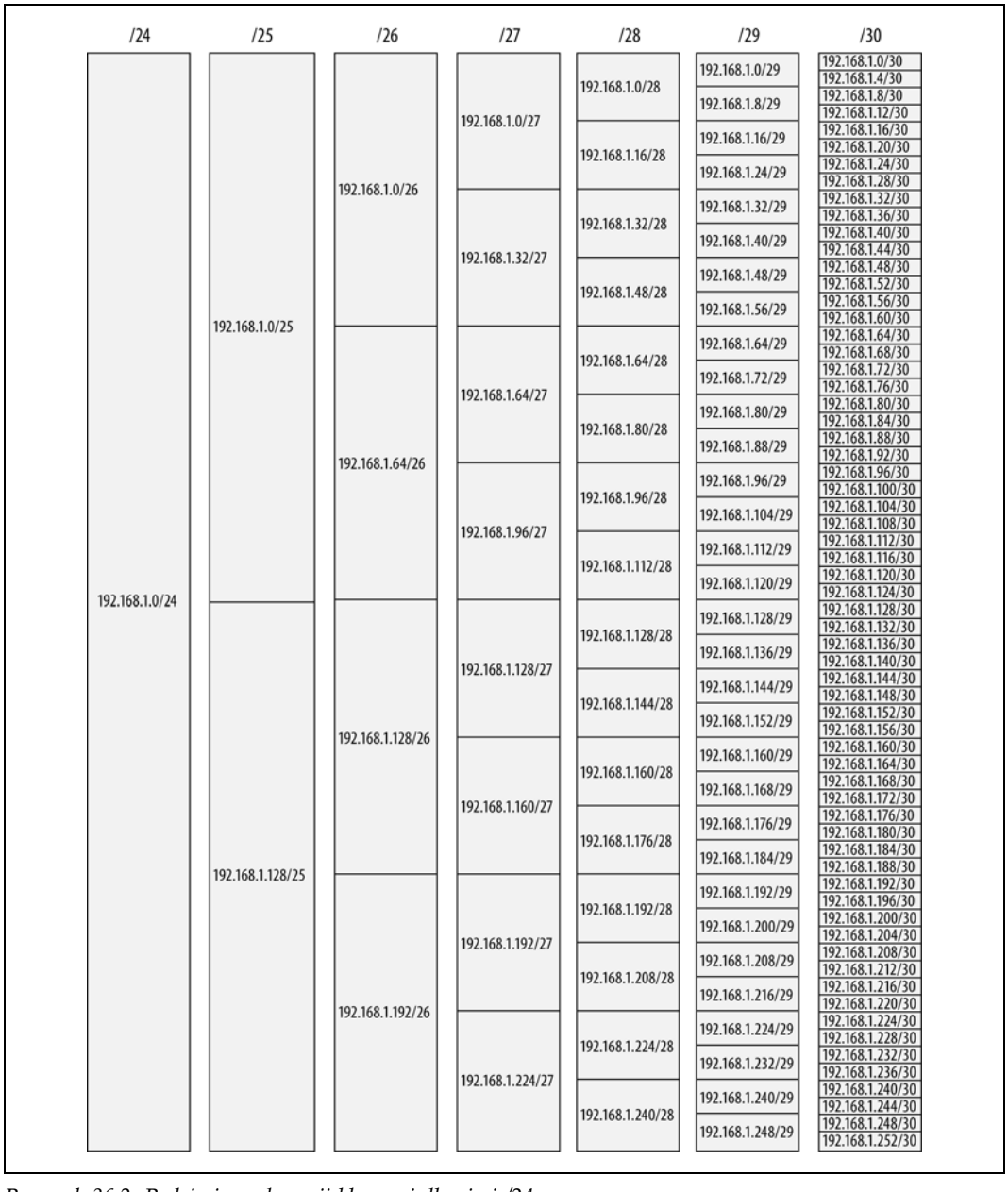
Zakładając, że firma B chciałaby, aby jej pracownicy mieli dostęp do urządzeń firmy A (z jakiego innego powodu miałaby łączyć sieci?), router łączący firmę B do internetu musi mieć trasę do sieci firmy A — w tym przypadku `25.0.0.0 255.0.0.0`. W tabeli routingu znajduje się trasa, wobec tego pakiety nie trafią do bramy domyślnej, a tym samym nie trafią do najciekawszej strony na świecie, tylko do firmy A. Przez to również Ewa nie będzie mogła dostać się do najciekawszej strony na świecie.

VLSM

Maski o zmiennej długości (ang. *Variable Length Subnet Masking*), w skrócie VLSM, to najlepsze, co spotkało sieci IP od czasu masek podsieci. W skrócie: zgodnie z tym, co będziemy określać jako **zasady tworzenia podsieci w adresowaniu klasowym** (ang. *classful subnetting rules*), jeśli sieć jest podzielona na podsieci, każda podsieć z sieci opartej na klasach (ang. *classful network*)

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

musi być tego samego rozmiaru. Rysunek 36.2 pokazuje, w jaki sposób sieć w klasie C może zostać podzielona. Są to jedyne możliwe podziały zgodnie z regułami związanymi z podziałem na klasy. Powodem tego jest binarny charakter maski podsieci.



Rysunek 36.2. Podsieci w adresacji klasowej dla sieci /24

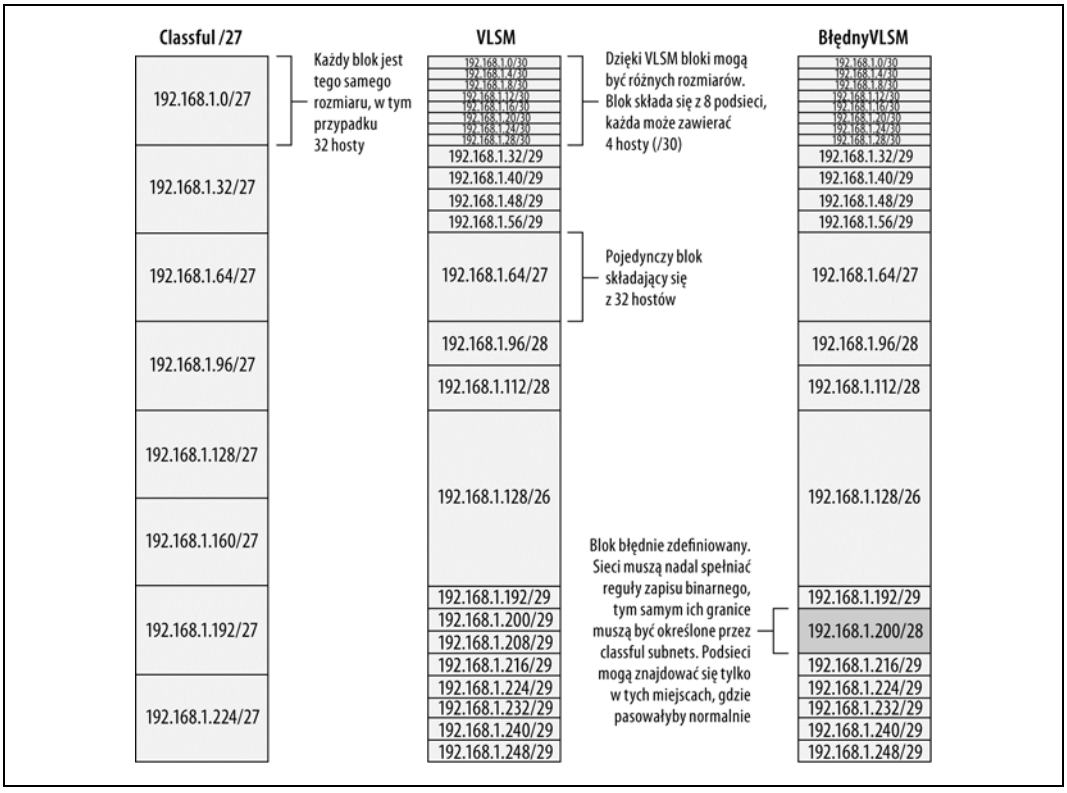
Kiedy wynaleziono sieć IP, dostępnych było więcej adresów, aniżeli ktoś mógłby kiedykolwiek zapragnąć; więcej, niż myślelibyśmy, że ktoś kiedyś będzie potrzebował. W miarę rozwoju cywilizacyjnego jednak zauważyliśmy, że w alarmującym tempie adresów przybywa.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kilka światłych umysłów wymyśliło, że zmiana reguł związanych z podsieciami da nam jeszcze trochę czasu. W rezultacie tego powstał VLSM.

VLSM określa, że jeśli sieć podzielona jest na podsieci, to **nie** muszą być one takich samych rozmiarów. Zastrzeżenie jest jednak takie, że muszą one nadal podlegać regułom zapisu binarnego, to znaczy muszą istnieć tam, gdzie byłyby zaklasyfikowane zgodnie z tradycyjnymi zasadami.

Rysunek 36.3 prezentuje przykład sieci podzielonej tradycyjnie na podsieci, właściwej sieci VLSM oraz nieprawidłowej sieci VLSM. W przykładzie nieprawidłowej sieci VLSM sieć `192.168.1.200/28` jest niedozwolona w związku z regułami podziału na podsieci. Pamiętaj, że podsieć musi znajdować się tam, gdzie pasowałaby normalnie, jeśli podzieliłbyś na podsieci całą sieć tą maską podsieci. Jeśli skonfigurowałbyś adres IP jako `192.168.1.200/28`, adres sieci wynosiłby `192.168.1.192`, a adres rozgłoszeniowy `192.168.1.207`.



Rysunek 36.3. Prawidłowe i nieprawidłowe podsieci VLSM

Zalety zastosowania VLSM powinny być oczywiste. Kiedyś musieliśmy zmarnować całą sieć dla stworzenia podsieci `/30` (jeśli podsieć `/30` została alokowana w sieci opartej na klasach, cała sieć musiała mieć tę samą maskę podsieci). Często 80% całej sieci było marnowane, ponieważ musiało być przydzielone dla podsieci `/30`, kiedy tylko 10% było potrzebne. Z VLSM możesz wykorzystać 25% dla podsieci `/30`, kolejne 25% dla podsieci `/27` i tak dalej. Chyba nie trzeba mówić, że VLSM miał znaczący wpływ na alokację adresów IP, zarówno w sieciach publicznych, jak i prywatnych.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

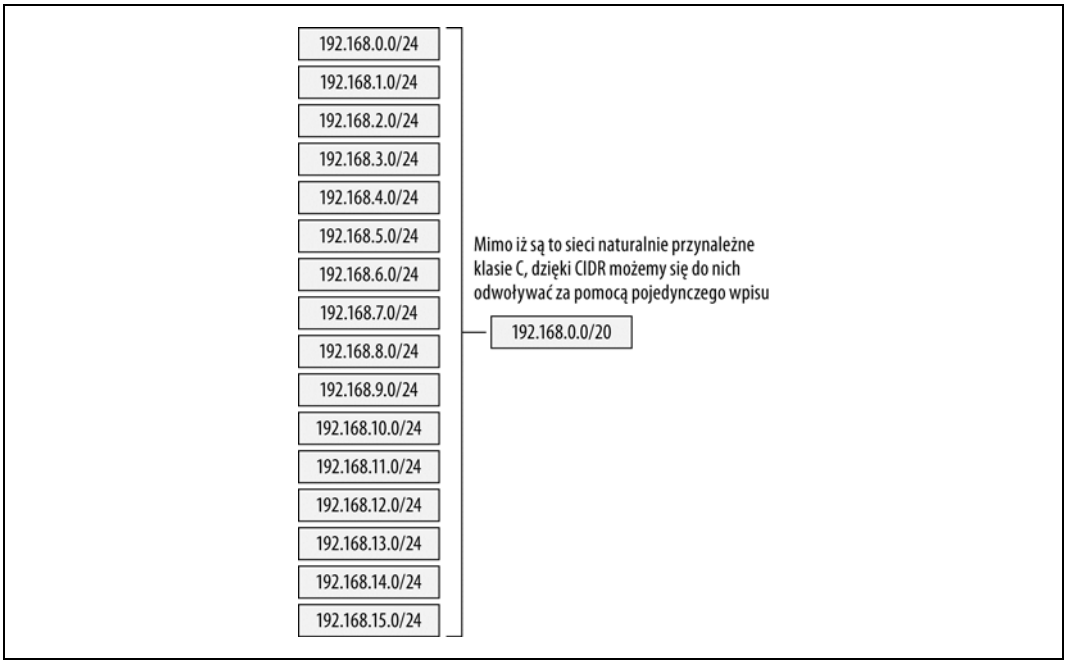
CIDR

Bezklasowy routing międzydomenowy (ang. *Classless Internet Domain Routing*), w skrócie CIDR, jest przydatnym narzędziem, jeśli operujesz dużą liczbą sieci IP — na przykład u dostawców usług internetowych (ISP). Podczas gdy VLSM miał ogromny wpływ na alokację przestrzeni adresów IP w sieciach prywatnych, CIDR wywarł równie imponujący wpływ na publiczne sieci przyznawane przez ISP.

CIDR stanowi rodzaj inwersji VLSM. Tak jak VLSM określa zasady podziału sieci, tak CIDR opisuje zasady odnoszenia się do grup sieci na podstawie pojedynczego zapisu trasy w tablicy routingu.

Agregacja tras może wydawać się rozwiązaniem problematycznym jeśli spotykałeś się jedynie z sieciami małych lub średnich firm — dla reszty agregacja przynosi rzeczywiste korzyści. Zazwyczaj małe i średnie przedsiębiorstwa używają jednej z prywatnych sieci opisanych w dokumencie RFC 1918. Jeżeli firma używa sieci `10.0.0.0/8`, dzieląc ją tak, aby maksymalizować wydajność, każda z tych części technicznie jest podsiecią. Podczas gdy VLSM dotyczy podsieci, CIDR odnosi się do sieci, które możemy określić jako **główna** (ang. *major*) lub **klasowa** (ang. *classful*).

Rysunek 36.4 prezentuje, jak pojedynczy zapis trasy (ang. *single route statement*) może odnosić się do 16 sieci klasy C. Trasa ta określana jest jako **zagregowana** (ang. *aggregate route*) lub też **zsumaryzowana** (ang. *summary route*).



Rysunek 36.4. CIDR — agregacja tras

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Routery podłączone bezpośrednio do internetu, z **pełnymi tablicami** routingu, zawierają setki tysięcy tras. Można to zaobserwować, podłączając się do jednego z wielu publicznie dostępnych serwerów tras. Aby to zilustrować, oto liczba tras na serwerze *AT&T IP Services Route Monitor* odczytana podczas pisania tej książki:

```
route-server>sho ip bgp summary
BGP router identifier 12.0.1.28, local AS number 65000
BGP table version is 75630579, main routing table version 75630579
324578 network entries using 39273938 bytes of memory
5975892 path entries using 310746384 bytes of memory
362530/55439 BGP path/bestpath attribute entries using 50754200 bytes of memory
76457 BGP AS-PATH entries using 1996260 bytes of memory
156 BGP community entries using 3744 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 402774526 total bytes of memory
Dampening enabled. 1280 history paths, 1747 dampened paths
BGP activity 1828537/1503956 prefixes, 50259731/44283836 paths, scan interval 60 secs
```

Podczas pisania pierwszej edycji tej książki liczba wpisów dla sieci (ang. *network entries*) na tym serwerze wynosiła 182 347. Blisko 4 lata później liczba ta wynosi 324 578. Rozsądne wydaje się stwierdzenie, że każdy router podłączony do internetu, który otrzymuje pełne tablice routingu, będzie zawierał podobną liczbę wpisów dla sieci. Zwróć uwagę na czwartą linię:

```
324578 network entries using 39273938 bytes of memory
```

Wpisy te zajmują 39 273 938 bajtów pamięci na routerze, przy czym cztery lata temu zajmowały 18 417 047 bajtów. Co ważniejsze jednak, każda zmiana na ścieżce do każdej z tych sieci powoduje powiadomienie wszystkich routerów na świecie podłączonych do BGP.

CIDR przyjął się dlatego, że powoduje zmniejszenie tych ogromnych tabel. Pomaga również mniejszym sieciom (a dokładniej ich administratorom) tworzyć mniejsze, lokalne tabele routingu. Są one bardziej logiczne, łatwiejsze do zrozumienia i prościej jest też wtedy rozwiązywać pojawiające się problemy.

Przyznawanie przestrzeni adresowej w sieciach IP

Jednym z powtarzających się problemów, który obserwuję od wielu lat, jest niewłaściwa alokacja przestrzeni dla sieci IP. Niezależnie, czy problem leży w adresach publicznych bezprawnie używanych, czy w braku zrozumienia korzyści wynikających z odpowiedniego przydzielania adresów, źle zaprojektowana sieć IP powoduje codzienne bóle głowy oraz zwiększa czas potrzebny na rozwiązywanie problemów. Nieprawidłowo zaprojektowany schemat może skutkować miesiącami korekt. Problem może pogarszać się z czasem, ponieważ nikt nie chce podjąć się zadania zmiany adresów IP.

Oto kilka prostych zasad, które mogą pomóc, kiedy będziesz projektował sieć, używając przestrzeni adresowej IP. Pierwsza zasada:

> Kiedy przydzielasz zakres adresów IP, staraj się przyznawać blok, do którego możesz odnieść się jednym wpisem na liście dostępu (ang. *access list entry*).

Postępowanie zgodnie z tą prostą zasadą spowoduje, że wszystko, od przyznania adresu serwerowi do zaprojektowania globalnego schematu sieci, stanie się o wiele prostsze.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Załóżmy, że zostałeś poproszony o stworzenie sieci dla nowej farmy serwerów. Nie będzie ona zawierała więcej niż 30 serwerów. Ty, jako bystry administrator, pytasz, co to będą za serwery. Odpowiedź brzmi, że będą to serwery Oracle.

Przeciętny administrator stwierdziłby po krótkim namyśle, że sieć `/27` (`255.255.255.254`) będzie rozwiązaniem, ponieważ daje ona 32 adresy dla hostów. Dwa z nich są zajęte na adres sieci oraz *broadcast*, ale nie zostawić żadnej przestrzeni na rozwój? To prowadzi nas do kolejnej reguły:

> Zawsze przyznawaj większy zakres adresów IP, niż jest to wymagane.

Jeżeli Twój szef żąda 30 adresów IP, daj mu blok 64 adresów, jeśli prosi o 60, daj mu 128. Na pierwszy rzut oka może się to wydawać marnotrawstwem, lepiej jest jednak posiadać nadmiar. Musisz nauczyć się balansować między ewentualną potrzebą rozbudowy, a ogólną dostępnością adresów. Głupotą jest przypisanie 256 adresów przy tradycyjnym trendzie wzrostowym, wtedy kiedy potrzeba ich 30. Zakładanie wzrostu o 30% wydaje się dobrą praktyką. Zaokrąglanie w górę, dopóki nie osiągnie się granic binarnych (16, 32, 64 itd.), czyni tę regułę dość łatwą do zaimplementowania.

Aby nieco urozmaicić, załóżmy, że w Twojej sieci istnieje farma serwerów. Jej adres sieci to `10.100.100.0/24`. Aktualnie jest tam 10 serwerów oraz port na routerze łączący sieć z resztą firmy. Załóżmy, że przypisano adresy od 1 do 11 (ludzie lubią przyznawać adresy w pewnym porządku).

Jeśli zostaniemy poproszeni o kolejne 30 adresów, dlaczego nie wykorzystać puli 12 – 42? Możesz to zrobić i wszyscy będą zadowoleni. To jest jednak ten moment, w którym powinieneś być bystrzejszy od przeciętnego administratora. Zamiast przyznawać przypadkową listę adresów, bazując na ostatnim przyznanym, alokuj blok z wykorzystaniem naszej pierwszej zasady: **kiedy przydzielasz zakres adresów IP, staraj się przyznawać blok, do którego możesz odnieść się jednym wpisem na liście dostępu.** Gdy przyznasz zakres 32 – 63, nie tylko przydzielisz odpowiednią liczbę adresów, ale dodatkowo możesz odnieść się do przyznanego zakresu w następujący sposób:

```
Access-list 101 permit ip any 10.100.100.32 255.255.255.224 eq web
```

A co z naszą drugą zasadą? **Zawsze przyznawaj większy zakres adresów IP, niż jest to wymagane.** Ponieważ poproszono o 30 adresów, pomyśl „do przodu" i przydziel 64 adresy. Musimy coś zmienić, dlatego że zgodnie z regułami dla podsieci nie możesz przyznać ciągu kolejnych 64 adresów, poczynając od 32. Spoglądając raz jeszcze na rysunek 36.3, możesz zauważyć, że będziesz musiał rozpocząć od wielokrotności 64. Przydziel więc 64 – 127:

```
Access-list 101 permit ip any 10.100.100.64 255.255.255.192 eq web
```

Nadal tylko jeden wpis na liście ACL!

Teraz możesz stworzyć reguły bezpieczeństwa, bazując na wcześniej nieistniejących logicznych grupach. Porównaj stworzoną przez nas listę ACL z listą ACL obejmującą adresy przydzielone z zakresu od 12 do 42. Potrzebowałbyś wszystkich linii poniżej, aby osiągnąć to, co robi jedna z nich:

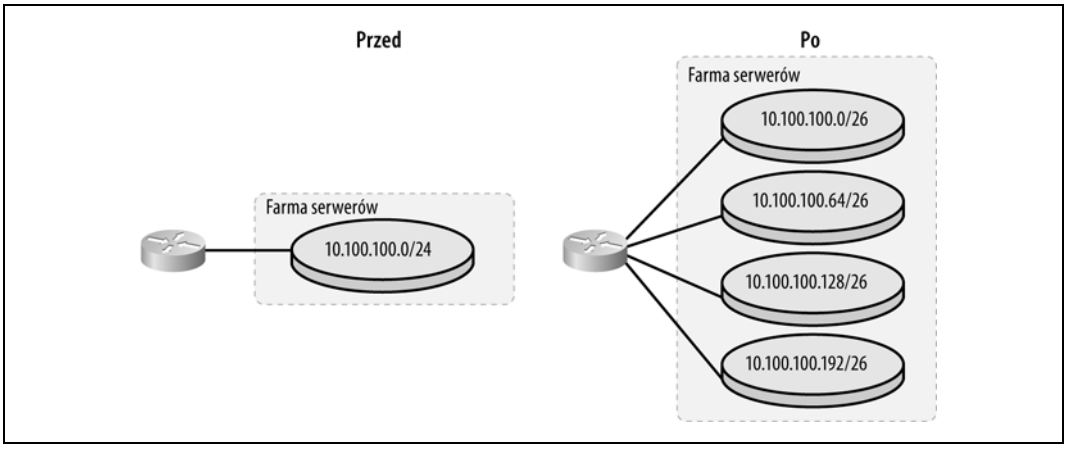
```
Access-list 101 permit ip and 10.100.100.12 255.255.255.252 eq web
Access-list 101 permit ip and 10.100.100.16 255.255.255.248 eq web
Access-list 101 permit ip and 10.100.100.32 255.255.255.250 eq web
Access-list 101 permit ip and 10.100.100.40 255.255.255.254 eq web
Access-list 101 permit ip and 10.100.100.42 255.255.255.255 eq web
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przydzieliłeś dwa razy więcej adresów, aniżeli wynosi aktualne zapotrzebowanie, i możesz się do nich odnieść listą dostępu pięć razy krótszą.

Zastosowanie powyżej opisanych zasad w stosunku do grupy podobnych serwerów ma jeszcze jedną zaletę: bez zmiany schematu IP można je usunąć z sieci i umieścić w dedykowanych podsieciach. Wymaga to nieco planowania, ale często bardzo się opłaca.

Rysunek 36.5 prezentuje, jak można zrealizować ten pomysł. Dzieląc sieć `10.100.100.0/24` na logiczne grupy odpowiadające granicom podsieci, możesz później dzielić sieć na podsieci bez konieczności zmiany IP serwerów. Będziesz musiał zmodyfikować oczywiście maski podsieci oraz bramy domyślne.



Rysunek 36.5. Podział na podsieci istniejącego zakresu

Przedsięwzięcia takie mają miejsce częściej, niż myślisz. Może Twój szef zażyczy sobie, żeby wszystkie serwery Oracle z przyczyn bezpieczeństwa miały dedykowaną sieć. Nie ma problemu! Ponieważ byłeś na tyle bystry, że przydzieliłeś zakresy IP zgodnie z granicami podsieci, cały zakres może zostać przeniesiony na inny interfejs. Zmieni się tylko maska podsieci, a wszystkie adresy IP pozostaną takie same.

Podstęp polega na tym (prawda, że zawsze jest jakiś podstęp?), aby być czujnym. Wystarczy, że jeden adres zostanie przydzielony niezgodnie z regułami, a przeniesienie istniejącej sieci będzie wymagało przenumerowania.

Przydzielanie podsieci IP

Kiedy tworzysz podsieci w sieciach IP, zadbaj o przydzielanie adresów IP w logiczny sposób. Powinieneś:

- Zadbać o jak największą pozostałą przestrzeń adresową (np. największe możliwe podsieci w pozostałej przestrzeni).
- Pozwolić tak wielu podsieciom, jak to tylko możliwe, by były rozszerzane o przyległą przestrzeń adresową.

Sam zobaczysz, że zastosowanie się do dwóch powyższych zaleceń będzie próbą pogodzenia sprzeczności.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Spotkałem się z trzema metodami przydzielania adresów IP. Będę nazywał te metody w sposób następujący: **sekwencyjną** (ang. *sequential*), **dzielenia na pół** (ang. *divide by half*) oraz **odwrotnie binarną** (ang. *reverse binary*).

Metoda sekwencyjna

Większość ludzi ma skłonność do przydzielania adresów IP oraz podsieci IP w porządku numerycznym. Oznacza to, że przydzielają numerycznie pierwszą podsieć, potem kolejną i tak dalej. Jeśli taka osoba będzie dzieliła sieć `192.168.100.0/24` na podsieci tak, aby uzyskać maskę `/30`, najprawdopodobniej przydzieli kolejno: 0, 4, 8, 12 itd. Administratorzy najczęściej mają skłonności właśnie do sekwencyjnego przydzielania adresów. Działa, jest zrozumiałe i łatwo wskazać kolejną sieć.

Z trzech metod, które tutaj omówię, ta jest najmniej pożądana, ale najczęściej stosowana. Istnieje kilka poważnych wad tej metody. Po pierwsze, w żadnej z podsieci, oprócz prawdopodobnie ostatniej, nie ma miejsca na rozbudowę. Jeśli sieć `192.168.100.16` rozwinie się i będzie potrzebowała 18 adresów zamiast 16, cała podsieć musi zostać przenumerowana do podsieci o większej przestrzeni adresowej. Po drugie, metoda ta nie pozwala zaadresować największego dostępnego bloku. Przyszła dostępność podsieci jest kolejnym problemem. Pamiętaj, że dzięki VLSM nie jesteś ograniczony tym, iż każda podsieć musi mieć ten sam rozmiar. Dokładnie planując, możesz bardzo efektywnie wykorzystać dedykowaną Ci sieć.

Rysunek 36.6 prezentuje, jak możemy sekwencyjnie alokować serię podsieci `/28` z sieci `192.168.`
`↪1.0/24`. Przydzielamy kolejną dostępną za każdym razem, kiedy potrzebujemy podsieci `/28`. Podzieliłem podsieci tak, aby pokazać największą możliwą przestrzeń adresową po każdej alokacji.

W pierwszej edycji tej książki stwierdziłem, że sieć rozgłoszeniowa jest nieużyteczna. Trudno przezwyciężyć stare przyzwyczajenia, szczególnie w przypadku takich starych gamoni jak ja. Nadal boli mnie korzystanie z ostatniej sieci, ale wskazuje to raczej na to, jak długo tego nie robiłem, niż na cokolwiek innego. Dokumentacja Cisco stwierdza: (*http://www.cisco.com/en/US/tech/tk648/tk361/technologies_tech_note09186a0080093f18.shtml*):

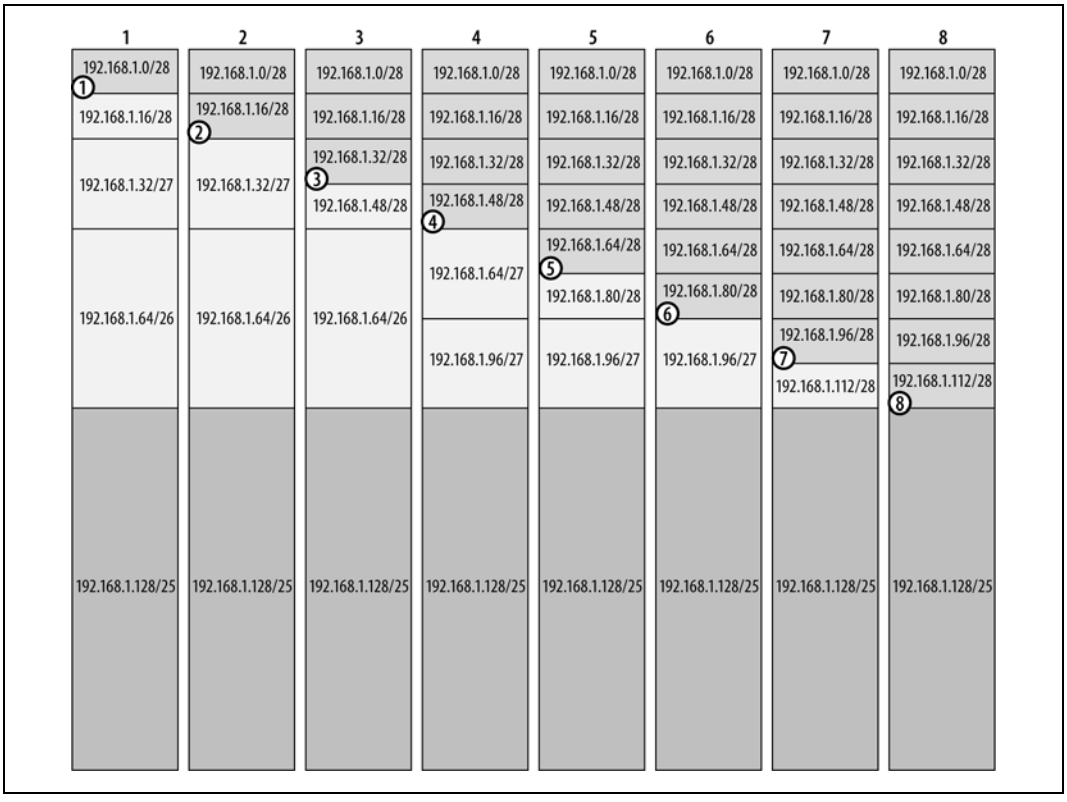
> W kwestii użycia podsieci 0 (ang. *zero subnet*) oraz podsieci z samymi jedynkami (ang. *all-ones subnet*) RFC 1878 stwierdza: „Praktyka ta (wyłączenia podsieci 0 oraz sieci z samymi jedynkami) jest zdezaktualizowana. Współczesne oprogramowanie będzie w stanie wykorzystać wszystkie zdefiniowane podsieci". Aktualnie wykorzystanie podsieci 0 oraz podsieci z samymi jedynkami jest generalnie akceptowane i większość dostawców wspiera ich zastosowanie. Jednak w określonych sieciach, szczególnie tych używających starszego oprogramowania, użycie tych podsieci może prowadzić do wystąpienia problemów.

Metoda dzielenia na pół

Metoda ta zaleca podział najmniejszego dostępnego kawałka adresowego na pół za każdym razem, kiedy należy przydzielić nową sieć. Pomysł polega na utrzymywaniu największych możliwych bloków adresowalnej przestrzeni. W tym celu przydzielamy najpierw środkową podsieć, pozostawiając dostępny spory blok adresów `/25`.

Rysunek 36.7 pokazuje zastosowanie metody dzielenia na pół. W tym przykładzie ponownie skorzystałem z sieci `192.168.1.0/24`, którą podzieliłem na podsieci `/28`. Zastosowanie tej metody pozwala na poszerzanie każdej podsieci o przestrzeń, jaka pozostała po zużytej podsieci

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 36.6. Sekwencyjne przydzielanie podsieci

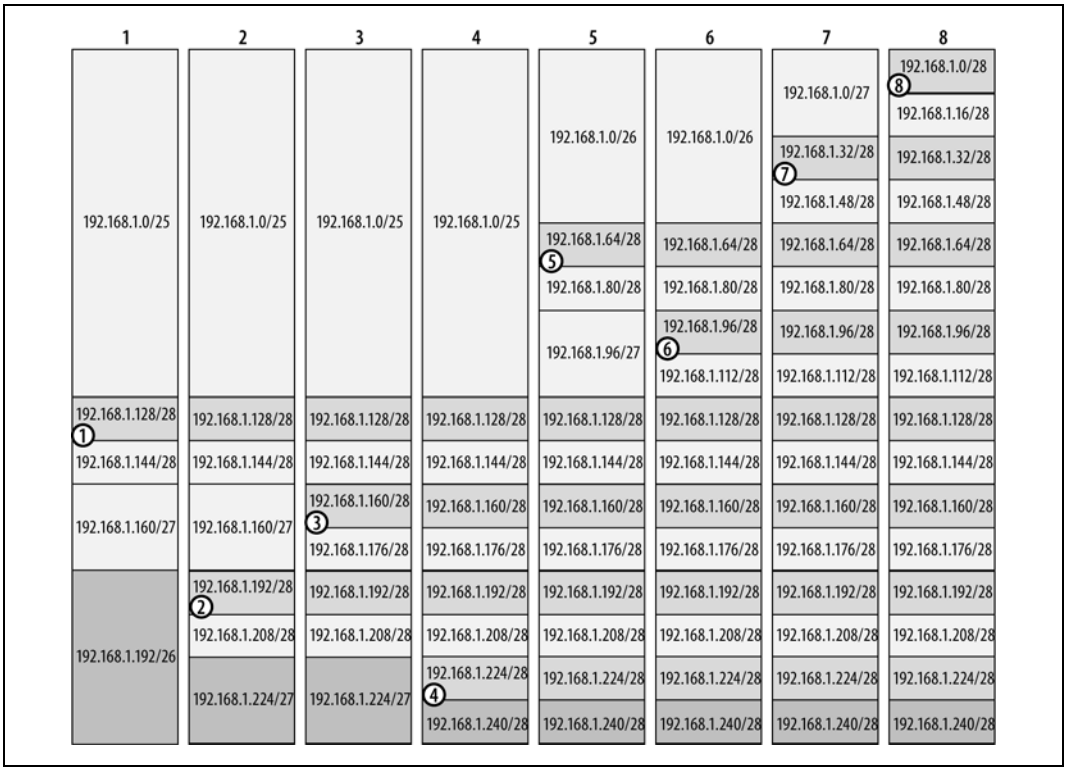
równej rozmiarowi podsieci, której użyto. Metoda ta pozwala również na przydzielenie w dowolnym momencie największej dostępnej podsieci. Lubię tę metodę, ponieważ jest prosta i w sposób rozsądny zaspokaja potrzebę rozbudowy podsieci przy utrzymaniu, tak długo, jak tylko się da, dostępu do największej możliwej przestrzeni adresowej.

Metoda odwrotna binarna

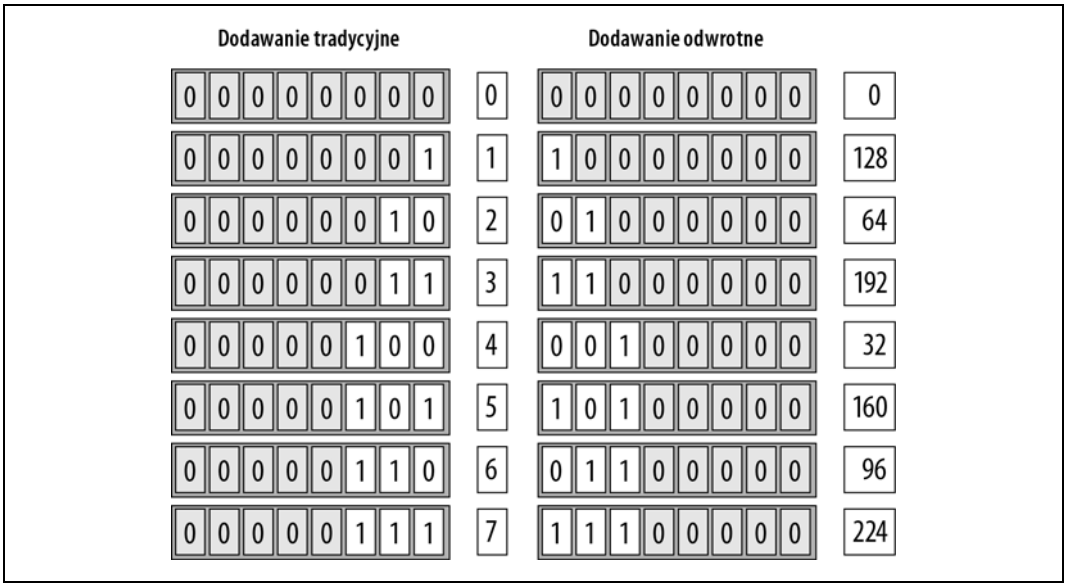
Metoda odwrotna binarna (ang. *reverse binary*) jest rekomendowanym przez Cisco sposobem przydzielania adresów podsieci IP. Metoda ta polega na przydzielaniu adresów podsieci na podstawie binarnych rachunków przy założeniu zamiany najbardziej i najmniej znaczących bitów. Rysunek 36.8 pokazuje, jak metoda ta działa. Poprzez odwrócenie bitu najbardziej znaczącego tworzymy lustrzany obraz liczb zapisanych binarnie (prawa strona). Liczby te odpowiadają podsieciom, które powinny być użyte niezależnie od rozmiaru. Może to być kontynuowane tak długo, jak długo jest to potrzebne.

Poprzez przydzielanie podsieci tą metodą docelowo rozrzucamy sieci w taki sposób, że żadne dwie podsieci nie będą do siebie przylegały, chyba że będą one różnych rozmiarów (bardzo prawdopodobne w przypadku VLSM). Algorytm ten oferuje również rozsądny kompromis, pozwalając na alokację największej dostępnej przestrzeni adresowej przy zapewnieniu możliwości rozwoju sieci. Prawdziwą różnicą między tą metodą a metodą dzielenia na pół jest rozmiar największego pozostałego bloku źródłowej sieci IP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 36.7. Przyznawanie podsieci metodą dzielenia na pół

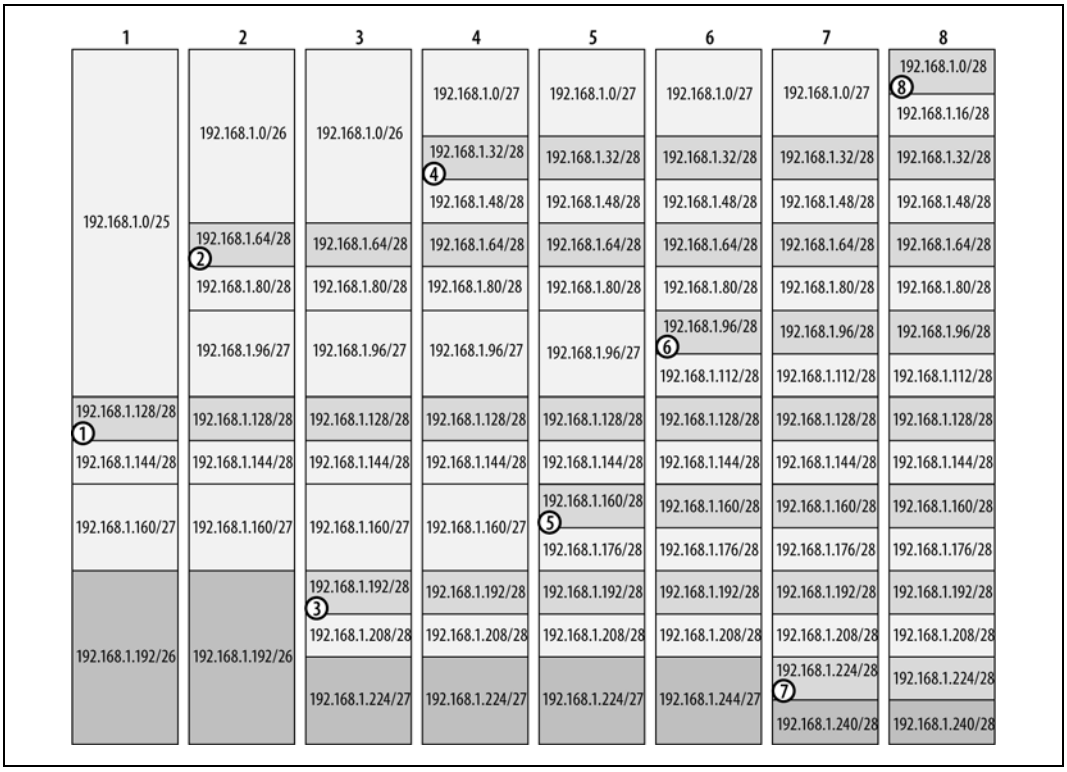


Rysunek 36.8. Przyznawanie podsieci metodą odwrotną binarną

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Wyjątkiem w powyższych regułach są dla mnie sieci `/30`. Nigdy nie zostaną one rozszerzone, ponieważ w większości dedykowane są połączeniom typu punkt-punkt (ang. *point-to-point*). Dla `/30` z reguły dedykuje się całą sieć `/24` (lub większą, w zależności od potrzeb).

Zastosowanie metody odwrotnej binarnej przy dzieleniu podsieci `192.168.1.0/24` pokazane jest na rysunku 36.9. Każda przydzielona podsieć może zostać rozszerzona bez zmiany adresów IP. Rozsądna równowaga została osiągnięta z zachowaniem dużej puli pozostałych dostępnych adresów. Różnica między metodą odwrotną binarną i dzielenia na pół zanika przy przyznaniu ósmej podsieci, kiedy to każda podsieć ma rozmiar podsieci przyległej do niej.



Rysunek 36.9. Przyznawanie podsieci IP metodą odwrotną binarną

Przyznawanie adresów metodą odwrotną binarną z matematycznego punktu widzenia jest najlepszym wyborem, jednak może być trudne do zrozumienia. Niezależnie od algorytmu, jakiego używasz, w celu uniknięcia chaosu utrzymuj zawsze odpowiednią dokumentację.

Łatwe podsieciowanie

Z podsieciowaniem (ang. *IP subnetting*) w świecie związanym z sieciami problem ma dość sporo osób. Znam wielu doświadczonych konsultantów pracujących w branży od ponad 15 lat i nadal uciekających się do kalkulatorów IP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Podsieciowanie może być zniechęcającym tematem dla tych, którzy nie mieli z nim wcześniej do czynienia. Zasady podsieciowania opierają się na regułach operacji binarnych (i odrobinie matematyki), jak na przykład eXclusive-OR (XOR), który dla wielu jest niepamiętanym aspektem dawno zapomnianych studiów.

Tak naprawdę są dwa momenty, w których przeciętny sieciowiec powinien wiedzieć, jak naprawdę działa matematyka związana z podsieciami: kiedy uczy się do swojego pierwszego certyfikatu z sieci (CCNA) i gdy uczy się do swojego ostatniego certyfikatu z sieci (CCIE). Właściwie nie wiem, czy potrzebujesz matematyki do zajęć CCNA, ale Cisco tego uczy. Egzamin CCIE powoduje, że z maskami sieci robisz naprawdę dziwne rzeczy. Wymagają one pełnego zrozumienia matematyki związanej z tematem. Rzeczy takie występują jednak rzadko lub w ogóle w praktycznych zastosowaniach i nie są tematem tej książki.

Gdy chcesz samodzielnie dzielić pule adresów IP na podsieci, musisz zrozumieć kilka rzeczy. Po pierwsze, Cisco, jak każdy inny większy producent, chce nauczyć Cię robić to na swój sposób. Dlaczego? Aby ich testy były trudniejsze do zdania i żeby każdy certyfikowany posługiwał się tym samym językiem.

Wydaje się, że szczególnie Cisco chce, aby ludzie myśleli w sposób, który powoduje, że trudniej przychodzi zrozumieć, o co właściwie chodzi. Tak na przykład w żargonie Cisco oryginalna klasa sieci C z maską `255.255.255.224` składa się z 6 sieci z 30 hostami w każdej z nich.

Nazwijcie mnie pedantem, ale jest to nieprawda. Maska podsieci wskazuje na 8 sieci i 32 hosty w każdej z nich. Z punktu widzenia Cisco jest jednak 6 **dostępnych** sieci i 30 **dostępnych** hostów. Mimo iż jest to prawda, uważam, że wnosi zamieszanie, szczególnie w kontekście naszej wcześniejszej dyskusji pokazującej podsieć 0 i podsieć z samymi jedynkami jako w pełni możliwe do wykorzystania.

Przyczyna, dla której sposób Cisco powoduje dezorientację, jest jasna: nie ma prostej metody, aby to udowodnić. Używając metody, którą przedstawię, będziesz miał zawsze sposób na jej udowodnienie. W tym przypadku: $8 \cdot 32 = 256$.

Wszystko, co ma coś wspólnego z podsieciowaniem, ma również coś wspólnego z liczbą 256. Właściwie to będzie nasza pierwsza zasada:

> Każde działanie będzie w stanie dać 256, a każdy wynik będzie podzielny przez 256.

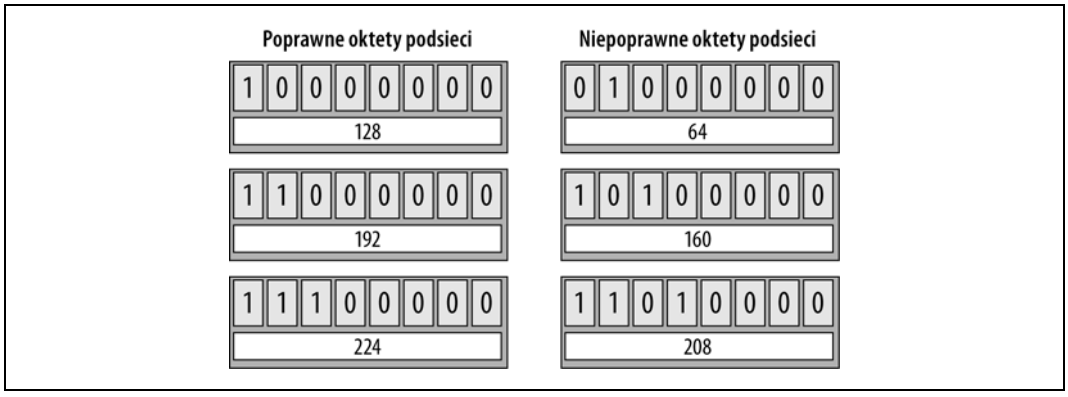
Patrząc na maskę podsieci, maksymalna wartość dla oktetu to `255`. Pamiętaj, że wartości zaczynają się od 0, zatem `255` jest 256. możliwym numerem. Jest to bardzo ważne, ponieważ wiele będzie później od tego zależało.

Druga reguła zadziwia wielu ludzi. Zaskoczyła również mnie, kiedy ją odkryłem, pisząc swój własny kalkulator liczący podsieci:

> W masce podsieci możliwych jest tylko dziewięć wartości dla każdego oktetu: 0, 128, 192, 224, 240, 248, 252, 254 oraz 255.

Maski podsieci mają swoistą specyfikę. Mianowicie możesz dodawać albo odejmować bity jedynie w określonym porządku i muszą być one dodawane od lewej do prawej. Rysunek 36.10 prezentuje bity maski zastosowane poprawnie oraz błędnie, pokazuje również rezultat w postaci zapisu dziesiętnego.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 36.10. Poprawne oraz niepoprawne oktety podsieci

Egzaminy CCIE wymagają wiedzy związanej z manipulowaniem maskami podsieci w nienaturalny sposób, choć istnieją niewielkie szanse, że zobaczysz coś takiego w praktyce. Możliwe jest wykorzystanie masek w mało oczywisty sposób, jednak nie powinieneś tego robić poza laboratorium testowym.

Ponieważ jest tylko 8 bitów w oktecie, jeśli ograniczysz się do dodawania i odejmowania bitów, otrzymasz skończoną liczbę wartości w kodzie binarnym. Wartości te pokazane są na rysunku 36.11.

Maska	Zapis binarny	Współczynnik
0	0000 0000	1:256
128	1000 0000	2:128
192	1100 0000	4:64
224	1110 0000	8:32
240	1111 0000	16:16
248	1111 1000	32:8
252	1111 1100	64:4
254	1111 1110	128:2
255	1111 1111	256:1

Rysunek 36.11. Możliwe wartości oktetów podsieci

Przyjrzyj się kolumnie zatytułowanej **Współczynnik**. Tego nie zobaczysz w tekstach producentów, a szkoda, ponieważ uważam, że w praktyce jest to istota masek podsieci.

Na rysunku 36.12 ułożyłem współczynniki zgodnie z tym, jak odnoszą się one do oktetów na każdej z pozycji w masce podsieci. Nie musisz tego zapamiętywać, wszystko stanie się wkrótce jasne. Przyjrzyj się, jak działa wzorzec. Współczynnik, którego nauczysz się używać dla klasy sieci C, będzie funkcjonował identycznie dla każdej innej klasy. Przesuwając współczynnik do właściwej pozycji w masce podsieci, możesz dowiedzieć się, jak powinien wyglądać oktet.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Maska	Pierwszy oktet		Drugi oktet		Trzeci oktet		Czwarty oktet	
0	1	256 * 256 * 256 * 256	256 * 1	256 * 256 * 256	256 * 256 * 1	256 * 256	256 * 256 * 256 * 1	256
128	2	128 * 256 * 256 * 256	256 * 2	128 * 256 * 256	256 * 256 * 2	128 * 256	256 * 256 * 256 * 2	128
192	4	64 * 256 * 256 * 256	256 * 4	64 * 256 * 256	256 * 256 * 4	64 * 256	256 * 256 * 256 * 4	64
224	8	32 * 256 * 256 * 256	256 * 8	32 * 256 * 256	256 * 256 * 8	32 * 256	256 * 256 * 256 * 8	32
240	16	16 * 256 * 256 * 256	256 * 16	16 * 256 * 256	256 * 256 * 16	16 * 256	256 * 256 * 256 * 16	16
248	32	8 * 256 * 256 * 256	256 * 32	8 * 256 * 256	256 * 256 * 32	8 * 256	256 * 256 * 256 * 32	8
252	64	4 * 256 * 256 * 256	256 * 64	4 * 256 * 256	256 * 256 * 64	4 * 256	256 * 256 * 256 * 64	4
254	128	2 * 256 * 256 * 256	256 * 128	2 * 256 * 256	256 * 256 * 128	2 * 256	256 * 256 * 256 * 128	2
255	256	1 * 256 * 256 * 256	256 * 256	1 * 256 * 256	256 * 256 * 256	1 * 256	256 * 256 * 256 * 256	1

Rysunek 36.12. Współczynniki oktetów podsieci

Weźmy jako przykład sieć z klasy C `192.168.1.0 255.255.255.0`: jeśli zastosujemy maskę `255.255.`
`↪255.224`, w rezultacie otrzymamy 8 podsieci i 32 hosty w każdej z nich. Spójrz na rysunek 36.11, a zobaczysz, że współczynnik dla oktetu podsieci `224` wynosi 8:32.

Współczynnik wydaje się współgrać z liczbą podsieci oraz z hostami w jednej podsieci klasy C. Jeśli spojrzysz na inne kolumny, zauważysz, że współczynnik jest ten sam, ale znajduje się na innej pozycji w zapisie.

Spójrzmy na pojedynczy przykład. Rysunek 36.13 przedstawia współczynniki dla oktetu podsieci wynoszącego `224`.

Maska	Pierwszy oktet		Drugi oktet		Trzeci oktet		Czwarty oktet	
224	8	32 * 256 * 256 * 256	256 * 8	32 * 256 * 256	256 * 256 * 8	32 * 256	256 * 256 * 256 * 8	32

Rysunek 36.13. Współczynniki dla 224

Oto, co w praktyce mówi rysunek 36.13, na przykładzie sieci `10.0.0.0`:

- 10.0.0.0 224.0.0.0 = 8 podsieci i 536 870 912 (32·256·256·256) hostów;
- 10.0.0.0 255.224.0.0 = 2048 (256·8) podsieci i 2 097 156 (32·256·256) hostów;
- 10.0.0.0 255.255.224.0 = 524 288 (256·256·8) podsieci i 8192 (32·256) hosty;
- 10.0.0.0 255.255.255.224 = 134 217 728 (256·256·256·8) podsieci i 32 hosty.

Zauważ, że w każdym przypadku gdzieś w zapisie pojawia się współczynnik 8:32.

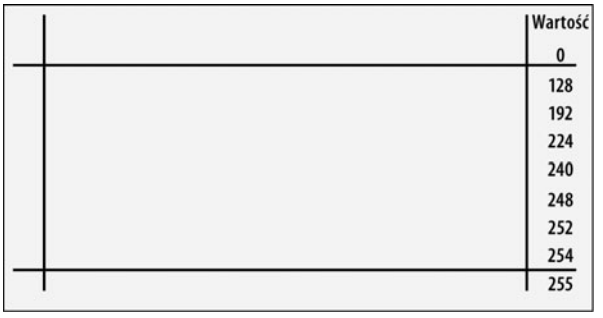
Nadal mi to jednak wygląda na ogrom paskudnej matematyki, a więc jeszcze to uprośćmy. Prowadzi nas to do trzeciej zasady:

> Jeśli znasz te zasady, jedyne, co musisz umieć robić, to dzielić razy dwa i mnożyć razy dwa.

Maski podsieci, podobnie jak wszystko, co jest związane z komputerami, bazują na systemie binarnym. Ponieważ system binarny opiera się na potęgach dwójki, wszystko, co musisz umieć robić, to podwajać liczbę lub dzielić ją przez dwa.

Przypomnij sobie dziewięć możliwych wartości dla masek podsieci (dwie są proste: 0 i 255) i zapisz je na kartce papieru (jak pokazano na rysunku 36.14). Po tym, gdy to zrobisz, wszystko, co powinieneś zapamiętać, leży przed Tobą.

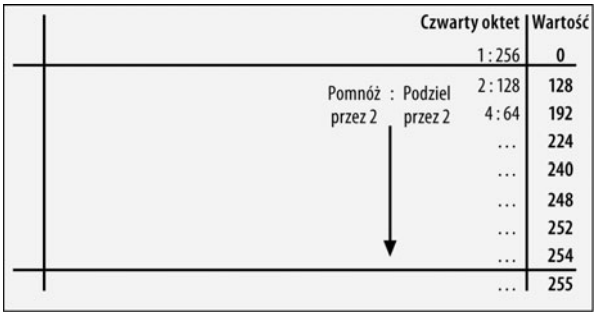
Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 36.14. Arkusz z podsieciami — krok 1.

Resztę uzupełnisz, używając najprostszych, jak to tylko możliwe, obliczeń matematycznych: dzielenia przez dwa i mnożenia razy dwa. Zapisz liczby 1:256 obok 0 na szczycie. Jest to łatwe do zapamiętania. Jak dużo naturalnych sieci jest w klasie C? 1. Ile hostów jest w klasie C? 256. Pamiętaj, że to, co się odnosi do podsieci, dąży do liczby 256.

Teraz weź liczbę na lewo od dwukropka i zwiększ ją dwa razy (1·2 = 2), następnie weź liczbę na prawo od dwukropka i podziel ją przez dwa (256/2 = 128). Powtarzaj tę czynność, dopóki nie dojedziesz do końca arkusza. Właśnie rozgryzłeś sieć — współczynniki hostów dla klasy C! Twoja kartka powinna wyglądać jak ta na rysunku 36.15.



Rysunek 36.15. Arkusz z podsieciami — krok 2.

Cała rzecz jest tak prosta, że kiedy podchodzę do intelektualnie wyczerpujących egzaminów, jak na przykład CCIE, tworzę sobie taki arkusz, tak aby nie musieć myśleć w trakcie egzaminu. Kiedy masz już wszystkie współczynniki, jesteś gotowy, aby móc pracować z klasą C, ale co z pozostałymi? Masz dwa wyjścia: nosić ze sobą zawsze tę książkę (co do mnie osobiście przemawia) albo skorzystać z tego, czego się właśnie nauczyłeś.

Jako przykładu użyjmy prywatnej sieci `172.16.0.0 255.255.0.0`, czyli `/16`. W związku z tym, że w małych firmach rzadko używa się innej podsieci niż `/24`, ludzie często czują się zakłopotani, kiedy widzą inną adresację. W rzeczywistości jednak, jeśli używasz mojej metody, nie ma to żadnego znaczenia. Podzielimy sieć na osiem kawałków. Na podstawie Twojego arkusza z podsieciami: który oktet ma liczbę 8 z lewej strony dwukropka? Zastąp zero umieszczone najbardziej na lewo w masce podsieci tym numerem, a otrzymasz swoją odpowiedź: maska podsieci, której potrzebujesz, to `255.255.224.0`. Proste.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

A teraz rozważmy bardziej złożony problem, taki, z którym możesz się spotkać na egzaminie. Jakiej maski podsieci potrzebujesz, aby w sieci umieścić 1200 hostów, mając do dyspozycji sieć `172.16.0.0 255.255.0.0`? Oto prosty sposób na rozwiązanie tego zadania. Podsieć `/24` zawiera 256 hostów. Podwajaj tę liczbę, dopóki nie otrzymasz wyniku większego niż 1200:

256·2 = 512

512·2 = 1024

1024·2 = 2048

Powstałe kolumny (rysunek 36.16.) zawierające oktety podsieci określają granice klasowe (ang. *classful boundries*). Trzecia kolumna w arkuszu wskazuje na przyszłą binarną progresję.

	Trzeci oktet	Czwarty oktet	Wartość
		1 : 256	**0**
		2 : 128	**128**
	...	4 : 64	**192**
	...	8 :32	**224**
	...	16 :16	**240**
Pomnóż przez 2	2048	32 : 8	**248**
	1024	64 : 4	**252**
	512	128 : 2	**254**
	256	256 : 1	**255**

Rysunek 36.16. Arkusz z podsieciami — krok 3.

Liczby te odnoszą się do liczby hostów w podsieci. Możesz je zapisywać tak długo, jak potrzebujesz. Kiedy dojdziesz do szczytu arkusza, rozpocznij nową kolumnę z lewej strony. Wartości te będą stanowić liczbę hostów dla Twojego drugiego i pierwszego oktetu.

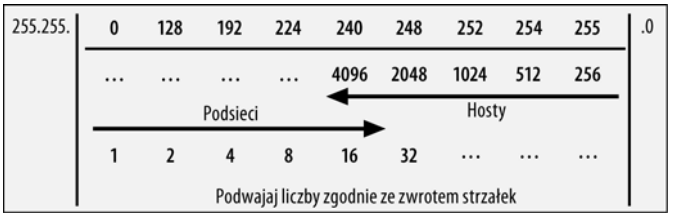
Na arkuszu widać, że wymagana maska podsieci to `255.255.248.0`. Sieć `172.16.0.0` z maską podsieci `255.255.248.0` w rezultacie daje 32 podsieci z 2048 hostami w każdej z nich. Aby to matematycznie sprawdzić, należy wykonać działanie 8·256 = 2048.

Podczas testów uważaj na podchwytliwe pytania. Korzystając z tej metody, musisz pamiętać, że pierwsza i ostatnia podsieć mogą być nieużywane zgodnie z klasycznymi zasadami podsieciowania, podobnie pierwszy i ostatni adres hosta. Dlatego jeśli spotkasz się z pytaniem: „Jaka będzie maska podsieci dla sieci, która wymaga 1024 hostów?”, odpowiedzią nie będzie 252, jak mogłoby się wydawać, ponieważ dwa z hostów są nieużyteczne (adres rozgłoszeniowy i adres sieci). Jeśli w sieci potrzeba 1024 działających urządzeń, musisz dysponować siecią zdolną do utrzymania 2048 hostów. Jest to jedno z ograniczeń zastosowania systemu binarnego.

Podobnie jeśli potrzebujesz 8 użytecznych hostów w sieci, musisz zapewnić nie 8, ale 16 adresów.

Rysunek 36.17 przedstawia to samo co poprzedni arkusz, lecz w innym formacie. Niektórzy ludzie wolą układ poziomy od pionowego. Lubię ten model, ponieważ bity ułożone są w ten sam sposób co w masce podsieci. Sam wybierz swój układ. W tej książce nie ma żadnych testów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 36.17. Poziomy układ arkusza podsieci

Pamiętaj, prawidłowa maska podsieci może mieć tylko jedną z dziewięciu wartości w każdym pojedynczym oktecie. Jeśli zaczniesz z jedynką w dolnym lewym rogu oraz wartością 256 w górnym prawym rogu, następnie podwoisz te liczby pod każdą kolejną wartością, szybko dowiesz się, jaka powinna być maska podsieci i ile hostów lub podsieci powinna mieć.

Kiedy już zapamiętałeś dziewięć wartości oktetu dla podsieci (pamiętaj, dwie są proste!) i potrafisz w pamięci mnożyć i dzielić przez 2, odkryjesz, że wszystkie te liczby są oczywiste. Będziesz je widział zawsze, kiedy będziesz choć przez chwilę w pobliżu komputera. Wszystkie są potęgą dwójki. Kiedy zaczniesz rozpoznawać schematy, które przedstawiłem, obliczenie jakiejkolwiek maski podsieci będzie kwestią szybkich i prostych obliczeń w pamięci.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 37.

IPv6

Bez wątpienia słyszałeś najnowsze doniesienia, że w najbliższym czasie wyczerpie się pula wolnych adresów IP. Wybacz moją śmiałość, ale słyszę to od ponad 10 lat.

W czasie redagowania niniejszego wydania książki NRO (ang. *Number Resource Organization*) poinformowała 3 lutego 2011 roku, że pula wolnych adresów IPv4 się wyczerpała. Możesz o tym przeczytać pod adresem: *http://www.nro.net/news/ipv4-free-pool-depleted*.

Istnieje skończona liczba adresów IP, a w początkowej fazie rozwoju internetu zarządzano nimi w bardzo zły sposób. Mechanizmy takie jak NAT, CIDR oraz VLSM zmieniły to, zwiększając dostępną pulę adresów. Było to około 10 lat temu, a aktualnie ponownie jesteśmy informowani, że pula adresów IP się wyczerpuje. Podczas gdy wszyscy panikują, ja staram się przypominać zdenerwowanym menedżerom, że problem dotyczy jedynie publicznych adresów IP. Jednak biorąc pod uwagę to, że istnieje skończona liczba tych adresów, prędzej czy później dojdziemy do pewnej granicy.

IPv4 (aktualnie dobrze znany standard) wykorzystuje 32-bitową adresację. Daje to 4 294 967 296 ($4{,}2 \cdot 10^9$) adresów IP, ale spora liczba tych adresów jest zarezerwowana lub nierutowalna. Na przykład 16 milionów z nich zawiera się w prywatnej przestrzeni adresowej `10.0.0.0/8`. Biorąc pod uwagę to, że istnieją 4 miliardy adresów publicznych, czy rzeczywiście się one kończą? Gdzie wobec tego podziały się te adresy?

Mam osiem adresów IP, pomimo że jedynie dzierżawię je od swojego dostawcy. Moja matka z kolei używa jednego. Przeciętna osoba może wykorzystywać jeden adres IP, jeśli posiada internet kablowy lub DSL. Firmy podłączone bezpośrednio do internetu mogą mieć 256 adresów, pomimo że prawdopodobnie nie używają ich wszystkich. Istnieje wiele rozwijających się krajów, które konsumują coraz to większe liczby adresów IP. A niech to! Przecież na Ziemi żyje obecnie 7 miliardów ludzi. Nie wystarczy więc dla każdego, nawet gdybym oddał swoje osiem adresów, a nie mam najmniejszego zamiaru tego robić.

Aby rozwiązać ten problem, organizacja IETF (ang. *Internet Engineering Task Force*) zaproponowała kolejną wersję IP nazwaną IPv6.

Standard IPv6 został pierwotnie opublikowany w roku 1998 wraz z RFC 2460, a według stanu na dzień wydania tej książki zmodernizowano go wraz z RFC 5095, 5722 oraz 5871. Jest wiele

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jeśli nie jesteś obeznany z notacją wykładniczą, pozwól, że Ci ją wyjaśnię, ponieważ będzie jej sporo w tym rozdziale. Notacja wykładnicza pozwala na przedstawianie ogromnie dużych liczb w łatwym do odczytania formacie. Ujmując rzecz w prosty sposób: weź liczbę i przesuwaj przecinek, dopóki nie otrzymasz liczby jednocyfrowej z jednym miejscem po przecinku. Liczba miejsc, o jakie przesunąłeś przecinek, będzie potęgą liczby 10. Liczba 1000 zostanie zapisana jako $1{,}0 \cdot 10^3$. Inaczej można o tym myśleć (choć technicznie niepoprawnie) jak o „jedynce z trzema zerami". Poniżej przytoczyłem inne przykłady, które pomogą Ci zrozumieć notację wykładniczą. Jak można zauważyć, notacja wykładnicza jest prostsza, zarówno jeśli chodzi o odczyt, jak i druk, aniżeli ciąg zer.

$1{,}0 \cdot 10^6$ = 1 000 000 (milion)

$1{,}0 \cdot 10^9$ = 1 000 000 000 (miliard)

$1{,}0 \cdot 10^{12}$ = 1 000 000 000 000 (bilion)

$1{,}0 \cdot 10^{15}$ = 1 000 000 000 000 000 (biliard)

Więcej przedstawionych w przystępny sposób informacji o nazewnictwie dla dużych liczb znajdziesz na stronach Wikipedii, w artykule zatytułowanym *Liczebniki główne potęg tysiąca*, pod adresem *http://pl.wikipedia.org/wiki/Liczebniki_g%C5%82%C3%B3wne_pot%C4%99g_tysi%C4%85ca*.

Zapytasz, dlaczego nie IPv5? Wersja IP identyfikowana jest w polu wersji pakietu IP. Wersja 4 jest naturalnie używana dla IPv4, ale wersja 5 została już wykorzystana do czegoś określanego jako Internet Stream Protocol (ST, ST+ oraz ST2), dlatego też IETF musiała przyjąć kolejną cyfrę. Jeśli jesteś zainteresowany dowiedzeniem się czegoś więcej o Internet Stream Protocol, odsyłam do dokumentu RFC 1819.

innych RFC, które obejmują IPv6 oraz zagadnienia związane z modyfikacją innych protokołów w sposób, jaki umożliwi im współpracę z IPv6. W RFC 2460 wprowadzono następujące zmiany w stosunku do IPv4:

Rozszerzone możliwości adresacji
IPv6 zwiększa rozmiar adresu IP z 32 bitów do 128, wspierając przy tym więcej poziomów hierarchii dla adresacji, dużo większą liczbę adresowalnych węzłów oraz prostszą automatyczną konfigurację adresów. Została poszerzona skalowalność routingu typu **multicast** poprzez dodanie pola zakresu dla adresów typu **multicast**. Stworzony został nowy typ adresu, nazwany adresem **anycast** (ang. *anycast address*), służący do wysyłania pakietów do dowolnego z zadanych adresatów.

Uproszczony format nagłówka
Niektóre pola nagłówka IPv4 zostały pominięte lub też przyjęte jako opcjonalne, tak aby zredukować koszt obsługi pakietu oraz ograniczyć zapotrzebowanie na przepustowość dla nagłówków IPv6.

Ulepszone wsparcie dla rozszerzeń i opcji
Zmiany w sposobie kodowania nagłówka IP pozwalają na wydajniejsze przesyłanie, mniej rygorystyczne limity dla długości opcji IP oraz znacznie większą elastyczność przy wprowadzaniu nowych opcji w przyszłości.

Możliwość etykietowania przepływu
Nowa funkcjonalność została dodana w celu umożliwienia etykietowania pakietów należących do szczególnego rodzaju ruchu, takiego, dla którego nadawca żąda specjalnej obsługi, na przykład dedykowane ustawienia *quality of service* albo *real-time service*.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Uwierzytelnianie i prywatność

Dla IPv6 zostały wyspecyfikowane rozszerzenia wspierające uwierzytelnianie, integralność danych oraz opcjonalnie ich poufność.

IPv6 wprowadziło znaczące zmiany w nagłówku warstwy IP. W celu zapoznania się z tymi zmianami sugeruję lekturę dokumentów RFC. Zamiast zanudzać Cię diagramami pakietów IP, zamierzam skupić się na tym, co powinieneś wiedzieć, aby zacząć pracę z IPv6, oraz na tym, jak należy konfigurować routery.

Adresacja

Schemat adresacji IPv6 jest nieco bardziej skomplikowany niż IPv4. Przede wszystkim IPv4 używa tylko 32 bitów adresu, IPv6 128 bitów, co równa się $3{,}40292367 \cdot 10^{38}$ adresom. Daje to w przybliżeniu 2^{95} adresów na jednego mieszkańca Ziemi w 2011 roku. Innymi słowy: dużo! Dla liczbowych maniaków napiszę, że $3{,}40 \cdot 10^{38}$ to 340 sekstylionów (*http://pl.wikipedia.org/wiki/Sekstylion*).

IPv6 zaprojektowano tak, aby zapewnić wydajną agregację (konsolidację) tras. Podsieci w IPv6 mają zawierać 2^{64} adresów. Aby właściwie to sobie wyobrazić, zauważ, że cały zakres IPv4 to zaledwie $^{1}/_{64}$ przestrzeni adresowej zawartej w jednej podsieci IPv6. Zgodnie z tym, co napisałem, pula adresów IPv6 jest naprawdę ogromna.

Adresy IPv6 nie wykorzystują liczb dziesiętnych, jak ma to miejsce w przypadku IPv4. Jest to pierwsza rzecz, z którą zmaga się wielu ludzi na początku swojej przygody z IPv6. Adresy IPv6 składają się z 4-cyfrowych grup liczb szesnastkowych (heksadecymalnych) oddzielonych dwukropkami. Szczęśliwie, wiodące zera mogą zostać wyeliminowane w każdej z grup oddzielonych dwukropkami. Na przykład:

```
AA76:0000:0000:0000:0012:A322:FE33:2267
```

może być reprezentowany jako:

```
AA76:0:0:0:12:A322:FE33:2267
```

Adresy te są identyczne, a tylko ich pisemna reprezentacja wygląda inaczej. Dodatkowo każdy nieprzerwany ciąg zer może zostać zamieniony na podwójne dwukropki. Tak więc:

```
AA76:0000:0000:0000:0012:A322:FE33:2267
```

może zostać zapisany jako:

```
AA76::12:A322:FE33:2267
```

Podsumowując, poniższe adresy są identyczne:

```
AA76:0000:0000:0000:0012:A322:FE33:2267
AA76:0:0:0:12:A322:FE33:2267
AA76::12:A322:FE33:2267
```

Jeśli zastanawiasz się, dlaczego użyto notacji szesnastkowej, porównaj przedstawione poniżej adresy, wersję heksadecymalną oraz jej dziesiętny odpowiednik:

```
HEX: AA76:0000:0000:0000:0012:A322:FE33:2267
DEC: 170.118.0.0.0.0.0.0.0.18.163.34.254.51.34.103
```

Nawet z tymi wszystkimi zerami i dwucyfrowymi oktetami dłuższym adresem będzie ten w notacji dziesiętnej. To uzmysławia, o ile więcej jest adresów IPv6 — każdy oktet oznacza 256 razy więcej adresów, a oktetów tych mamy o 12 więcej niż w przypadku IPv4.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W środowiskach, gdzie IPv6 współpracuje z IPv4, adresy IPv4 mogą być zapisywane w notacji IPv6. Możesz to zrobić poprzez ustawienie pierwszych 80 bitów na 0 oraz kolejnych 16 bitów na 1, następne bity z kolei będą tymi z adresu IPv4. Co jest jeszcze lepsze, większość implementacji pozwoli Ci na odniesienie się do adresu IPv4 w jego kropkowej notacji. I tak na przykład adres `10.20.30.40` zostanie zapisany jako:

```
0000:0000:0000:0000:0000:0000:ffff:10.20.30.40
```

albo:

```
0:0:0:0:0:0:ffff:10.20.30.40
```

albo:

```
::ffff:10.20.30.40
```

Jak sam widzisz, rzecz jest prosta, choć wymaga zrozumienia i przestudiowania. Mimo że zmiana nie będzie łatwa, niedługo nie będziemy mieli innego wyjścia. Aktualnie można już znaleźć oferty pracy, w których wymaganiem jest znajomość IPv6, więc jeśli ma to uczynić Cię atrakcyjniejszym na rynku pracy, nie jest to chyba nic złego.

Maski podsieci

Maskowanie w IPv6 jest podobne do tego używanego w IPv4, gdzie liczy się długość prefiksu. Długość maski w IPv6 może wydawać się dziwna ze względu na długość samego adresu. Adresy typu `/64` dla IPv4 wyglądają nonsensownie, ale w IPv6 spotkamy ich wiele.

To, co odróżnia IPv6, to sposób, w jaki adresy odnoszą się do masek podsieci. W IPv4 z CIDR nie istnieją klasy, dlatego też w masce podsieci może zostać wykorzystana prawie dowolna liczba bitów. W przypadku IPv6 występują następujące reguły:

Zakres adresów sieciowych
: Pierwsze 48 bitów adresu IPv6 zarezerwowane jest dla routingu internetowego:

```
ffff:ffff:ffff:0000:0000:0000:0000:0000
```

Zakres dla podsieci
: 16 bitów od numeru 49 do 64 definiuje podsieć:

```
0000:0000:0000:ffff:0000:0000:0000:0000
```

Zakres dla urządzeń
: Ostatnie 64 bity są adresem urządzenia:

```
0000:0000:0000:0000:ffff:ffff:ffff:ffff
```

Możesz skorzystać z części adresu urządzenia, aby w przyszłości wykorzystać to na potrzeby podsieci, jednak jest to niezalecane. Wkrótce dowiesz się dlaczego.

Typy adresów

Podobnie jak w IPv4 niektóre adresy są adresami specjalnymi, identyfikowanymi przez swoje najbardziej znaczące bity. Są one ujęte w dokumencie RFC 4291 (podrozdział 2.4) i przedstawione w tabeli 37.1.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Tabela 37.1. Rodzaje adresów IPv6

Typ adresu	Binarny prefiks	Notacja IPv6	Podrozdział
Nieokreślony	`00...0` (128 bitów)	`::/128`	2.5.2
Loopback	`00...1` (128 bitów)	`::1/128`	2.5.3
Multicast	`11111111`	`FF00::/8`	2.7
Lokalny dla łącza	`1111111010`	`FE80::/10`	2.5.6
Globalny	(pozostałe)		

Adres nieokreślony

Adres nieokreślony to adres składający się z samych zer. Pakiety z nieokreślonym przeznaczeniem zostaną odrzucone przez routery. Adres źródłowy może być nieokreślony dla urządzeń, które jeszcze nie nauczyły się swoich adresów.

Adres loopback

Adres *loopback* w notacji IPv6 to adres: `0:0:0:0:0:0:0:0:1` albo `::1/128`. Jest on analogiczny do tego stosowanego w IPv4 `127.0.0.1`.

Maski podsieci wymagają niezwykłej ostrożności. `/32` w notacji IPv4 wskazuje na maskę złożoną z samych jedynek lub też specyficzny adres hosta. `/32` w IPv6 znaczy zupełnie coś innego, kumuluje wiele tras w internecie. Tzw. host routes w notacji IPv6 to `/128`, a nie `/32`.

Multicast

Multicast jest to coś nieco innego w notacji IPv6, choć główna idea pozostaje ta sama. Pierwszą rzeczą, którą powinieneś odnotować, jest to, że w IPv6 nie istnieje koncepcja ruchu rozgłoszeniowego (*broadcast*). Jeśli chcesz wysłać pakiet do wszystkich nodów, powinieneś użyć adresu *multicast link-local all nodes* `ff02::1`.

Adresy globalne (ang. *global unicast*)

Adres globalny można określić jako jeden ze zwykłych adresów, który zostanie przypisany do Twojego urządzenia. Adres globalny musi być 64-bitowy, poza przypadkiem, w którym zaczyna się od `000`, tak jak to przedstawiono w tabeli 37.1.

Adresy lokalne dla łącza (ang. *link-local unicast*)

Adresy lokalne dla łącza opisane są w dokumencie RFC 2373 (podrozdział 2.5.8) w sposób następujący:

> Lokalne adresy dla łącza zostały zaprojektowane, aby umożliwić adresowanie pojedynczego łącza. Wykorzystywane jest to na potrzeby automatycznej konfiguracji, wykrywania sąsiednich hostów lub też kiedy nie mamy dostępu do żadnych routerów.
>
> Routerom nie wolno przesyłać pakietów ze źródłowym bądź docelowym adresem lokalnym dla łącza.

O co więc chodzi? Jeśli kiedykolwiek próbowałeś uruchomić system Windows i nie mógł on znaleźć serwera DHCP, prawdopodobnie zaobserwowałeś, że używa on adresu `169.254.x.x`. Adresy `169.254.1.0 – 169.254.254.255` stanowią blok adresów lokalnych (ang. *link-local adressess*), nierutowalnych w obrębie sieci internet.

Adresy lokalne dla łącza w notacji IPv6 znajdują się w sieci `fe80::/10`. IPv6 wymaga od systemu operacyjnego automatycznego przypisania adresu lokalnego do każdego z interfejsów. Format takiego adresu jest złożeniem prefiksu routingu oraz adresu MAC. Adres MAC ma

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

jednak wtedy w środku `ff:fe`, jak również `02` dodane na początku. Dlaczego? Umieszczenie w środku `ff:fe` konwertuje adres MAC na format Extended Unique Identifier 64 (EUI-64). `02` z kolei jest rezultatem rzutowania siódmego bitu na pierwszy oktet adresu MAC. Bit ten jest bitem **lokalnie administrowanym**, wskazującym, że OS zmienił adres. Poniżej znajduje się fragment z maszyny linuksowej. Wszystko, co zrobiłem, to konfiguracja adresu IPv4. Jest tam również adres lokalny dla łącza IPv6. Jak widać, końcówka adresu pasuje do adresu MAC zmodyfikowanego w sposób, jaki opisałem powyżej:

```
[gad@cozy ~]$ ifconfig eth0
eth0 Link encap:Ethernet HWaddr 00:02:55:b7:da:9d
  inet addr:192.168.1.200 Bcast:192.168.1.255 Mask:255.255.255.0
  inet6 addr: fe80::202:55ff:feb7:da9d/64 Scope:Link
  UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
  RX packets:193321787 errors:0 dropped:0 overruns:0 frame:0
  TX packets:196670319 errors:0 dropped:0 overruns:0 carrier:0
  collisions:0 txqueuelen:1000
  RX bytes:3569684877 (3.3 GB) TX bytes:3127898656 (2.9 GB)
  Interrupt:16
```

Poniżej podobny przykład z routera Cisco. Na początku adres MAC interfejsu:

```
R1#sho int f0/0 | inc Hardware
  Hardware is MV96340 Ethernet, address is 0012.43b9.2c70 (bia 0012.43b9.2c70)
```

A oto automatycznie skonfigurowany adres lokalny dla łącza:

```
R1#sho ipv6 interface f0/0 | inc link-local
  IPv6 is enabled, link-local address is FE80::212:43FF:FEB9:2C70
```

W adresie lokalnym możesz zaobserwować `FF:FE` w środku adresu MAC oraz wiodącą dwójkę będącą rezultatem rzutowania bitowego.

Podsieci

Pierwsze 48 bitów zarezerwowane jest dla routingu internetowego, dlatego też prawdopodobnie nie będziesz miał do czynienia z maskami większymi niż `/48`, o ile nie będziesz zajmował się takim właśnie ruchem. Kolejne 8 bitów to podsieć, umożliwiają one dostęp do 65 535 podsieci z około 18 trylionami urządzeń w każdej z nich.

Świetnie, ale czy potrzebuję alokować ponad 18 trylionów urządzeń dla każdego łącza typu *point-to-point*? Słowo **potrzebuję** nie jest tu może najtrafniejsze, należałoby raczej powiedzieć **powinienem**. Istnieje tutaj kilka mocnych argumentów.

Po pierwsze, zostaliśmy przyzwyczajeni do oszczędzania, jeśli chodzi o wykorzystywanie adresów IP; stało się to wręcz naszą drugą naturą. Po co używać 256 adresów IP dla łącza typu *point-to-point*, skoro potrzebujesz tylko 4? Ten rodzaj myślenia ocalił nas w przeszłości, a przyzwyczajenia, zwłaszcza te dobre, trudno przezwyciężyć. Faktem jest jednak, że mamy tak wiele adresów IPv6, że nie dojdziemy do ich granic przez naprawdę długi czas. Prędzej Skynet i terminatorzy zniszczą cywilizację, którą znamy.

Istnieją protokoły, które wymagają 64-bitowego adresu IPv6. Najczęściej przywołuje się protokół *Stateless Address Autoconfiguration*. Wymóg ten oznacza mianowicie, że maska podsieci nigdy nie powinna być mniejsza niż `/64`; właściwie wszystkie sieci IPv6 w powszechnym użyciu powinny mieć maskę podsieci o rozmiarze `/64`. ARIN (ang. *American Registry for Internet Numbers*), zgodnie z *http://www.arin.net/policy/nrpm.html#six54*, wskazało na następujący sposób przydzielania sieci w IPv6:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Użytkownicy końcowi otrzymują adresy od swoich dostawców (LIR lub ISP). W interesie dostawcy leży taki rozmiar przydziału, aby użyć najmniejszej wartości /64 (jeśli przewidujemy jedną podsieć dla lokalizacji), a maksymalnie wartości /48. W wyjątkowym przypadku niezwykle dużych lokalizacji może zostać przyznany większy przydział.

Poniższe wskazania mogą się okazać pomocne (choć należy pamiętać, że to tylko wskazania):

/64, jeśli wiemy, że potrzeba jednej i tylko jednej podsieci;

/56 dla małych lokalizacji, takich, których przyrost w najbliższych pięciu latach szacuje się na kilka podsieci;

/48 dla dużych lokalizacji.

Mimo że technicznie możliwe jest użycie paradygmatu VLSM w technologii IPv6, to jednak nie powinieneś tego robić. Najlepiej używaj wszędzie maski /64. Jeśli tego nie zrobisz, to i tak Twój lokalny adres interfejsu będzie miał maskę /64. Tak na przykład /48 daje Ci 2^{16} podsieci z maską /64. Myślę, że możesz poświęcić kilka z nich dla swojego łącza szeregowego.

Wyobraź sobie — każda podsieć IPv6 Twojej sieci będzie miała prefiks /64. Nie trzeba już rozważać, ile hostów zmieści się w danej podsieci, ponieważ może ich być aż 18 trylionów. Innymi słowy, takie rozważania nie mają już żadnego praktycznego uzasadnienia.

IETF rekomenduje zastosowanie masek /64 dla wszystkich podsieci IPv6, nawet dla łączy typu *point-to-point*. Organizacja ta zaleca również, aby każda sieć domowa stosowała maski /48. Można o tym przeczytać w dokumencie RFC 3177. Oto jego fragment:

W szczególności zalecamy:

- użytkownicy abonenci korzystający z łączy stałych jak również dostępu na żądanie powinni otrzymać maskę /48;
- małe i duże przedsiębiorstwa powinny otrzymać maskę /48;
- bardzo duzi abonenci powinni otrzymać maskę /47 lub nieco krótszy prefiks, albo też wiele podsieci w masce /48;
- sieci mobilne, np. samochody, telefony komórkowe z dodatkowym interfejsem sieciowym (takim jak na przykład bluetooth albo 802.11b) powinny otrzymać statyczny prefiks /64, tak aby możliwa była komunikacja wielu urządzeń w ramach jednej podsieci;
- pojedyncze komputery typu PC bez konieczności korzystania z podsieci, wdzwaniające się na przykład z pokoi hotelowych, powinny otrzymać swój adres /128 dla połączenia PPP jako część prefiksu /64.

Przeczytaj uważnie! IETF rekomenduje, aby każdy samochód i telefon komórkowy otrzymały maski /64! To właśnie jest efekt oszałamiająco dużej liczby dostępnych adresów IPv6. Nagle mój najnowszy smartfon z jednym nędznym adresem IPv4 okaże się przeżytkiem.

NAT

NAT okazał się wielkim wynalazkiem, który pozwolił nam oddalić moment, gdy skończą się wolne adresy IPv4. Jak stosować NAT w technologii IPv6?

Nie stosować.

Mimo że od pewnego czasu NAT jest wspierany przez IPv6, jest tak wiele adresów, że nie musimy go wykorzystywać. Ukrycie adresu IP pod innym, jak to się dzieje w przypadku mechanizmów NAT, to oczywiście pewna forma zabezpieczenia, kontrargumentem będzie jednak to, że bezpieczeństwo rozumiemy w tym przypadku jedynie jako zatuszowanie informacji. Przed dołączeniem mechanizmów NAT do specyfikacji IPv6 sporo na ten temat dyskutowano.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Istnieją funkcje NAT-PT oraz NAPT-PT, które pozwalają na translacje adresów IPv6 na IPv4 (NAPT-PT umożliwia również translacje portów), ale są one zaprojektowane do tego, aby służyć jako narzędzia do migracji z IPv4 do IPv6.

Zanim wprowadzono NAT-a, każdy otrzymywał adres publiczny. IPv6 zatem prowadzi nas do miejsca, z którego wyszliśmy. Ja osobiście preferowałbym świat bez technologii NAT.

Prosta konfiguracja routera

Posiadam dwa routery połączone za pomocą przełącznika. R1 ma nadany adres IPv4 `10.0.0.1/24` na `f0/0`, a R6 adres `10.0.0.6/24` na `f0/0`. Aby uzyskać lokalny adres IPv6 dla łącza na tych interfejsach, nie musiałem robić zupełnie niczego poza ich włączeniem (istnienie adresu IPv4 nie ma żadnego wpływu na testowanie):

```
R1(config)#int f0/0
R1(config-if)#ipv6 address autoconfig
```

Uruchomienie komendy `show interface` nie pokaże adresu IPv6:

```
R1#sho int f0/0
FastEthernet0/0 is up, line protocol is up
  Hardware is MV96340 Ethernet, address is 0012.43b9.2c70 (bia 0012.43b9.2c70)
  Internet address is 10.0.0.1/24
  MTU 1500 bytes, BW 100000 Kbit, DLY 100 usec,
    reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  Full-duplex, 100Mb/s, 100BaseTX/FX
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:00:00, output 00:00:03, output hang never
  Last clearing of "show interface" counters never
  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  5 minute input rate 3000 bits/sec, 0 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
    3403 packets input, 1970499 bytes
    Received 3393 broadcasts, 0 runts, 0 giants, 0 throttles
    0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
    0 watchdog
    0 input packets with dribble condition detected
    755 packets output, 77707 bytes, 0 underruns
    0 output errors, 0 collisions, 1 interface resets
    0 babbles, 0 late collision, 0 deferred
    0 lost carrier, 0 no carrier
    0 output buffer failures, 0 output buffers swapped out
```

Aby zobaczyć automatycznie skonfigurowany adres lokalny IPv6, należy skorzystać z polecenia `show ipv6 interface` *`nazwa_interfejsu`*:

```
R1#sho ipv6 int f0/0
FastEthernet0/0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::212:43FF:FEB9:2C70
  No global unicast address is configured
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FFB9:2C70
  MTU is 1500 bytes
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
ICMP error messages limited to one every 100 milliseconds
ICMP redirects are enabled
ND DAD is enabled, number of DAD attempts: 1
ND reachable time is 30000 milliseconds
```

A oto adres lokalny IPv6 urządzenia R6:

```
R6#sho ipv6 interface f0/0 | inc link-local
IPv6 is enabled, link-local address is FE80::214:6AFF:FEA2:D438
```

Co powinien zrobić każdy sieciowiec konfigurujący dwa urządzenia w sieci? Oczywiście skorzystać z polecenia `ping`. Spróbujmy wykonać ping z R6 do R1. Ping w IPv6 to nadal to samo polecenie, choć musimy użyć słówka `ipv6`:

```
R1#ping ipv6 FE80::214:6AFF:FEA2:D438
Output Interface: f0/0 //interfejs wyjściowy f0/0
% Invalid interface. Use full interface name without spaces (e.g. Serial0/1)
                        //Nieprawidłowy interfejs. Użyj pełnej nazwy interfejsu bez spacji (np. Serial0/1)
```

Niestety, choć jest to niewygodne, nie mamy innego wyjścia:

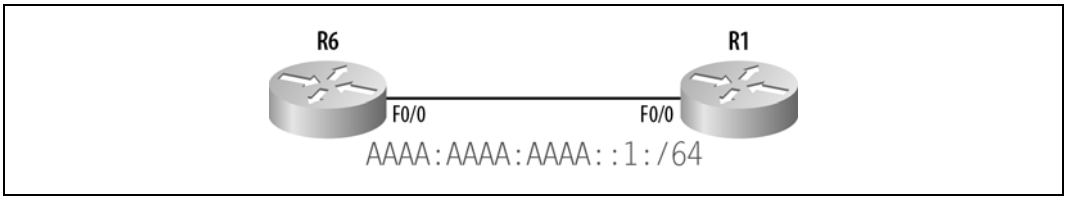
```
Output Interface: fastethernet0/0
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to FE80::214:6AFF:FEA2:D438, timeout is 2 seconds:
Packet sent with a source address of FE80::212:43FF:FEB9:2C70
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 0/0/4 ms
```

Ping działa tak, jak byśmy tego oczekiwali, dlaczego jednak musieliśmy wpisać nazwę interfejsu źródłowego? A dlatego, że sieć dla adresu IP jest taka sama na każdym interfejsie — pamiętaj, że mamy do czynienia z adresem lokalnym dla łącza, którego router nie potrafi przełączyć między interfejsami. To, co najbardziej boli, to przerażająca długość adresu IPv6. Wszyscy lubimy posługiwać się adresami IP, co wywołuje wrażenie swoistej erudycji. Jestem przekonany, że adresy IPv6 sprawią, iż docenimy i zaczniemy właściwie wykorzystywać mechanizm DNS.

No dobrze, ale pakiety z adresem lokalnym dla łącza jako zaprojektowane w celu wykorzystania w obrębie jednej sieci nie są przecież przekazywane przez routery. Jaki jest zatem w tym przypadku cel posiadania routera, skoro nie potrafi on przekazywać ruchu między interfejsami?

Celem jest właśnie przesyłanie pakietów. Jak wiemy z definicji, każdy interfejs na routerze powinien być w innej podsieci. Według mnie to właśnie dzięki temu IPv6 może stać się marzeniem każdego administratora sieci.

IPv6 pozwala Ci adresować interfejs bez przypisywania adresu urządzeniu. Odkąd możemy stworzyć adres poprzez translację adresu MAC na format EUI-64, możemy też szybko połączyć urządzenia, wskazując im wyłącznie to, w jakiej podsieci się znajdujemy. Przykład znajduje się na rysunku 37.1.



Rysunek 37.1. Prosta sieć IPv6 z dwoma routerami

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

R1 właśnie otrzymał adres lokalny dla łącza, ale jak już sobie powiedzieliśmy, będzie to bezużyteczne, jeśli chodzi o jego kontakt ze światem zewnętrznym. Poleceniem `ipv6 address` *`adres /maska`* `eui-64` przypiszę adres do interfejsu `F0/0`. Powiedzmy, że będę przypisywał adresy z bloku `AAAA:AAAA:AAAA::/48`. Zacznę od podsieci `:1::`

```
R1(config-if)#ipv6 address AAAA:AAAA:AAAA:1::/64 eui-64
```

Jest to bezpieczne w laboratorium, ale pamiętaj, że poza nim powinieneś posiadać dostarczony przez dostawcę odpowiedni prefiks `/64`, np. przez ARIN.

Teraz interfejs będzie miał adres `aaaa:aaaa:aaaa:1::/64`, a adres hosta będzie adresem w formacie EUI-64 wygenerowanym na podstawie adresu MAC. `:1:` reprezentuje podsieć, którą przypisałem z zakresu `/48 aaaa:aaaa:aaaa`. Możemy to zobaczyć, wydając komendę `show ipv6 interface` *`nazwa_interfejsu`*:

```
R1#sho ipv6 int f0/0
FastEthernet0/0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::212:43FF:FEB9:2C70
  Global unicast address(es):
    AAAA:AAAA:AAAA:1:212:43FF:FEB9:2C70, subnet is AAAA:AAAA:AAAA:1::/64 [EUI]
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FF00:0
    FF02::1:FFB9:2C70
  MTU is 1500 bytes
  ICMP error messages limited to one every 100 milliseconds
  ICMP redirects are enabled
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  ND advertised reachable time is 0 milliseconds
  ND advertised retransmit interval is 0 milliseconds
  ND router advertisements are sent every 200 seconds
  ND router advertisements live for 1800 seconds
  Hosts use stateless autoconfig for addresses.
```

Interfejs ma adres lokalny dla łącza `FE80::212:43FF:FEB9:2C70`, zdefiniowany tak ze względu na prefiks `FE80::2`, oraz globalny adres typu *unicast* `AAAA:AAAA:AAAA:1:212:43FF:FEB9:2C70`. Ostatnie 14 znaków szesnastkowych pokrywa się w obu adresach, ponieważ odnoszą się one do adresu MAC interfejsu (z dodanym w środku ciągiem `FF:FE`).

Pomyślmy nad tym przez chwilę. Wyobraźmy sobie, że wszystkie urządzenia w sieci byłyby skonfigurowane w ten właśnie sposób. Wtedy, jeśli chcielibyśmy coś przenumerować, jedyne, co musielibyśmy zrobić, to zmienić adres sieci. Nie musielibyśmy już przyznawać adresów IP, nikt nie musiałby utrzymywać stale nieaktualnych arkuszy kalkulacyjnych, nikt nie miałby też pojęcia o czymś takim jak adresy hostów. To ostatnie może nie brzmi jak zaleta, ale mimo to wierzę, że nią jest.

Maniacy tacy jak ja, którzy lubią kontrolować, wskazywać, czy router ma być zaadresowany pierwszym, czy ostatnim adresem w sieci, nie mają szczęścia. Co zatem pozostaje? To, co powinniśmy robić przez ten cały czas — stosować mechanizmy DNS. Jeśli wdrożenie IPv6 przyniosłoby ze sobą jedynie popularyzację mechanizmów DNS, uznałbym je za sukces.

Dobrze, jesteś maniakiem kontroli wszystkiego i chcesz sam skonfigurować swoje urządzenie. Oczywiście możesz to zrobić. Mogę przypisać adres hosta `1` do interfejsu:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
R1(config-if)#ipv6 address aaaa:aaaa:aaaa:1::1/64
```

Spójrzmy:

```
R1#sho ipv6 interface f0/0
FastEthernet0/0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::212:43FF:FEB9:2C70
  Global unicast address(es):
    AAAA:AAAA:AAAA:1::1, subnet is AAAA:AAAA:AAAA:1::/64 [ANY]
    AAAA:AAAA:AAAA:1:212:43FF:FEB9:2C70, subnet is AAAA:AAAA:AAAA:1::/64 [EUI]
[fragment usunięty]
```

Chwila! Dlaczego adres EUI nadal tam jest? To nie IPv4, w IPv6 możesz przypisać wiele adresów do interfejsu i nie musisz przy tym używać słówka `secondary`. Bądź jednak ostrożny, jeśli zechcesz usunąć adres EUI, musisz to zrobić ręcznie:

```
R1(config)#int f0/0
R1(config-if)#no ipv6 address AAAA:AAAA:AAAA:1::/64 eui-64
```

Jak widać poniżej, po usunięciu adresu EUI pozostał wyłącznie adres `::1`:

```
R1#sho ipv6 interface f0/0
FastEthernet0/0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::212:43FF:FEB9:2C70
  Global unicast address(es):
    AAAA:AAAA:AAAA:1::1, subnet is AAAA:AAAA:AAAA:1::/64 [ANY]
 [fragment usunięty]
```

Pamiętasz, jak uzyskałem połączenie między R1 i R6? Skonfigurowałem interfejs `f0/0` na R6 komendą `ipv6 address autoconfig`. Kiedy na niego patrzę, widzę, że stało się coś fascynującego. Na początku spójrzmy na konfigurację interfejsu, aby się przekonać, że nie zwariowałem (jak gdyby to było wystarczającym dowodem):

```
interface FastEthernet0/0
  ip address 10.0.0.6 255.255.255.0
  duplex auto
  speed auto
  ipv6 address autoconfig
```

A oto efekt wykonania komendy `show ipv6 interface` *`nazwa_interfejsu`*:

```
R6#sho ipv6 int f0/0
FastEthernet0/0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::214:6AFF:FEA2:D438
  Global unicast address(es):
    AAAA:AAAA:AAAA:1:214:6AFF:FEA2:D438, subnet is AAAA:AAAA:AAAA:1::/64 [PRE]
      valid lifetime 2591981 preferred lifetime 604781
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FFA2:D438
    MTU is 1500 bytes
    ICMP error messages limited to one every 100 milliseconds
    ICMP redirects are enabled
    ND DAD is enabled, number of DAD attempts: 1
    ND reachable time is 30000 milliseconds
    Default router is FE80::212:43FF:FEB9:2C70 on FastEthernet0/0
```

Router ten poza adresem lokalnym łącza ma również adres globalny oraz adres routera. Skąd te wszystkie informacje? R6 nauczył się ich od routera R1.

Routery wysyłają coś, co określane jest w IPv6 jako ogłoszenie routera (ang. *router advertisements*). Jest to odpowiedź na zapytanie o router (ang. *router solicitation*). Mimo że nie jest to w pełni

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

identyczne, można myśleć o tym troszkę jak o DHCP. R6 nauczył się, że prefiks sieci to `AAAA:AAAA:AAAA:1::/64`, a skoro jest automatycznie konfigurowalny, dodał swój adres EUI-64 do prefiksu i sam stworzył pełny adres IPv6.

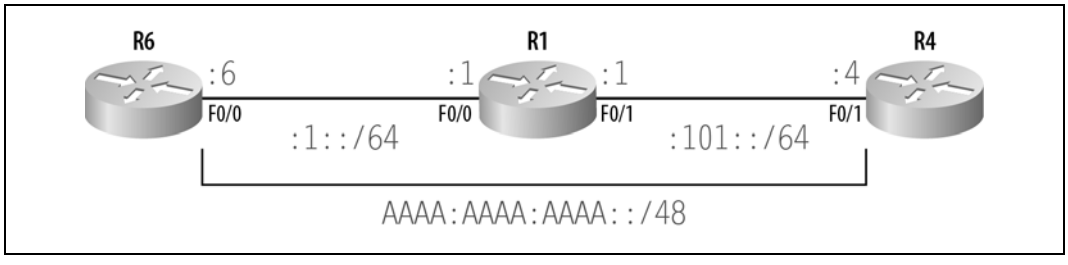
Przypiszmy jakiś łatwo odczytywalny adres IPv6. Przypiszę adres hosta `::6` do interfejsu `f0/0` na R6. Pamiętaj, że nasz prefiks to `AAAA:AAAA:AAAA::`, a podsieć `:1::`

```
R6(config)#int f0/0
R6(config-if)#ipv6 address aaaa:aaaa:aaaa:1::6/64
```

Szybki `ping` pozwoli sprawdzić, czy konfiguracja jest poprawna. Spróbujmy z adresem `::1` na R1:

```
R6#ping ipv6 aaaa:aaaa:aaaa:1::1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to AAAA:AAAA:AAAA:1::1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 0/157/788 ms
```

Sukces! Zauważ, że tym razem nie musieliśmy określać interfejsu. Stało się tak, ponieważ przypisaliśmy do interfejsu sieć typu *non-link local*. Sprawmy, aby nasza mała sieć stała się nieco bardziej interesująca. Dodam trzeci router, R4, zajmując interfejs `F0/1` na R1. Będzie to wyglądało tak jak na rysunku 37.2.



Rysunek 37.2. Sieć IPv6 złożona z 3 routerów

Na początek interfejs `f0/1` na R1, ten sam prefiks, który został użyty wcześniej, podsieć `:101`:

```
interface FastEthernet0/1
  ipv6 address AAAA:AAAA:AAAA:101::1/64
```

Teraz na R4 dodam adres hosta `::4` z tym samym prefiksem i podsiecią:

```
R4(config)#int f0/1
R4(config-if)#ipv6 address aaaa:aaaa:aaaa:101::4/64
```

Jeszcze tylko sprawdzenie za pomocą polecenia `ping`:

```
R4#ping ipv6 aaaa:aaaa:aaaa:101::1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to AAAA:AAAA:AAAA:101::1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 0/146/728 ms
```

`ping` do innych podsieci na interfejsie R1 zakończy się jednak niepowodzeniem:

```
R4#ping ipv6 aaaa:aaaa:aaaa:1::1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to AAAA:AAAA:AAAA:1::1, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Dlaczego? Czy R1 nie powinien przesyłać pakietów poprzez połączone interfejsy? Domyślnie w IPv6 nie powinien. Aby to umożliwić, w urządzeniach Cisco musisz włączyć *IPv6 unicast-routing*:

```
R1(config)#ipv6 unicast-routing
```

Kiedy już to zrobisz, ping na inne interfejsy R1 zakończy się sukcesem:

```
R4#ping ipv6 aaaa:aaaa:aaaa:1::1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to AAAA:AAAA:AAAA:1::1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 0/0/4 ms
```

Zauważ też, że na polecenie `ping` odpowiada interfejs `f0/0` na R6:

```
R4#ping ipv6 aaaa:aaaa:aaaa:1::6
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to AAAA:AAAA:AAAA:1::6, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 0/0/4 ms
```

Mimo że na R1 nie ma skonfigurowanego routingu, R1 będzie domyślnie rutował lokalnie połączone interfejsy. Zarówno R6, jak i R4 nie mają skonfigurowanych bram domyślnych ani też statycznych tras. O co więc chodzi?

Pamiętasz komunikaty ogłoszenie routera, o których pisałem wcześniej? Wskazują one na R1, określając go jako router, i tę właśnie informację wykorzystuje IPv6. Jeśli spojrzymy na efekt użycia komendy `show ipv6 interface f0/1` na urządzeniu R4 przed i po wywołaniu polecenia `ipv6 unicast-routing` na R1, zobaczymy interesującą różnicę.

Przed:

```
R4#sho ipv6 int f0/1
FastEthernet0/1 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::212:80FF:FE31:E259
  Global unicast address(es):
    AAAA:AAAA:AAAA:101::4, subnet is AAAA:AAAA:AAAA:101::/64
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FF00:4
    FF02::1:FF31:E259
  MTU is 1500 bytes
  ICMP error messages limited to one every 100 milliseconds
  ICMP redirects are enabled
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
```

Po:

```
R4#sho ipv6 int f0/1
FastEthernet0/1 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::212:80FF:FE31:E259
  Global unicast address(es):
    AAAA:AAAA:AAAA:101::4, subnet is AAAA:AAAA:AAAA:101::/64
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FF00:4
    FF02::1:FF31:E259
  MTU is 1500 bytes
  ICMP error messages limited to one every 100 milliseconds
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
  ICMP redirects are enabled
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  Default router is FE80::212:43FF:FEB9:2C71 on FastEthernet0/1
```

Interfejs pokazuje domyślny router jako adres IPv6 EUI-64 interfejsu `F0/1` na R1. Idę o zakład, że coś podobnego zaobserwujemy na interfejsie `f0/0` routera R6:

```
R6#sho ipv6 int f0/0
FastEthernet0/0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::214:6AFF:FEA2:D438
  Global unicast address(es):
    AAAA:AAAA:AAAA:1::6, subnet is AAAA:AAAA:AAAA:1::/64
    AAAA:AAAA:AAAA:1:214:6AFF:FEA2:D438, subnet is AAAA:AAAA:AAAA:1::/64 [PRE]
      valid lifetime 2591990 preferred lifetime 604790
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FF00:6
    FF02::1:FFA2:D438
  MTU is 1500 bytes
  ICMP error messages limited to one every 100 milliseconds
  ICMP redirects are enabled
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  Default router is FE80::212:43FF:FEB9:2C70 on FastEthernet0/0
```

Całkiem fajne, czyż nie? Może tak, może nie. Ze swojego doświadczenia wiem, że umożliwienie routerom budowania swoich własnych sieci jest dość niebezpieczne. Kiedy każda sieć jest w stanie rozmawiać z każdą inną, programiści zaczynają używać ścieżek, które nie powinny istnieć. Ich późniejsza blokada może spowodować problemy z aplikacjami. Co gorsza, w testach, jak na przykład w egzaminach praktycznych CCIE, wykorzystuje się tę automatyczną konfigurację interfejsów, aby utrudnić realizację zadań laboratoryjnych. Naturalne jest dla mnie pytanie, jak zapobiec takiemu zachowaniu.

Można wyłączyć komunikaty ogłoszenia routera, wydając polecenie `ipv6 nd suppress-ra` na interfejsie podstawowym:

```
R1(config-if)#int f0/0
R1(config-if)#ipv6 nd suppress-ra
R1(config-if)#int f0/1
R1(config-if)#ipv6 nd suppress-ra
```

Kiedy to zrobię, inne routery nie będą otrzymywać komunikatów ogłoszenia routera oraz uczyć się domyślnych routerów, a tym samym dynamiczna sieć zniknie.

A teraz w kroku finalnym dodajmy tradycyjny routing i domyślną bramę na R6 oraz R4.

Na początku dodam statyczny routing na R6, tak aby zezwolić na dostęp do podsieci `:101::`

```
R6(config)#ipv6 route aaaa:aaaa:aaaa:101::/64 aaaa:aaaa:aaaa:1::1
```

Na R4 dodam zamiast tego bramę domyślną. W IPv6 domyślna trasa różni się nieco od tej w IPv4, głównie sposobem, w jaki zapisywane są adresy. Kusić może zapis: `0:0:0:0::/0` pasujący do każdej trasy. Pokazuje on, że jesteś jeszcze nowicjuszem w kwestiach IPv6 (mimo że zadziała):

```
R4(config)#ipv6 route 0:0:0:0::/0 aaaa:aaaa:aaaa:101::1
```

Ekspert od IPv6 odwoła się do trasy `0/0`, używając `::/0`:

```
R4(config)#ipv6 route ::/0 aaaa:aaaa:aaaa:101::1
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Stosując polecenie `show ipv6 route`, sprawdźmy więc konfigurację tras naszych urządzeń:

R6:

```
R6#sho ipv6 route
IPv6 Routing Table - 6 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
      U - Per-user Static route
      I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary
      O - OSPF intra, OI - OSPF inter, OE1 - OSPF ext 1, OE2 - OSPF ext 2
      ON1 - OSPF NSSA ext 1, ON2 - OSPF NSSA ext 2
C AAAA:AAAA:AAAA:1::/64 [0/0]
    via ::, FastEthernet0/0
L AAAA:AAAA:AAAA:1::6/128 [0/0]
    via ::, FastEthernet0/0
L AAAA:AAAA:AAAA:1:214:6AFF:FEA2:D438/128 [0/0]
    via ::, FastEthernet0/0
S AAAA:AAAA:AAAA:101::/64 [1/0]
    via AAAA:AAAA:AAAA:1::1
L FE80::/10 [0/0]
    via ::, Null0
L FF00::/8 [0/0]
    via ::, Null0
```

R1:

```
R1#sho ipv6 route
IPv6 Routing Table - 8 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
      U - Per-user Static route
      I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary
      O - OSPF intra, OI - OSPF inter, OE1 - OSPF ext 1, OE2 - OSPF ext 2
      ON1 - OSPF NSSA ext 1, ON2 - OSPF NSSA ext 2
C AAAA:AAAA:AAAA:1::/64 [0/0]
    via ::, FastEthernet0/0
L AAAA:AAAA:AAAA:1::1/128 [0/0]
    via ::, FastEthernet0/0
L AAAA:AAAA:AAAA:1::2/128 [0/0]
   via ::, FastEthernet0/0
L AAAA:AAAA:AAAA:1:212:43FF:FEB9:2C70/128 [0/0]
    via ::, FastEthernet0/0
C AAAA:AAAA:AAAA:101::/64 [0/0]
    via ::, FastEthernet0/1
L AAAA:AAAA:AAAA:101::1/128 [0/0]
    via ::, FastEthernet0/1
L FE80::/10 [0/0]
    via ::, Null0
L FF00::/8 [0/0]
    via ::, Null0
```

R4:

```
R4#sho ipv6 route
IPv6 Routing Table - 5 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
      U - Per-user Static route
      I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary
      O - OSPF intra, OI - OSPF inter, OE1 - OSPF ext 1, OE2 - OSPF ext 2
      ON1 - OSPF NSSA ext 1, ON2 - OSPF NSSA ext 2
S ::/0 [1/0]
    via AAAA:AAAA:AAAA:101::1
C AAAA:AAAA:AAAA:101::/64 [0/0]
    via ::, FastEthernet0/1
L AAAA:AAAA:AAAA:101::4/128 [0/0]
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
    via ::, FastEthernet0/1
L FE80::/10 [0/0]
    via ::, Null0
L FF00::/8 [0/0]
    via ::, Null0
```

R6 wykorzystuje statyczny routing, R1 ma trasy bezpośrednio połączone, a R4 domyślnie wskazuje na R1. Wszystko działa świetnie:

```
R4#ping ipv6 aaaa:aaaa:aaaa:1::6
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to AAAA:AAAA:AAAA:1::6, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 0/1/4 ms
```

IPv6 to temat, o którym można przeczytać dużo więcej. Mam nadzieję jednak, że ten rozdział pozwoli rozwiać wiele obaw związanych z przyszłością. IPv6 nie jest naprawdę niczym trudnym. Mimo złożonej budowy adresów nieco praktyki pozwoli Ci się nimi biegle posługiwać.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 38.

NTP

NTP (ang. *Network Time Protocol*) jest internetowym protokołem synchronizacji czasu na zegarach urządzeń, którego źródłem odniesienia jest sieć. Jeżeli na Twoim sprzęcie sieciowym bądź komputerze wymagany jest precyzyjny czas, powinieneś skorzystać z NTP. Innymi słowy: zawsze powinieneś używać NTP.

Pierwotnie NTP został zdefiniowany w dokumencie RFC 958. Ostatnią udokumentowaną wersją jest wersja 4. Aby dowiedzieć się więcej o NTP, zajrzyj na stronę *http://www.ntp.org*.

W dokumencie RFC 1129, dostępnym jedynie w formacie PostScript oraz PDF, znajdziemy kilka interesujących informacji związanych z tym, jak właściwie działa NTP. Jeśli jesteś podekscytowany, widząc wzór na rysunku 38.1, RFC 1129 z pewnością Cię zainteresuje.

$$e_{i+1} = \frac{d_i}{K_g}\frac{q^{n_i}-1}{q-1} + \frac{1}{K_f}\sum_{j=1}^{i}\frac{n_j d_j}{a_{j-1}u_{j-1}}$$

Rysunek 38.1. Fragment dokumentu RFC 1129

Jeżeli nie chcesz zagłębiać się w matematykę, tak jak większość ludzi, których znam, a chciałbyś jedynie wiedzieć, co należy zrobić, aby NTP zadziałało u Ciebie, czytaj dalej.

Czym jest dokładny czas?

Jak mierzymy czas? Jak długo trwa sekunda? Kto zdefiniował czas trwania sekundy i dlaczego ma to znaczenie? Oto pytania, o których większość ludzi nie myśli. Większość z nas uznała za prawdziwe to, co nam narzucono.

Każde elektroniczne urządzenie, które posiadasz, od Twojego komputera po Twój telewizor, opiera się na czasie właściwym do pewnego stopnia. Moc w Twoim domu oscyluje na poziomie około 60 cykli na sekundę (60 Hz). Telewizja typu HD 1080 odświeża z prędkością 60 (120, 240 albo nawet 480) ramek na sekundę. Modemy, T1, DS3 i wszystkie inne usługi sieciowe wysyłają informacje z określoną liczbą bitów na sekundę. Gdy komunikują się dwa urządzenia, muszą przyjmować wspólną długość trwania sekundy.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przez długi czas definiowano sekundę jako 1/86 400 **dnia gwiazdowego** (ang. *sidereal day*) (24 godziny × 60 minut × 60 sekund = 86 400 sekund). Dzień gwiazdowy to czas, w jakim Ziemia obraca się raz (360 stopni) wokół własnej osi. Astronomowie powiedzą, że dzień gwiazdowy to czas potrzebny gwieździe na dwukrotne przekroczenie południka. **Dzień słoneczny** (ang. *solar day*) jest krótszy od dnia gwiazdowego o 4 minuty i to właśnie jest powodem występowania roku przestępnego. Kalendarz musi zostać „zresetowany" w związku z akumulacją tych czterominutowych błędów występujących każdego dnia. Wobec tego 1 sekunda gwiazdowa = 1,00278 sekundy słonecznej. Sekunda słoneczna zdefiniowana jest w ten sam sposób (1/86 400 dnia słonecznego), używa jednak Słońca jako punktu odniesienia, a nie odległej gwiazdy. W związku z tym, że Ziemia obraca się wokół Słońca, pozycja Słońca na niebie zmienia się każdego dnia.

Jak widzisz, określenie długości sekundy nie jest tak proste, jak mogłoby się wydawać. Obecnie definicję jednej sekundy zatwierdzono jako: czas trwania 9 192 631 770 okresów promieniowania odpowiadający przejściu między dwoma poziomami struktury nadsubtelnej stanu podstawowego atomu cezu 133. Prawda, że to wszystko tłumaczy? Definicja ta odnosi się do zegarów atomowych. Zegary te są najbardziej precyzyjnymi znanymi urządzeniami. Uznawane są za źródło odniesienia dla wszystkich innych zegarów.

Jeśli uznamy zegary atomowe za złoty standard, według którego inne zegary powinny być regulowane, możemy również określić, jak dokładne są one w porównaniu do atomowego wzorca.

W NTP każdy zegar umiejscowiony jest w strukturze **stratum**. Aktualny zegar atomowy jest źródłem czasu stratum 0. Urządzenia, które obsługują zapytania NTP i są bezpośrednio podłączone do zegara atomowego, znajdują się w stratum 1. Jeśli stworzysz serwer NTP, który będzie pobierał czas z serwera stratum 1, Twój serwer stanie się stratum 2 i tak dalej. Pojęcie **strata** w NTP odnosi się do liczby poziomów zegarów pomiędzy urządzeniem a źródłem czasu.

Pojęcie *strata* definiuje się w systemach telekomunikacyjnych nieco inaczej niż w NTP. W telekomunikacji warstwy stratum odnoszą się do parametru *holdover performance* oscylatora, w przypadku gdy synchronizacja zostanie utracona. Jeśli oscylator jest dokładny do $1{,}0 \cdot 10^{11}$, określany jest jako stratum 1. Oscylatory stratum 2 są dokładne do $1{,}6 \cdot 10^{8}$. Innymi słowy, poziom dokładności nie jest związany z odległością od zegara źródłowego jak w przypadku NTP.

Projektowanie NTP

NTP najczęściej nie jest projektowane, lecz po prostu implementowane w najprostszy możliwy sposób. Wielu ludzi błędnie wierzy, że wszystko, czego potrzebują, aby rozwiązać problem z czasem, to skonfigurowanie pojedynczego NTP. Założenie to utrwala się, ponieważ zazwyczaj taka konfiguracja działa. Co jednak stałoby się z Twoją siecią w przypadku, kiedy pojedyncze źródło czasu przestałoby odpowiadać lub stałoby się niedokładne?

NTP nauczyłem się na własnych trudnych doświadczeniach. Skonfigurowałem pojedyncze źródło czasu dla przełączników w rdzeniu dużej sieci. Sądziłem, że jestem sprytny, ponieważ wszystkie moje urządzenia w sieci pobierały czas z tych właśnie przełączników. Oznaczało to, że zamiast setek urządzeń tylko te musiały pobierać czas z internetu.

Pewnego dnia źródło czasu przestało odpowiadać, choć nie wiedzieliśmy o problemie. Przełączniki rdzenia (6509s) ciągle pełniły funkcję serwerów NTP, tak więc wszyscy uważali, że czas

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

jest właściwy. W tym przypadku wszystkie urządzenia miały podobny czas, nie był to jednak tzw. *real time* (*Coordinated Universal Time* albo UTC). Mimo to różnica między czasem UTC a tym raportowanym była nieznaczna — może minuta różnicy na przestrzeni kilku miesięcy.

Lekcja 1.: Zawsze korzystaj z więcej niż jednego źródła NTP. NTP nie tylko przełączy się do innego w przypadku awarii, ale też wybierze najbardziej dokładne z dostępnych. Skonfiguruj przynajmniej trzy źródła NTP dla najważniejszych urządzeń.

W tym samym czasie musieliśmy zrestartować przełączniki rdzenia w związku z pracami utrzymaniowymi. Następny dzień był katastrofą. W jakiś sposób wszystkie urządzenia w sieci — włączając w to wszystkie serwery i stacje robocze — „uznały", że ten dzień jest dniem poprzednim.

Lekcja 2.: Niektóre urządzenia Cisco mają zegar i kalendarz. Zegar utrzymuje czas uruchomionej sesji. Kiedy urządzenie zostanie zrestartowane bez dostępu do źródła NTP, odczyta czas ze swojego kalendarza. Zawsze konfiguruj urządzenia wyposażone w kalendarz tak, aby aktualizowały go automatycznie na podstawie NTP.

Przełączniki po restarcie załadowały czas ze swojego wewnętrznego kalendarza i wykorzystały ten czas, aby ustawić swoje zegary. Ponieważ nie konfigurowałem kalendarzy, domyślnie przyjęły błędne ustawienia daty i godziny. Inicjalizacja NTP byłaby w stanie to naprawić, jeśli mogłaby nawiązać połączenie z właściwym zegarem. Niestety, nie mogła. Jako że każde urządzenie w sieci otrzymywało czas z przełączników rdzenia, wszystkie utraciły dzień. Wywołało to problemy, ponieważ programy replikujące dane do serwerów backup „stwierdziły", że ostatni zapis danych jest nieprawidłowy, bo był on z przyszłości.

Dzięki mojej niezdolności do poprawnego wdrożenia NTP firma zakończyła dzień utratą transakcji z całego dnia. To nie był dobry dzień.

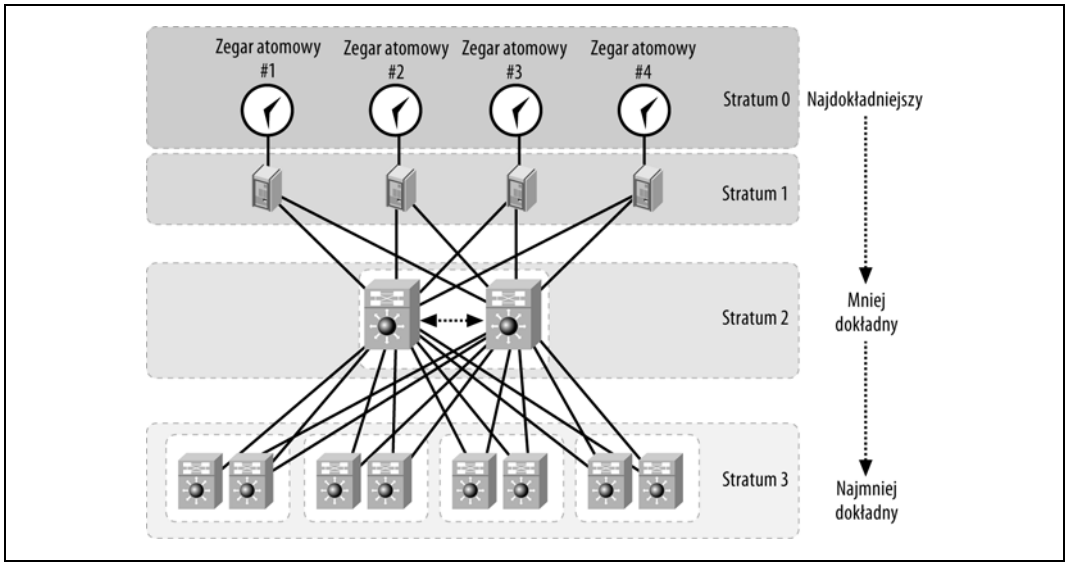
Problem sprowadzał się do mojego przekonania, że potrzebuję tylko jednego serwera NTP. Założyłem, że nie będzie żadnych problemów z czasem, ponieważ skonfigurowałem urządzenia sieciowe, aby pobierały czas z zewnątrz.

Projektując NTP, nadal lubię ograniczać liczbę urządzeń, które będą odpytywać internet. Zazwyczaj ustawiam dwa przełączniki rdzeniowe jako serwery NTP dla całej sieci. Konfiguruję te dwa przełączniki (załóżmy parę wysokiej dostępności) tak, aby otrzymywały czas z internetu, z przynajmniej trzech serwerów czasu. Przełączniki funkcjonują jako **NTP peers** (przełączniki równorzędne). Oznacza to, że są sobie równe, jeśli chodzi o warstwę stratum. Dzięki temu mogą one również siebie nawzajem traktować jako źródła czasu.

Później konfiguruję wszystkie pozostałe urządzenia sieciowe tak, aby otrzymywały czas z przełączników rdzenia sieci. Zakładając, że wszystkie serwery czasu w internecie są w stratum 1, przełączniki rdzenia powinny działać w stratum 2, a pozostałe przełączniki w sieci w warstwie stratum 3. Hierarchia ta pokazana jest na rysunku 38.2.

Starsze wersje Windows, jak Windows Server 2000, nie wspierają NTP. Zamiast tego wspierają *Simple NTP* (SNTP). W związku z tym ograniczeniem nie możesz wykorzystać urządzeń Cisco jako serwerów NTP dla starszych maszyn Windows bez zainstalowania na nich dedykowanego klienta NTP. Obecnie rzadko jest to problemem, ale powinieneś uważać na starsze systemy operacyjne, które nie współpracują dobrze z NTP.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 38.2. Hierarchia NTP w sieci korporacyjnej

Aby znaleźć publicznie dostępny serwer czasu, wpisz w wyszukiwarce *public NTP servers*. Niektóre serwery wymagają rejestracji, podczas gdy inne są w pełni otwarte.

Konfiguracja NTP

NTP jest aplikacją typu klient-serwer. Urządzenia biorące udział w NTP są albo serwerami NTP, które dostarczają czas do innych urządzeń, albo klientami NTP odpytującymi serwery NTP. Serwery są również klientami i mogą być łączone równorzędnie między sobą. Konfiguracja IOS i NX-OS jako klienta jest łatwiejsza, dlatego od niej zaczniemy.

Klient NTP

W celu skonfigurowania urządzenia IOS albo NX-OS pobierającego czas z serwera NTP, skorzystaj z komendy `ntp server`. Możesz wypisać tyle serwerów, ile chcesz, każdy w osobnej linii. Poniżej wypisałem siedem publicznie dostępnych serwerów NTP. Zastosowanie tak wielu pomoże zademonstrować zachowanie NTP.

Na przełącznikach Nexus z wieloma VDC tylko domyślny VDC może mieć skonfigurowane NTP.

```
ntp server 132.163.4.102
ntp server 193.67.79.202
ntp server 152.1.58.124
ntp server 128.118.46.3
ntp server 129.6.15.29
ntp server 64.236.96.53
ntp server 208.184.49.9
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kiedy już skonfigurowałeś serwery NTP, powinieneś zacząć otrzymywać właściwy czas (zakładając, że serwery są osiągalne).

Aby zobaczyć status serwerów NTP w IOS, użyj komendy `show ntp associations`. W tym rozdziale wykorzystam system IP-PBX, który zbudowałem w rozdziale 30. Oto router 2811 z uruchomioną wersją IOS 15.0(1) M4:

```
R1-PBX#sho ntp associations
    address      ref clock       st    when poll reach delay offset  disp
~127.127.1.1    .LOCL.          7       7   16  37   0.000  0.000   437.71
~128.118.46.3   .INIT.          16      -   64  0    0.000  0.000   15937.
~208.184.49.9   .INIT.          16      -   64  0    0.000  0.000   15937.
+~129.6.15.29   .ACTS.          1       63  64  36   23.087 -4.829  939.32
*~64.236.96.53  .ACTS.          1       8   64  37   18.705 -4.804  438.60
+~193.67.79.202 .PPS.           1       2   64  37   109.59 -5.209  439.79
~132.163.4.102  .ACTS.          1       56  64  17   118.13 15.596  938.89
~192.168.1.250 64.236.96.53     2       32  64  37   2.802  -2.104  439.79
~152.1.58.124  .INIT.          16       -   64   0   0.000  0.000   15937.
* sys.peer, # selected, + candidate, - outlyer, x falseticker, ~ configured
```

Zaprezentowano wiele informacji, jednak potrzebujesz wiedzieć, co oznaczają tylko niektóre z nich. `Address` to oczywiście serwer, który skonfigurowałeś. Na lewo od serwera znajduje się jeden bądź więcej znaków wskazujących status serwera. Klucz dla tych symboli znajduje się w ostatniej linijce wyniku. Tabela 38.1 zawiera ich bardziej szczegółowe wyjaśnienie. Linia rozpoczynająca się od gwiazdki wskazuje na serwer, z którym router zsynchronizował swój czas. Zwróć uwagę na niektóre wpisy mające stratum równe 16 z parametrami `delay` i `offset` równymi 0. Są to martwe serwery NTP. Szukając w sieci, znajdziesz zapewne wiele martwych serwerów NTP. Listę tych serwerów zaczerpnąłem z pierwszej edycji książki. Wtedy, w roku 2007, wszystkie te adresy odpowiadały poprawnie.

Zadbaj o to, by sprawdzać konfigurację NTP każdego roku, tak aby wiedzieć, że serwery NTP nadal są aktywne.

Tabela 38.1. Znaki wskazujące status serwera NTP (ang. association characters)

Znak	Opis
*	Zsynchronizowany
#	Prawie zsynchronizowany
+	Wybrany do możliwej synchronizacji
-	Kandydat do wybrania
~	Statycznie skonfigurowany

Kolejna kolumna prezentuje zegar odniesienia dla danego serwera. Mogłoby to być użyteczne, jednak podawane jest tylko informacyjnie. Na przykład pierwszy serwer NTP pokazuje jako zegar odniesienia ACTS (ang. *Automated Computer Time Service*). ACTS jest usługą *dial-up* oferowaną przez instytut NIST (ang. *National Institiute of Standards and Technology*). Drugi serwer pokazuje odniesienie do GPS, wskazując, że otrzymuje czas z satelit GPS (ang. *Global Positioning System*), które zwierają atomowe zegary.

Trzecia kolumna prezentuje stratum. Serwery pokazujące stratum 16 są niedostępne i nie dostarczają czasu.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Aby sprawdzić, czy Twój zegar jest zsynchronizowany, użyj komendy IOS `show ntp status`. Komenda ta pokaże status NTP wraz ze stratum:

```
R1-PBX#sho ntp stat
Clock is synchronized, stratum 2, reference is 64.236.96.53
nominal freq is 250.0000 Hz, actual freq is 249.9924 Hz, precision is 2**24
reference time is D117CE2C.7FAFBB96 (14:57:32.498 EST Tue Mar 1 2011)
clock offset is -4.8047 msec, root delay is 18.70 msec
root dispersion is 444.19 msec, peer dispersion is 438.60 msec
loopfilter state is 'CTRL' (Normal Controlled Loop), drift is 0.000030117 s/s
system poll interval is 64, last update was 3 sec ago.
```

Jeśli Twój zegar nie jest zsynchronizowany, polecenie `show ntp status` wskaże tego powód. W tym przykładzie usunę wszystkie serwery NTP, tak aby nie było żadnego zegara, do którego można się odnieść. Zauważ, że teraz stratum wynosi 16:

```
R1-PBX#sho ntp status
Clock is unsynchronized, stratum 16, no reference clock
nominal freq is 250.0000 Hz, actual freq is 249.9924 Hz, precision is 2**24
reference time is D117CB70.76B05259 (14:45:52.463 EST Tue Mar 1 2011)
clock offset is 0.0000 msec, root delay is 0.00 msec
root dispersion is 0.42 msec, peer dispersion is 0.00 msec
loopfilter state is 'CTRL' (Normal Controlled Loop), drift is 0.000030140 s/s
system poll interval is 64, last update was 376 sec ago.
```

Kolejnym sposobem na pokazanie, że NTP kieruje zegarem systemowym, jest komenda `show clock details`:

```
R1-PBX#sho clock detail
15:02:07.371 EST Tue Mar 1 2011
Time source is NTP
Summer time starts 02:00:00 EST Sun Mar 13 2011
Summer time ends 02:00:00 EDT Sun Nov 6 2011
```

Odkąd polecenie `show ntp association` nie istnieje w NX-OS, moja niedojrzała część osobowości nie może przestać śmiać się ze skrótu `show ntp ass`. Dla NX-OS w celu pokazania statusu serwerów NTP użyj komendy `show ntp peers`. Poniżej skonfigurowałem te same serwery na urządzeniu Nexus 7000:

```
NX-7K-1(config)# sho ntp peers
----------------------------------------------------
Peer IP Address                    Serv/Peer
----------------------------------------------------
132.163.4.102                  Server (configured)
193.67.79.202                  Server (configured)
152.1.58.124                   Server (configured)
128.118.46.3                   Server (configured)
129.6.15.29                    Server (configured)
64.236.96.53                   Server (configured)
208.184.49.9                   Server (configured)
```

Polecenie `show ntp peer-status` pokaże również tzw. *NTP peers*. Mój wynik wskazuje na brak połączenia, ponieważ Nexus 7000 w środowisku laboratoryjnym nie miał dostępu do internetu:

```
NX-7K-1(config)# sho ntp peer-status
Total peers : 4
* - selected for sync, + - peer mode(active),
- - peer mode(passive), = - polled in client mode
remote           local         st      poll     reach     delay      vrf
------------------------------------------------------------------------------
+132.163.4.102   0.0.0.0        16      16         0      0.00000    default
=193.67.79.202   0.0.0.0        16      64         0      0.00000    default
```


Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

```
=152.1.58.124    0.0.0.0        16      64         0      0.00000    default
=128.118.46.3    0.0.0.0        16      64         0      0.00000    default
=129.6.15.29     0.0.0.0        16      64         0      0.00000    default
=64.236.96.53    0.0.0.0        16      64         0      0.00000    default
=208.184.49.9    0.0.0.0        16      64         0      0.00000    default
```

Serwer NTP

Aby skonfigurować urządzenie IOS jako serwer NTP, wpisz komendę `ntp master`:

```
6509(config)#ntp master
```

Kiedy konfigurujesz urządzenie IOS, by było serwerem NTP, potrzebuje ono pięciu lub więcej minut, aby się zsynchronizować. Nie można tego wymusić.

Na urządzeniu, które jest wyposażone w kalendarz, musisz również wpisać komendę `ntp update-calendar`:

```
6509(config)#ntp update-calendar
```

Aby skonfigurować urządzenie z takim samym stratum co *peer NTP*, użyj komendy `ntp peer ip-address`:

```
6509(config)#ntp peer 10.10.10.1
```

Pozwoli to innemu urządzeniu w tym samym stratum współdzielić czas.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 39.

Awarie

Pewne awarie mogą zostać wykryte bardzo łatwo. Kiedy zawodzi router, przeważnie pojawiają się dość oczywiste symptomy. W przypadku gdy interfejs WAN (*WAN Interface Card*) zaczyna zniekształcać pakiety, problem może być nieco trudniejszy do zdiagnozowania. W tym rozdziale umieszczę kilka przykładów tego, co może być przyczyną awarii, oraz opiszę, jak efektywnie rozwiązywać pojawiające się problemy.

Błąd ludzki

Błąd ludzki (ang. *human error*) może być najtrudniejszy do wytropienia, a po wykryciu prawie niemożliwy do udowodnienia. Przekonanie człowieka do przyznania się do błędu, który popełnił, może sprawiać problem — w szczególności wtedy, kiedy osobą odpowiedzialną jesteś Ty sam!

Pracując kiedyś w sieci o zasięgu ogólnoświatowym, administrowałem zmianami list ACL związanymi z ruchem SNMP do routerów. Dodaliśmy właśnie nowy serwer do zarządzania siecią i chcieliśmy dodać jego adres do wszystkich routerów, tak aby mogły być monitorowane oraz zarządzane z centralnego biura.

Zamiast wcześniej napisać kod konfiguracyjny, przetestować go w laboratorium, a następnie wdrożyć w odpowiednim oknie zmian, postanowiłem, będąc pewnym siebie, zaimplementować zmiany natychmiast. Były one minimalne — tylko jedna linia kodu, co złego mogło się zatem stać?

Oczywiście namieszałem w jednym z routerów i tym samym usunąłem aktywną listę ACL na wejściowym interfejsie. W firmie, o której piszę, router ten był jedynym na cały kontynent, jakim jest Australia. Jako że z prostych problemów żaden jeszcze mnie nie spotkał, katastrofa miała miejsce w piątkową noc (czasu australijskiego) podczas wakacyjnego weekendu. Do wtorku nikt tam nie był osiągalny. W wyniku ludzkiego błędu (mojego) przez trzy dni międzynarodowa instytucja finansowa nie była w stanie dostać się do swoich serwerów w Australii. Lubię dorzucać takie wydarzenia do swojego „doświadczenia".

Dobrym sposobem zapobiegania takim rodzajom katastrof jest wykorzystanie komendy `reload`. Jest ona w stanie zaplanować restart routera po upływie określonego czasu. W związku z tym, że konfiguracja czyniąca router niedostępnym nie została zapisana, automatyczny restart routera przywróciłby ustawienia domyślne.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Aby zrestartować router po 15 minutach, należy zastosować komendę `reload in` jak poniżej:

```
TS-1#reload in 15
Reload scheduled in 15 minutes // przeładowanie nastąpi za 15 minut
Proceed with reload? [confirm] // kontynuować? [potwierdź]
```

Alternatywnie można skonfigurować godzinę, o której nastąpi restart routera, za pomocą komendy `reload at`:

```
TS-1#reload at 12:05 July 10
Reload scheduled for 12:05:00 UTC Mon Jul 10 2006 (in 566 hours and 33 minutes)
          // przeładowanie zaplanowano na poniedziałek 12:05:00 czasu UTC 10 lipca 2006 (za 566 godzin i 33 minuty)
Proceed with reload? [confirm] //kontynuować? [potwierdź]
```

Uważaj jednak, jeśli wprowadziłeś zmiany, router zapyta się o ich zapisanie przed restartem.

Pamiętaj, aby anulować restart, jeżeli zmiany zostały zaimplementowane poprawnie! W celu anulowania zaplanowanego restartu wykonaj komendę `reload cancel`:

```
TS-1#reload cancel
TS-1#
***
*** --- SHUTDOWN ABORTED --- // wyłączanie przerwano
***
```

Awaria wielu komponentów

Wiele sieci zaprojektowanych zostało tak, aby unikać pojedynczych punktów awarii. Awaria tych sieci następuje w sytuacji, w której nieprawidłowo zaczyna funkcjonować więcej niż jeden komponent. Awaria wielu komponentów może być spowodowana zawodnym zasilaniem, ale też zwykłym przypadkiem. Jedno urządzenie może również spowodować awarię innego.

Wiele przedsiębiorstw korzysta z zapór sieciowych Cisco PIX albo ASA, umieszczając je w redundantnych parach. Jest to sprawdzona i niezawodna droga do zagwarantowania, że Twoja zapora sieciowa nie stanie się pojedynczym punktem awarii. Twój firewall będzie jednak pojedynczym punktem awarii, jeśli zapasowa zapora ogniowa ulegnie awarii, a nikt nie będzie o tym wiedział. Problem polega na tym, że wielu ludzi tworzy scenariusze *failover*, ale nie monitoruje urządzeń. Miałem kiedyś klienta, którego zapasowy firewall pozostawał w trybie *offline* od miesięcy. Kiedy podstawowa zapora uległa awarii, cała sieć stała się niedostępna, co spowodowało krzyki kadry zarządzającej. Wszyscy założyli, że zapory działają prawidłowo, nie sprawdzając tego. Narzędzia do zarządzania siecią niczego nie pokazały, inaczej niż w poprzednim przypadku dużej przerwy w dostępie do systemów. Jedna z największych katastrof, której byłem świadkiem, wynikała z tego, że system nie potrafił wrócić z zapasowej do swojej podstawowej instancji.

Czasami wiele urządzeń zawodzi z im tylko znanych powodów. W jednej z instalacji podwójny moduł zarządzający w 6509 uległ dość złożonej awarii. Podstawowy moduł nie działał, z kolei podstawowy MSFC pozostał aktywny. W związku z tym, że MSFC jest fizycznie przywiązany do modułu zarządzającego, tak aby utrzymać łączność z siecią w przełączniku, cała lokalizacja została odcięta, mimo że nastąpiło poprawne przełączenie na *moduł* zapasowy. Ciągle aktywne MSFC nie potrafiło zachować łączności. Tego typu problem nie zdarza się w przypadku, kiedy używasz najnowszej wersji IOS, ale w tamtej chwili powiedziałbym, że to, co właśnie obserwujemy, nie jest możliwe.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Łańcuch katastrof

Jeśli chodzi o redundancję, nowoczesne samoloty pasażerskie są jednymi z najbardziej zdumiewających urządzeń na świecie. Gdy zawiedzie jedno urządzenie, inne przejmuje jego funkcjonalność. Jeśli awarii ulegną dwa systemy, trzeci jest w stanie przejąć ich obowiązki przy ograniczonych parametrach. Samoloty pasażerskie mają nawet wielu pilotów! Mimo tej całej bezpiecznej inżynierii samoloty ulegają awariom i spadają. Kiedy się to stanie, przeważnie odkrywa się, że przyczyną rozbicia była seria zdarzeń, pojedynczo dość niegroźnych, ale w połączeniu oznaczających katastrofę. Taki łańcuch zdarzeń określany jest jako **łańcuch wypadków** (ang. *accident chain*).

Również mocno redundantne sieci mogą cierpieć w efekcie łańcucha wypadków. Dla określenia wpływu, jaki tak złożona awaria może mieć na biznes (nie wspominając o wypłatach), lubię używać określenia **łańcuch katastrof** (ang. *disaster chain*).

Wyobraź sobie sieć, której rdzeń stanowią dwa przełączniki Cisco 6509. W związku z podniesieniem wersji zaplanowano niedostępność najpierw jednego, potem drugiego. W związku z tym, że urządzenia są redundantne, każde z nich mogło zostać wyłączone bez wpływu na sieć. Pierwszy przełącznik został wyłączony, z czego nie wynikł żaden incydent. Podczas pracy jednak zaplątałem się w przewody zasilające drugiego przełącznika, wypinając go z zasilania. Oczywiście źródło zasilania przełącznika 6509 było wyposażone w zaciski do zabezpieczenia kabli zasilających, ale inżynier ostatnio przy nich pracujący zapomniał je docisnąć. Każdy 6509 ma dwa źródła zasilania, które są podłączone do innych obwodów, zatem wyciągnięcie jednego z przewodów nie powinno być problemem. Inżynier, który zapomniał docisnąć zacisków, stwierdził jednak, że przewody będą wyglądać lepiej, jeśli się je ściśnie i zwiąże razem. Potknąłem się zatem o kable związane razem, co oznaczało, że wypiąłem całe zasilanie zapewniające dostęp do sieci przełącznika 6509. Efekt: całkowita awaria sieci.

Podsumowując, niedostępność spowodowała następująca kombinacja zdarzeń:

- Urodziłem się pozbawiony kompletnie koordynacji ruchowej.
- Ostatni inżynier pracujący z systemem związał razem przewody i zostawił je w miejscu, w którym mogłem się o nie potknąć.
- Ostatni inżynier, który pracował z systemem, nie docisnął zacisków przy przewodach zasilających.
- Wyłączyłem jeden z przełączników 6509.
- Potknąłem się o przewody zapasowego przełącznika 6509.

Żadne z tych zdarzeń samo nie spowodowałoby niedostępności. Właściwie nawet dowolne dwa z nich by jej nie spowodowały. Jeśli kable zasilające nie zostałyby zaciśnięte, a przewody nie zostałyby związane razem, moja wielka stopa wypięłaby prawdopodobnie tylko jeden przewód, pozostawiając przełącznik z jednym aktywnym źródłem zasilania.

Odkrycie przyczyn w łańcuchu katastrof może wymagać przeprowadzenia prawdziwego dochodzenia. W moim przypadku potknięcia się o przewody przesłuchujący mnie menedżer zapytał: „Jak to się mogło stać?”. Posiadanie gotowej odpowiedzi na to pytanie może ocalić Twoją pracę, szczególnie wtedy, kiedy odpowiedzią jest to, że jesteś niezdarą.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Brak testów awarii

Pracowałem kiedyś w zespole projektującym dużą infrastrukturę dla strony internetowej typu *e-commerce*. Mówiąc dużą, mam na myśli 8 przełączników Cisco 6509 oraz ponad 200 fizycznych serwerów (większość z kilkoma maszynami wirtualnymi) dostarczających zawartość 1 GB na sekundę. Wszyscy byli zestresowani, ponieważ terminy nagliły. W związku z terminami związanymi z etapami projektu wyeliminowano jedną z kluczowych faz: przetestowanie wystąpienia awarii.

Kiedy uruchomiliśmy ośrodek produkcyjnie, jedno z urządzeń zawiodło. Ośrodek został zaprojektowany tak, aby nie miał pojedynczego punktu awarii, jednak strona przestała działać prawidłowo. Okazało się, że urządzenie zapasowe zostało błędnie skonfigurowane, ale nie było to widoczne przed samym przełączeniem. W związku z tym, że awaria uniemożliwiła łączność z ośrodkiem, nie mieliśmy innej możliwości dostania się do niesprawnego sprzętu niż przejechanie się do zapasowego centrum danych. Awaria ta, która nigdy nie powinna się wydarzyć, kosztowała 2 godziny niedostępności ośrodka. Trwała do momentu, gdy ktoś z kablem konsolowym udał się do centrum danych.

Gdyby wykonano testowanie *failover*, problem zostałby wykryty podczas testów, a niedostępności można by uniknąć. Projekt był poprawny, ale jego implementacja już nie. Nalegaj zawsze na testy typu *failover* w środowiskach wysoko dostępnych. Testowanie awarii powinno być regularnie planowane i cyklicznie wykonywane w dedykowanych oknach czasowych.

Rozwiązywanie problemów

Cała ta książka napisana została właśnie w oparciu o rozwiązywanie problemów (ang. *troubleshooting*). Widziałem ludzi, których można określić jako prawdziwych *troubleshooters*, jak i takich, którzy się do tego zupełnie nie nadają. Wydaje się, że niektórzy po prostu wyczuwają źródło skomplikowanego problemu, podczas gdy inni nie są w stanie dowiedzieć się, co się stało, nawet jeśli **sami** są przyczyną problemu.

Najciekawsze problemy to najczęściej takie, które powodują największe szkody. Są to wydarzenia, które mogą przerwać Twoją karierę. Miałem związek z awarią strony internetowej, której skutkiem był zerowy przychód podczas niedostępności strony. Stykałem się z awariami w bankach, gdzie każda minuta kosztowała miliony dolarów strat w niezrealizowanych transakcjach. Najlepsze awarie to te szybko usuwane oraz takie, których powodem nie były moje błędy. Te, które wynikły z mojej winy, jako że byłem ich przyczyną, identyfikowałem tak szybko, jak to było możliwe. Ludzie popełniają błędy. Gdy ukrywają ich prawdziwą przyczynę, nie chcąc przyjąć odpowiedzialności, awarie najczęściej trwają dużo dłużej aniżeli wtedy, kiedy osoby rozwiązujące problem są odpowiednio poinformowane.

Niezależnie od problemu lub sytuacji jest kilka rzeczy, o których należy pamiętać podczas wykrywania przyczyn awarii. Oto ich krótka lista.

Zachowaj spokój

Pracowałem kiedyś z byłym sierżantem marynarki. Został on postrzelony w ramię podczas walki, po czym przez miesiące poddawany był rehabilitacji. Aktualnie pracował dla mnie jako starszy inżynier ds. sieci, zajmując się wsparciem globalnej sieci liczącej ponad 10 000 komputerów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Pewnego dnia nastąpiła awaria. Podczas prób znalezienia i rozwiązania problemu za plecami byłego sierżanta stał sam wiceprezes. Inżynier z każdą upływającą minutą stawał się coraz bardziej zestresowany. Zauważyłem, że przeszkadza mu to w efektywnej pracy nad problemem. Pierwsze, co zrobiłem, to wyprosiłem wiceprezesa, obiecując mu co 10 minut przekazywać status sprawy. Drugą rzeczą było napisanie na tablicy „Spokojnie, nikt do Ciebie nie strzela". Były sierżant uśmiechnął się do mnie i od razu można było zauważyć, że się rozluźnił. Problem został rozwiązany przed przekazaniem pierwszego statusu wiceprezesowi.

Awarie sieci przeważnie nie są kwestią związaną z narażaniem ludzkiego życia. W ekstremalnych przypadkach ryzykujesz swoją posadą, jednak nawet jej utrata pozwoli Ci wrócić do domu, uściskać dzieci, małżonkę, psa albo konsolę Xbox, jeśli to Ci sprawia przyjemność.

Na im większy stres sobie pozwolisz, tym dłużej trwać będzie awaria.

Rejestruj swoje działania

Podczas pracy przy awarii zapisuj swoje działania oraz czas ich wykonania. Możliwe, że będziesz je musiał umieścić w raporcie po awarii. Rejestrowanie działań, które wykonałeś, może się później okazać niezwykle cenne. Posłuży również temu, aby w przyszłości nie marnować czasu, powtarzając czynności, które nie przyniosły efektu.

Dowiedz się, co uległo zmianie

Awarie nie występują bez przyczyny. Nawet jeśli nie ma oczywistego uszkodzenia urządzenia, to jeśli wszystko działało, a teraz nie działa, coś musiało się zmienić. W niektórych środowiskach ludzie się przyznają, w innych nie. Niezależnie od tego pierwsze pytanie, które powinieneś zadać, brzmi: „Co się zmieniło?". Jeśli coś zmieniło się bezpośrednio przed wystąpieniem awarii, istnieje duża szansa, że zmiana ta związana jest z awarią. Dowiedz się, czy możesz odwrócić zmianę, upewnij się jednak wcześniej, jakie to może mieć następstwa.

Awarie mogą być skutkiem pozornie niezwiązanych z nimi zmian. Kiedy pracowałem z komputerami typu *mainframe*, byłem świadkiem kilku nieco dziwnych scenariuszy awarii. Jeden z klientów skarżył się, że jego maszyna restartuje się codziennie między północą a godziną pierwszą w nocy. Trwało to od miesiąca i zdarzało się każdej nocy, przyczyny były jednak niejasne. Klient był bardzo wyrozumiały i pomocny, jednak nie wykonywały się kopie zapasowe i to powodowało problemy. Wysłaliśmy w końcu kogoś na miejsce, aby sprawdził, co się dzieje.

Wysłana osoba zauważyła, że każdej nocy, około północy, na parkingu parkuje swój samochód pewien oficer policji. Kiedy opuszcza parking, wzywa przez radio dyspozytora. W momencie kiedy policjant naciskał przycisk mikrofonu radia, system się restartował.

Odkryliśmy, że na miesiąc przed naszym przybyciem administrator odłączył grupę terminali od komputera *mainframe*. Usunął terminale, lecz pozostawił kable szeregowe, które łączyły je z serwerem. Przewody te pełniły rolę anteny. Kiedy policjant dotykał swojego radia, kable otrzymywały sygnał, a impuls elektryczny powodował restart.

Jeżeli nie jesteś w stanie odnaleźć przyczyny problemu, przyjrzyj się wykonanym zmianom, niezależnie od tego, jak bardzo wydają się niezwiązane z problemem.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Najpierw sprawdź warstwę fizyczną!

Zasady tej nauczył mnie mój szef, kiedy pracowałem jeszcze u jednego z dostawców ISP. Uczył rozwiązywania problemów inżynierów pracujących w obszarze związanym z telekomunikacją dla jednego z operatorów regionalnych (RBOC, ang. *Regional Bell Operating Company*). Założenie było takie, że najwięcej awarii ma swoją przyczynę w fizycznych uszkodzeniach. Okablowanie i sprzęt powinny być sprawdzone w pierwszej kolejności. Z tych dwóch najpierw sprawdzić należy okablowanie. Zapamiętaj tę zasadę, a będziesz zaskoczony, jak często się ona sprawdza. Do dzisiaj słyszę jego wrzaski podczas awarii: „Warstwa fizyczna jako pierwsza!".

Nie zakładaj, udowodnij

Jeśli coś z góry założysz, zwróci się to przeciwko Tobie. Kiedy usłyszysz kogoś (włącznie z samym sobą) mówiącego: „To nie może być to, ponieważ…", podejmij wysiłek udowodnienia tego twierdzenia. Podczas awarii tendencją inżynierów jest przekonywanie samych siebie, że coś jest prawdą tylko dlatego, że tak uważają. Nie zakładaj, że coś jest prawdą, dopóki tego nie udowodnisz.

Znajdź problem

Problemy są często zapowiadane przez wiele symptomów. Dodatkowo niektóre problemy powodują inne. Na przykład możesz twierdzić, że napotkałeś złożony problem związany z routingiem, ponieważ pakiety przesyłane są w sposób, który uważasz za niewłaściwy. Problem może być jednak spowodowany uszkodzeniem interfejsu. Możesz spędzić wiele godzin, szukając złożonego problemu routingu, tymczasem będzie on leżał w zepsutym interfejsie. Spróbuj znaleźć prawdziwą przyczynę problemu. Złożone usterki się zdarzają, lecz pojedyncze awarie są znacznie częściej spotykane.

Nie wypatruj zebr

Miałem kiedyś doktora, który powiedział mi „Jeśli słyszysz tętent kopyt, nie zakładaj, że to zebry". Stwierdzenie to podobne jest do zasady Brzytwy Ockhama, którą przytacza się zazwyczaj w taki sposób: „Ze wszystkich teorii zazwyczaj ta najprostsza okazuje się prawidłowa". To, iż często boli mnie głowa, nie oznacza, że mam guza mózgu.

Ta sama zasada odnosi się do każdego problemu, przed którym stajesz. Gdy widzisz poprzestawiane bity, nie oznacza to, że kosmici zmieniają Twoje dane, chcąc zniszczyć świat. Sprawdziłbym w pierwszej kolejności okablowanie. NORAD (ang. *North American Aerospace Defense Command*) nie będzie już kontrolował moich rozmów!

Wykonaj kontrolę warstwy fizycznej

Jeśli kręcisz się w kółko i nic nie ma sensu, przeprowadź kontrolę warstwy fizycznej. Są szanse, że poczyniono pewne założenia, które taki audyt może obalić. Czasami dokumentacja jest nieaktualna albo po prostu błędna. Jeśli masz wątpliwości, narysuj schemat sieci od początku. Wykorzystaj tablicę lub tablet. Nie interesuje Cię przepiękna dokumentacja, lecz faktyczne dane.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

W przypadku restartów komputera *mainframe* audyt warstwy fizycznej pomógł, ponieważ otrzymaliśmy zadanie usunięcia kabli.

Pewnego razu zostałem zaangażowany w problem sieci, nad którym pewien zespół pracował od ponad tygodnia. Sieć była zwodniczo prosta: dwa łącza T1, cztery routery i dwa serwery. Zespół ludzi, niezależnie od wysiłku, jaki wkładano, nie potrafił odnaleźć przyczyny problemu. Przejrzałem ich notatki i zaleciłem im wykonanie kompletnego audytu fizycznego, każdego urządzenia, interfejsu i adresu sieciowego. Poinstruowałem również, aby zbadali łączność pomiędzy każdym z węzłów sieci. W ciągu godziny zostałem poinformowany, że odkryli przyczynę. Karta WIC na T1 miała problem, co powodowało deformację pakietów. Problem nie był oczywisty, dopóki nie przeszło się procesu technicznej kontroli. Źródło kłopotu odkryto dzięki testom łączności między wszystkimi punktami tej prostej sieci.

Przekazywanie problemów

Jeśli nie wiesz, co się zepsuło, przekaż problem do kogoś innego. Zazwyczaj Twój szef poprosi Cię o skontaktowanie się z Cisco, jeśli problem dotyczy sieci (i Twoim dostawcą jest Cisco). W przypadku gdy istnieje wewnętrzny dział, do którego możesz zgłosić problem, zrób to. Gdy czujesz, że problem jest poza Twoim zasięgiem, nie marnuj czasu, dzwoń po posiłki. Czasem dotarcie do właściwego człowieka może zająć wiele czasu. Pracowałem w środowiskach, w których w momencie pojawienia się problemu zaleceniem było, niezależnie od tego, czy potrafiliśmy rozwiązać problem, czy nie, zgłosić go natychmiast w centrum wsparcia technicznego firmy Cisco (TAC). Jeśli udało się go rozwiązać, anulowaliśmy zgłoszenie, jeśli nie, byliśmy już na drodze dotarcia do odpowiedniej osoby po drugiej stronie słuchawki.

Zespołowe rozwiązywanie problemów

Zespół rozwiązujący problem wymaga lidera. Mógłbym napisać całą książkę tylko na ten temat. Jeśli jesteś członkiem zespołu rozwiązującego problem, pracuj tylko nad częścią, która została Ci powierzona. Zajmując się pracą innego członka zespołu, marnujesz nie tylko swój czas, ale też drugiej osoby.

Stanie nad osobą próbującą rozwiązać problem i wykrzykiwanie swoich pomysłów również nie pomaga. Gdy jesteś przekonany o słuszności swojego rozwiązania, ale nikt Cie nie słucha, znajdź sposób na jej udowodnienie i wymuś, by Cię wysłuchano.

Janitor Principle

Zasada *Janitor Principle* mówi, że tłumaczenie problemu komuś, kto zupełnie go nie rozumie, może spowodować, iż zobaczysz coś, czego wcześniej nie dostrzegłeś. Jest to zdumiewająco potężne narzędzie. Wytłumaczenie komuś złożonego problemu wymaga zredukowania go do prostych składowych, co może prowadzić do spojrzenia na zagadnienie z innego punktu widzenia. Zmiana sposobu patrzenia jest często tym, czego potrzeba, aby rozwiązać problem.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 40.

Maksymy GAD-a

Lata pracy w branży uświadomiły mi, że światem IT kierują określone siły. Są one widoczne w prawie każdym aspekcie życia, lecz ich aplikacja nigdzie nie jest tak oczywista jak w przypadku IT.

W każdej sytuacji, w której inżynier nie robi tego, co chciałby, albo gorzej — tego, co uważa za słuszne — siły te wkraczają do gry. Wierzę, że im więcej ludzi je zrozumie, tym mniej konfliktów zdarzać się będzie między inżynierami a ich przełożonymi.

Moje inicjały to GAD, wielu ludzi tak właśnie na mnie woła. Z największą pokorą przedstawiam maksymy GAD-a.

Maksyma 1.

Wiodące siły w dziedzinie projektowania sieci to:

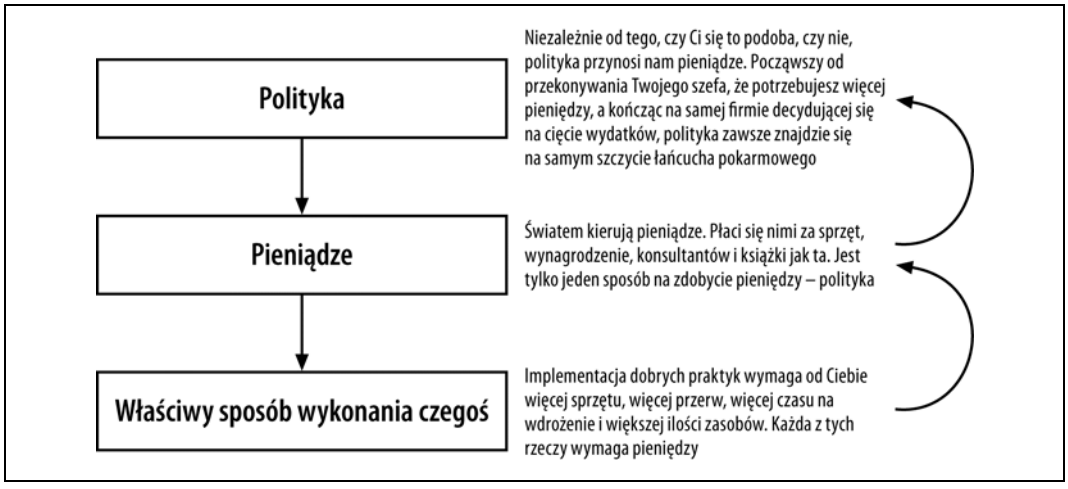
- polityka,
- pieniądze,
- właściwy sposób wykonania czegoś.

Maksyma 1.: Projektowanie sieci zależy od polityki, pieniędzy, właściwego sposobu wykonania — w tej kolejności.

W skrócie opisuje je rysunek 40.1. Idea jest naprawdę prosta. Inżynier dąży do „właściwego sposobu wykonania", który zazwyczaj opisuje dobra praktyka. Aby to zrobić, niezbędne są jednak pieniądze. Zdobycie ich wymaga biegłości w polityce. Inaczej mówiąc: jeśli chcesz wykonać coś we właściwy sposób, będziesz potrzebował pieniędzy, a droga do nich prowadzi przez politykę. Kiedy to czytasz, słyszę w myślach Twoje jęki: „Nienawidzę polityki!". Prawda jednak jest taka, że nienawidzisz **brudnej** polityki.

Spójrzmy bliżej na wspomniane trzy elementy oraz ich wzajemne relacje.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com



Rysunek 40.1. Trzy zasady projektowania sieci według GAD-a

Polityka

Polityka to nie tylko prezydent nawołujący do głosowania lub wiceprezydent zatrudniający swojego brata zamiast osoby najlepiej nadającej się na to stanowisko. Polityka to między innymi **sztuka intryg i manipulacji politycznej uprawiana przez jednostkę lub grupę mającą za cel przejęcie kontroli lub władzy**. Grupa to, wierz mi lub nie, dwie lub więcej osób. Każda rozmowa z drugą osobą zawiera jakąś formę polityki. Za każdym razem, kiedy witasz się z kimś, Twoja postawa i nastawienie powodują, że osoba ta wyrabia sobie pewne zdanie na Twój temat. To jest polityka. Jeżeli pójdziesz do szefa z genialnym pomysłem, chcąc uzyskać potrzebne fundusze, sądzisz, że Cię wesprze, jeśli jeszcze wczoraj nazwałeś go tępogłowym? Czy polityką będzie odrzucenie Twojego pomysłu? Może to Ty uprawiałeś politykę, rzucając w jego stronę obelgi.

Moim zdaniem polityka zaczyna wchodzić w grę zawsze wtedy, kiedy spotykają się dwie osoby. Jeśli masz reputację człowieka „trudnego", znajdzie to swój wpływ (uświadomiony lub nie), kiedy będziesz o coś prosił. I odwrotnie, jeżeli masz reputację osoby zrównoważonej, łatwej we współpracy, zadziała to na Twoją korzyść.

Teraz, kiedy zrozumiałeś, że polityka to nie zawsze coś złego, spójrz, jak w rzeczywistości wpływa ona na Ciebie. Stwierdziłem wcześniej, że bez polityki nie ma pieniędzy, co to jednak właściwie znaczy? Oto przykład.

Średnie przedsiębiorstwo rozważa wdrożenie nowej sieci. Odpowiedzialny za projekt jest dyrektor ds. technologii, który całe swoje życie, zanim został wybrany do zarządu, pracował jako programista. Jako główny inżynier proponujesz w pełni odporny na uszkodzenia projekt, który wymaga sprzętu o wartości 2 milionów dolarów i 6 miesięcy prac. Plan jest idealny. Znasz każdy szczegół związany ze sprzętem i jego działaniem. Będzie to prawdziwe dzieło.

Tymczasem dyrektor zerka na Twoją propozycję i po prostu stwierdza: „Zbyt drogie i zajmie zbyt dużo czasu. Budżet na sprzęt i wsparcie wynosi milion".

Czujesz, że zapiera Ci dech w piersiach. Twoja pierwsza odpowiedź to zapewne coś w stylu: „Ale to jest to, czego potrzebujemy" lub „Za mniej nie jesteśmy w stanie tego zrobić!".

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

To jest polityka. Dyrektor może być w stanie zdobyć dodatkowe fundusze; właściwie to jest to, co dyrektorzy potrafią najlepiej. Jeśli wyjaśnisz mu wszystkie korzyści wynikające z przyjęcia Twojego projektu, włącznie z tym, dlaczego należy wydać 2 miliony na sam sprzęt, może Cię wysłuchać. Jeszcze lepiej, jeśli wyjaśnisz, jak, już za 6 miesięcy, firmie zacznie się zwracać ta inwestycja. Masz wtedy spore szansę w walce o swój projekt.

Jeśli jednak zamiast tego zaczniesz się dąsać, skomleć i rozgłaszać, jak to firma jest źle zarządzana, troszczy się jedynie o pieniądze i jak to dyrektor nie rozumie technologii, zarządzający pomyśli, że nie jesteś w stanie zrozumieć realiów oraz pracować z ograniczeniami, jakie ze sobą niosą. Oboje możecie się mylić, ale, niezależnie od tego, to Ty nie dostaniesz funduszy. Dzieje się tak każdego dnia na całym świecie.

Oto dobra nowina: jeśli zamierzasz przedstawić swoje racje dyrektorowi i oczekujesz respektu oraz wsparcia, musisz nauczyć się grać na jego polu. Niezależnie od tego, czy Ci się to podoba, czy nie, to on ma pieniądze, władzę oraz wpływy. Jeśli nie wiesz, co oznacza ROI lub nie potrafisz rozmawiać o utracie wartości sprzętu, nie oczekuj, że otrzymasz fundusze.

Pieniądze

Pieniądze są tym, czego wszyscy chcemy. Jak już wcześniej powiedziałem, aby otrzymać pieniądze, których chcesz lub potrzebujesz, musisz zrozumieć politykę. Nie ma innej drogi. Jeśli otrzymujesz pieniądze, nie będąc wmieszanym w politykę, ktoś działa za Ciebie.

W przypadku kiedy nie ma wystarczająco dużo środków, istnieje ryzyko, że nie będziesz w stanie zrobić tego, co chciałbyś (to znaczy wykonać czegoś we właściwy sposób).

Właściwy sposób wykonania czegoś

Oto cel inżynierów. „Coś" oznacza cokolwiek, nad czym pracują, poświęcając swój czas. Inżynierowie wiedzą, czym jest „właściwy sposób wykonania", i często są sfrustrowani, kiedy nie pozwala się im czegoś wykonać w sposób, jaki uważają za odpowiedni.

Jeśli nie jesteś gotowy do gry, moja rada brzmi: „Przebolej to". Bez pieniędzy „we właściwy sposób" nie zrobisz niczego, bez uprawiania polityki nie zdobędziesz środków.

Jeżeli Twoja organizacja nie posiada zbyt wielu zasobów, po prostu nie masz szczęścia. Podobnie kiedy w Twojej organizacji trwa polityczna wrzawa, a Ty jesteś na etapie pozyskiwania środków lub ich akceptacji. Jeśli chcesz to zmienić, musisz nauczyć się polityki. To proste.

Kilka rad z dziedziny uprawiania polityki, które najbardziej cenię, pochodzi z książki *Jak zdobyć przyjaciół i zjednać sobie ludzi* autorstwa Dale'a Carnegiego (Studio EMKA). Polecam tę książkę każdemu, kto kiedykolwiek czuł się sfrustrowany jakąkolwiek formą interakcji z drugą osobą.

Maksyma 2.

Wiele lat temu pracowałem jako menedżer w dużej firmie. Mieliśmy problem z siecią: w związku ze słabym projektem występowało mnóstwo awarii, pojawiało się również wiele kwestii politycznych. Dysponowałem grupą bardzo bystrych inżynierów, którzy mieli pomysły na to, jak

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

poprawić zaistniałą sytuację. Wszyscy oni uważali, że ich propozycje zmian powinny zostać zaimplementowane, wyłącznie dlatego, że uważali je za słuszne (niektóre takie były, inne nie). Moim zadaniem było uświadomienie im, że nie mogą wykonywać takich zmian, jakie sami chcieliby.

Podjąłem wysiłek stworzenia trzech zasad, aby uzmysłowić im, że nie każda zmiana jest dobra, oraz zachęcić do analizowania implikacji zmian, które proponują. Zasady te były proste, jednowyrazowe. Dla mnie rozważanie dokonania zmiany usprawiedliwiała zgodność z jednym z trzech kryteriów:

- uproszczenie,
- standaryzacja,
- stabilność.

Maksyma 2.: Jedynymi właściwymi powodami wykonania zmiany w odpowiednio wyskalowanej sieci produkcyjnej są: uproszczenie, standaryzacja, stabilność.

Jeśli proponowana zmiana nie odpowiadała jednemu z tych trzech celów, nie brałem pod uwagę prośby o jej wprowadzenie. Na początku moi inżynierowi byli stale sfrustrowani. Chcieli efektów, a nie mogli zrobić tego, co uważali za konieczne. W przeciągu miesiąca jednak zaczęli rozumieć, a w ciągu 6 miesięcy sieć poprawiła się na tyle, że nie sprawiała już żadnych problemów. Stabilna sieć to stabilne usługi i dumni inżynierowie.

W perspektywie długoterminowej jest to jedyny sposób, aby osiągnąć stabilizację w niepewnym środowisku. Zastosowanie tych trzech kryteriów pozwoliło przejść z punktu, w którym sieć była przedmiotem kpin wyższej kadry zarządzającej, do miejsca, gdzie stała się ona powodem do dumy.

Oto omówienie poszczególnych zasad.

Uproszczenie

Zasada ta jest dość jasna. Aby było zgodne z tą regułą, proponowane rozwiązanie musi w jakiś sposób upraszczać sieć lub operacje w niej wykonywane. Przykładem jest usunięcie urządzenia lub zastąpienie złożonego schematu IP prostszym w zrozumieniu i utrzymaniu. Może zmiana ta wymusi ujednolicenie wersji oprogramowania na routerach. Wszystko, co sprawia, że operowanie siecią i rozwiązywanie jej problemów staje się prostsze, jest oszczędnością czasu i pieniędzy. Co proste, to dobre. Jeśli możesz z łatwością wyjaśnić, jak dana zmiana upraszcza sieć, istnieją duże szanse, że jest ona uzasadniona.

Standaryzacja

Efektywniejszą pod względem kosztów oraz łatwiejszą w zarządzaniu może uczynić sieć również standaryzacja. Gdy posiadasz 200 routerów tego samego modelu i podłączysz się do jednego z nich, będziesz wiedział, czego się spodziewać. Podczas gdy to może nie jest realne, posiadanie standardowych urządzeń dla każdej z funkcji oraz standardowego sposobu wdrażania usług **jest**. Możesz zastosować jeden typ routerów brzegowych, przełączników rdzenia w budynkach itd. Gdy implementujesz NTP, powinno ono być skonfigurowane wszędzie w ten sam sposób.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Standaryzacja pozwala na szybszy dostęp do części zapasowych, umożliwia również ewentualną wymianę na urządzenia pochodzące z mniej krytycznych obszarów.

Stabilność

Sposób poprawienia stabilności nie zawsze jest oczywisty. Jeśli wersja systemu IOS, ujednolicona na wszystkich bramkach VoIP, została zgłoszona jako możliwa przyczyna problemu, który jest w stanie wpłynąć na sieć, to kod ten powinien zostać uznany za niestabilny i należy wykonać jego aktualizację. Podobnie, jeśli istnieje urządzenie, które w przeszłości sprawiało problemy, należy je wymienić. Ten typ zmian jest uzasadniony, poprawiają one stabilność.

Maksyma 3.

Istnieje ograniczona liczba powodów, w związku z którymi firma pokryje koszty zmian istniejącej sieci. Powody te są zdumiewająco proste i dość oczywiste:

- niższe koszty,
- zwiększenie wydajności lub pojemności,
- zwiększenie niezawodności.

Dopóki nie udowodnisz, że Twój pomysł jest realizacją jednego z wymienionych celów, istnieją duże szanse, że spotkasz się z odmową przyznania Ci funduszy.

Maksyma 3.: Firmy wydają pieniądze tylko na projekty, które: obniżają koszty, zwiększają wydajność lub pojemność, zwiększają niezawodność.

Zbadajmy każdy z tych trzech elementów.

Niższe koszty

Pracowałem kiedyś dla firmy telekomunikacyjnej posiadającej łącza DS3 przecinające kulę ziemską. Pod względem projektowym sieć była perfekcyjna. Każdy z węzłów miał wiele punktów wejścia, opóźnienia były akceptowalne, odporność doskonała, a okresy niedostępności niezwykle rzadkie. Sieć była bliska osiągalnego ideału. Ja i mój zespół byliśmy dumni z pracy, jaką wykonaliśmy.

Pewnego dnia doszła do mnie informacja, że sieć musi zostać zmieniona. Dokładniej, należało zredukować budżet o jeden milion. Należało wybrać pomiędzy telekomunikacją a personelem. Witaj w świecie biznesu — powiedziano mi, że muszę zniszczyć swoją idealną sieć lub wskazać, kto musi odejść z pracy.

Wraz z zespołem postanowiliśmy przeprojektować sieć generującą 12 milionów kosztów rocznie tak, aby jej utrzymanie kosztowało 10 milionów, a ona sama była prawie tak dobra jak wcześniej. Nie byliśmy dumni, ponieważ kosztowało nas to wiele kompromisów, w które nie wierzyliśmy, ale organizacja uznała to za sukces.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kluczowe w tej historii jest to, że organizacja nie zyskała na odporności i pojemności sieci, mimo to zmiany zostały zatwierdzone, ponieważ obniżały one koszty utrzymania sieci o 2 miliony. Dla firmy rozwiązanie **wystarczająco dobre** było lepsze aniżeli **perfekcyjne**, dlatego że pozwalało rocznie oszczędzić 2 miliony.

Jeśli masz pomysł na zmiany w sieci, najlepszym sposobem na zdobycie aprobaty jest przedstawienie oszczędności, jakie wygeneruje ich wprowadzenie, szczególnie jeśli chodzi o redukcję stałych wydatków, jak na przykład: kosztów telekomunikacji, kosztów kontraktów związanych ze wsparciem, opłat związanych z konsultantami.

Choć mój przykład to zmiana raczej wymuszona aniżeli pożądana, istotne jest to, że dla firmy najwyższy priorytet mają koszty. Gdy masz pomysł na ograniczenie kosztów, są szanse, że zarządzający Cię wysłuchają. Z menedżerskiego punktu widzenia każdy inżynier chciałby wydawać. Pomysł niosący ograniczenie kosztów wyróżni Cię na tle innych.

Pamiętaj, że istnieją różne rodzaje kosztów, z którymi musi się zmagać biznes, ich przykładami są:

Koszty wyposażenia
: Sprzęt oraz okablowanie są przykładami kosztów wyposażenia (ang. *equipment costs*). Gdy jesteś w stanie zastąpić trzy routery dwoma, redukujesz koszty tego rodzaju.

Koszty stałe
: Koszty umów serwisowych, opłaty odprowadzane do dostawcy telekomunikacyjnego, jak np. rachunki za T1, rachunki telefoniczne i inne tego typu, stanowią koszty stałe (ang. *recurring costs*). Cięcie kosztów stałych ma dla firm bardzo istotne znaczenie.

Koszty zasobów ludzkich
: Twoje wynagrodzenie stanowi dla firmy koszt. Jeżeli w Twoim departamencie pracuje 10 osób zarabiających po 100 000 dolarów, a Ty zaprojektowałeś zmianę wymagającą 200 000 dolarów, która pozwoli zredukować zatrudnienie do dwóch osób, po mniej niż roku firmie zwróci się ta inwestycja.

Inżynierowie nie zawsze zdają sobie sprawę z kosztów ich utrzymania. Czy wiedziałeś na przykład, że kwota opodatkowania Twojej pensji to jedynie część tego, czego wymaga rząd? Jeśli zarabiasz 100 000 dolarów rocznie, opodatkowanie kosztuje firmę dodatkowo 25 000 dolarów. Nagle Twoje koszty to 125 tysięcy. Nie zapominajmy o innych obowiązujących pozycjach, jak plan emerytalny i opieka zdrowotna. Oba te składniki mogą kosztować kolejne 25 000 dolarów. Teraz kosztujesz firmę 150 tysięcy. Dostajesz premię? Dodaj ją do tej sumy. Nie zapomnij również, że przestrzeń biurowa także kosztuje. Firma musi też płacić za klimatyzację, Twój komputer, telefon stacjonarny i zapewne komórkowy. Podliczając wszystko, spokojnie możesz kosztować firmę dwukrotnie więcej, aniżeli wynosi Twoje wynagrodzenie.

Koszty konsultacji
: Koszty konsultacji (ang. *consulting costs*) są liczone inaczej niż koszty HR (dlatego właśnie czasami zwalniani są stali pracownicy, a na ich miejsce przyjmuje się konsultantów pracujących za tę samą lub wyższą stawkę). Jeśli jesteś w stanie przeprojektować sieć aktualnie utrzymywaną przez trzech konsultantów tak, aby wystarczył jeden, obniżasz koszty konsultacji.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zwiększenie wydajności lub pojemności

Zwiększenie wydajności najczęściej nie jest niczym złym. Wyjątkiem może być podnoszenie wydajności w przypadku, kiedy nie jest to potrzebne. Oddziały zatrudniające dziesięciu użytkowników prawdopodobnie nie potrzebują zapór sieciowych typu ASA 5580 zdolnych do obsłużenia ruchu 1 GB.

Z drugiej strony urządzenia, które według aktualnego trendu w krótkim czasie osiągną swoje maksimum, staną się niebawem problemem. Łącza telekomunikacyjne, które wykorzystują 70% swojej przepustowości, powinny zostać zmodernizowane. Routery, przełączniki, firewalle, których procesory utylizowane są na poziomie 50%, powinny zostać wymienione na silniejsze.

VoIP jest doskonałym przykładem na to, jak poważnymi problemami mogą być wąskie gardła w obszarze wydajności. Jeśli Twoja bramka będzie w stanie obsłużyć jedynie 50 równoczesnych połączeń, a w związku ze wzrostem obciążenia połączenia od klientów zaczną być odrzucane, bądź pewny, że znajdą się fundusze na modernizację bramki, która pozwoli obsłużyć większy ruch.

Bystrzy inżynierowie będą obserwować sieć oraz jej tendencje. Jeżeli będziesz w stanie zauważyć trend i przełożyć go na konkretną zmianę w następnym cyklu budżetowym, zostaniesz bardzo dobrze odebrany, ponieważ zmianę i jej budżet będzie można zaplanować z wyprzedzeniem. Pamiętaj, menedżerowie nie lubią niespodzianek. Postaraj się dostrzec te szanse, zanim pojawi się problem. Poinformowanie szefa o konieczności aktualizacji po tym, gdy nastąpi awaria, nie uczyni z Ciebie bohatera.

Zwiększenie niezawodności

Przełączniki modularne są bardziej niezawodne niż te o stałej konfiguracji. Dlaczego? Ponieważ mają one przeważnie redundantne zasilanie, a często też procesory. Niezawodność jest istotnym aspektem projektowania sieci. Pewniejsze komponenty, jak przełączniki modularne, bywają o wiele droższe od ich mniej niezawodnych odpowiedników, dlatego małe firmy z reguły ich nie wykorzystują.

W miarę jak małe firmy przekształcają się w średnie, a później w duże przedsiębiorstwa, dostrzegają, że nieplanowane niedostępności kosztują je więcej aniżeli wydatki na bardziej niezawodny sprzęt. Jest to punkt, od którego zaczynają przeznaczać środki na zwiększenie niezawodności.

Innym przykładem zwiększenia niezawodności jest zastąpienie pojedynczej zapory sieciowej ASA parą skonfigurowaną jako aktywny-oczekujący.

Niezawodność można łatwo sprzedać w przypadku, kiedy Twoja firma doświadczyła przerw związanych z pojedynczym punktem awarii, lub jeśli kadra zarządzająca miała styczność z tego rodzaju sytuacjami. Z pewnością istnieją menedżerowie zdający sobie sprawę, że posiadanie pojedynczego punktu awarii nie jest niczym właściwym, jednak ci, którzy przeżyli awarie, są bardziej podatni na wydatki związane z ich unikaniem.

Odwołanie się do jednej z powyższych przyczyn uczyni Twoją zmianę pożądaną, a połączenie pierwszej (niższy koszt) z którąś z kolejnych spowoduje, że pomysł prawie na pewno zostanie zrealizowany. Jeśli jesteś w stanie wyjść z propozycją łączącą wszystkie trzy, zarządzający będą Cię nosić na rękach i przeparadujesz holem niczym zwycięski bohater.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Z drugiej strony w przypadku, kiedy zdajesz sobie sprawę, co należy zrobić, aby zwiększyć wydajność i niezawodność, ale nikogo o tym nie powiadomisz, winą za ewentualną awarię zostaniesz obarczony Ty sam. Większość inżynierów wie, jak zadbać o sieć. Korzystając z tych porad, powinieneś być w stanie przekonać swojego szefa do zmian wartych implementacji.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ROZDZIAŁ 41.

Unikanie frustracji

W branży związanej z komputerami i sieciami pracuję od ponad 25 lat. W tym okresie byłem świadkiem wpadania ludzi we frustrację stale z tego samego powodu. Nauczyłem się wiele w ciągu całego tego czasu, ale niektóre lekcje nie przyszły łatwo. W tym rozdziale chcę Ci pomóc uniknąć niektórych frustracji, które prawdopodobnie czekają Cię na ścieżce kariery, a tym samym sprawić, że praca będzie sprawiała Ci więcej przyjemności.

Dlaczego wszystko jest takie zagmatwane

Nie wiem, jak często słyszałem zdanie: „Dlaczego wszystko jest takie zagmatwane?". Oczywiście język często bywa bardziej obrazowy, ale przyzwoitość nie pozwala mi na przytaczanie przykładów. Wystarczy powiedzieć, że wyrażenie „zagmatwane" zastępowane jest czymś bardziej satysfakcjonującym w chwilach frustracji i złości.

Sytuacja ta jest bardzo powszechna. Przedsiębiorstwo dochodzi do momentu, w którym potrzebuje niezawodnej sieci. Zarządzający uznają z jakichś powodów, że aktualna załoga nie jest w stanie przeprowadzić migracji w stronę odpowiedniego kształtu sieci, dzwonią więc po konsultantów. Wtedy przybywam ja, a witany jestem wyżej wspomnianym pytaniem. Czasami pochodzi od pracowników firmy, a czasami stawiane jest przez młodszych konsultantów, z którymi współpracuję.

Odpowiedź na nie opiera się na kilku prostych prawdach.

Prawda 1.: Firmy rozpoczynające działalność jako małe, a następnie doświadczające intensywnego wzrostu często mają skomplikowaną sieć, która nie jest w stanie spełnić stawianych jej wymagań.

Przekształcanie małych firm w duże często przebiega w bardzo szybkim tempie. Przedsiębiorstwa, które szybko rozwiązują problemy, to te, które aktualnie dobrze sobie radzą na rynku. Kiedy firmy są małe, zatrudniają specjalistów mających doświadczenie z serwerami. Profesjonaliści sieciowi nie są z reguły wymagani w małych przedsiębiorstwach, a ludzie od serwerów doskonale sobie radzą z utrzymaniem wszystkiego.

Kiedy małe firmy rosną, sieć również się rozwija. Doskonali inżynierowie będący ekspertami w dziedzinie skalowalności i projektowania serwerów niekoniecznie są też świetnymi fachowcami w obszarze sieci. Prowadzi to do sytuacji, w której sieć jest sklejana tak, aby jak najszybciej

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

rozwiązać problem, a nie ostrożnie planowana zgodnie z pełną wizją stanu docelowego. W żadnym razie nie jest to zarzut w stronę inżynierów od serwerów. Doskonale wykonują swoją pracę. Problem leży jednak w tym, że są oni inżynierami od serwerów, a nie sieci. Zapamiętaj, że małe firmy zwykle nie potrzebują inżynierów sieciowych. Wielu pracodawcom trudno zaakceptować to, że w miarę rozwoju muszą takich zatrudnić.

Prawda 2.: Sieci są częściej składane niż projektowane.

Sieć postrzegana jest zazwyczaj jako nic więcej jak tylko środek umożliwiający połączenie stacji roboczej z serwerem bądź internetem. Szczególnie w małych środowiskach sieć odgrywa mniejszą rolę niż kwestie związane z serwerami i innymi komputerami. Odnosi się to do prawdy 1., lecz jest też skutkiem niezrozumienia przez zarządzających znaczenia sieci. Spotykałem menedżerów mówiących w ten sposób: „Po prostu włóż serwer do tej zdalnej lokalizacji z VPN". Dla nich słowo VPN oznacza „magiczne coś, co zastępuje prawdziwą sieć".

Infrastruktura sieciowa jest często postrzegana w tych samych kategoriach co budynki, krzesła i inne tego typu elementy. Oznacza to, że nie jest traktowana jako źródło przewagi konkurencyjnej, a raczej jako konieczność konsumująca kapitał. Nie stanowi to źródła zysku, po co więc trwonić na to pieniądze? W ten sposób myśli spora część kadry zarządzającej i tak właśnie patrzy na sprzęt oraz zespół związany z siecią. Podejście to powszechnie występuje w małych firmach niezorientowanych technicznie, ale na pewno w wielu innych miejscach również. W związku z takim podejściem do sieci nie przykłada się także uwagi do szkoleń z tego obszaru. Ludzie poświęcający swoje kariery technologiom serwerowym wybierają kursy związane z serwerami, a nie te kształcące w tematyce sieciowej.

Prawda 3.: Projekty dotyczące sieci nie zajmują wysokiej pozycji na liście priorytetów kadry zarządzającej.

Wyższa kadra zarządzająca małych i średnich firm często nie rozumie, jakie znaczenie może mieć źle zaprojektowana sieć. Zazwyczaj kwestie sieci nie wzbudzają emocji, podczas gdy wieści o nowych wirusach oraz dziurach w bezpieczeństwie serwerów trafiają do CNN oraz na pierwsze strony „**New York Timesa"**. Kiedy sieć zawodzi, prawda dociera do menedżerów, jest już jednak za późno, wymagane są poważne zmiany. Oto moment, kiedy pojawiają się konsultanci.

Poświęciłem wiele tygodni, projektując schematy IP dla dużych firm. Projekt dla jednego z klientów był tak imponujący, że można śmiało powiedzieć, iż było to dzieło. Skalowalność rozwiązania zapewniała dziesiątki lat funkcjonowania, przy czym elastyczność była niesłychana. Projekt był tak imponujący, że tylko dwóch z nas, twórców rozwiązania, rozumie go do dzisiaj. O ile mi wiadomo, schemat ten nadal funkcjonuje. Założenia poczynione podczas projektowania, podobnie jak zasady grawitacji, nie muszą być zrozumiałe, aby tę sieć utrzymać.

Przesłanie tej historii jest jednak takie, że tygodnie spędzone nad rozwiązaniem sporo kosztowały klienta, zysk nie był oczywisty, a czas spędzony nad projektem dyskusyjny. Korzyści, jako trudne do oszacowania, nie łatwo było wykazać. Obliczenie czasu, który zostanie zaoszczędzony w związku z brakiem konieczności renumeracji serwerów, jest niemożliwe. Ile czasu

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

i wysiłku uniknie się na skutek tego, że do serwerów tego samego typu będzie można odwołać się pojedynczym wpisem na liście dostępu? Co z lokalizacją? Krajami? Zaprojektowaliśmy sieć, w której każdy kawałek sieci mógł stać się częścią większego fragmentu. Przy zmianie całej lokalizacji należało zmienić jedynie drugi oktet. Cały ten zaoszczędzony czas był jednak **ukryty**. Nie można opierać się na potencjalnych oszczędnościach. Wszystko to czyni schemat IP trudnym do sprzedania kadrze zarządzającej. Jednak w przypadku pracowników, którzy codziennie pracują z siecią, dobry schemat IP faktycznie wiele upraszcza.

Prawda 4.: Inżynierowie nie są biegli w komunikowaniu swoich potrzeb kadrze zarządzającej.

Spojrzenie kadry zarządzającej na sieć określa kombinacja ich własnego doświadczenia oraz informacji otrzymanych od inżynierów. Jeśli inżynierowie nie potrafią porozumiewać się z przełożonymi, rzutuje to na percepcję kadry zarządzającej. To swoiste dopełnienie innych kwestii, które poruszyłem. Kiedy inżynier wie, co powinno być zrobione, ale nie potrafi uzmysłowić tej potrzeby menedżerowi, zazwyczaj menedżer nie przeznaczy środków na proponowane rozwiązanie.

Pewnego razu pojechałem do klienta, aby rozwiązać problem związany z komunikacją klient-serwer. Klientem była agencja reklamowa, która pracowała z bardzo dużymi zdjęciami. Pojawiły się skargi związane z tym, że pliki graficzne o rozmiarze 80 MB bardzo długo przenosiły się z serwera na stację roboczą (było to w późnych latach 90., gdy pliki graficzne o rozmiarze 80 MB były niespotykane). Problem był wynikiem istnienia zbyt wielu urządzeń w sieci wykorzystującej huby. Zaproponowałem wymianę co najmniej głównego huba na przełącznik. Wtedy jednak przełączniki były drogie i tylko nieliczne firmy stać było na takie zmiany.

Kadra zarządzająca odmówiła dokonania zmian w sieci i poleciła mi szukać dalej. Uważali, że przyczyna może tkwić w serwerach, ponieważ był to obszar, który rozumieli. Nie mieli pojęcia, jak działa sieć, oraz nie zdawali sobie sprawy z jej ograniczeń.

Zrobiłem menedżerom krótką lekcję na temat działania sieci i wyjaśniłem, skąd się bierze problem. Zagwarantowałem im, że jeśli wymiana huba na przełącznik nie pomoże, będą mogli oddać urządzenie bez ponoszenia dodatkowych kosztów. Switch okazał się właściwym rozwiązaniem, a ja zamieniłem stałego klienta w lojalnego klienta.

Małe firmy lubią, kiedy każdy skupia się na tym, w czym jest dobry. W tym przypadku klient specjalizował się w reklamie. Nie dbał o komputery i sieć, ale je rozumiał, ponieważ używał ich w codziennej pracy. Na co dzień jednak nie miał związku ze sprzętem sieciowym, pracował, dopóki nie pojawiły się problemy.

Prawda 5.: Czasami problemem jest to, że kogoś problem nie obchodzi albo ów ktoś nawet o nim nie wie.

Ludzie od reklamy nie powinni musieć uczyć się, jak działa przełącznik. Sieć jest po to, aby im ułatwić pracę.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Jak przedstawiać swoje pomysły kadrze zarządzającej

Wpadłeś na wspaniały pomysł, który według Ciebie coś ulepszy. Może dotyczy to sieci, a może sposobu zatrudniania i rekrutowania ludzi do firmy, w każdym razie wiesz, że warto to zrobić. Problem jednak w tym, że nikt z kadry zarządzającej nie chce Cię wysłuchać. Według nich marnujesz ich czas. Kto ma więc rację? Zarówno oni, jak i Ty.

Dobry menedżer chce wiedzieć, o czym myślisz, i znajdzie czas, aby Cię wysłuchać. Jeśli Twój pomysł okaże się kiepski, otrzymasz jasną informację zwrotną oraz pomoc w weryfikacji pomysłu. Inni menedżerowie obawiający się niespodziewanego wpływu jakichkolwiek zmian nie zrobią nic. W przypadku gdy Twój menedżer należy do tego drugiego rodzaju, możesz albo starać się go przekonać, że zmiana przyniesie korzyści (przynajmniej dopóki stanowi realizację Twojego pomysłu), albo zupełnie go pominąć. W drugim przypadku radzę poprosić o pozwolenie lub przynajmniej poinformować swojego szefa o tym, że zamierzasz odwołać się do wyższych instancji. Nikt nie lubi być pomijany. Knuciem „na boku" możesz sobie wyłącznie przysporzyć wrogów.

Inżynierowie mają głowy pełne wspaniałych pomysłów. Zaprojektowani jesteśmy, aby ulepszać. Jednak wydaje się, że w naszym przypadku istnieje odwrotna relacja między wspaniałymi pomysłami a zdolnościami komunikacyjnymi. Sprowadza to na nas wiele kłopotów.

Podczas jednego ze spotkań zwróciłem uwagę wszystkich, ogłaszając, że cała sieć jest na granicy awarii i że zapłacimy miliony odszkodowań w związku z umowami SLA, których warunków nie spełnimy. W pokoju panowała napięta atmosfera, a moja wypowiedź wcale jej nie rozładowała. Cała uwaga skupiła się na mnie, a ja byłem gotowy, aby zrobić ruch. Wyjaśniłem długiemu stolikowi zasiadających tam dyrektorów, jak ich błędne decyzje przyczyniły się do aktualnego bałaganu. Wskazałem konkretne osoby, obarczając je winą. Wyrzuciłem z siebie wszystko. Byłem ich bohaterem i tylko ja mogłem uratować świat. Przekaz pochodził od potężnego i wszechmocnego boga inżynierów i biada temu, kto nie chce słuchać.

Moja krytyka była bezwzględna. Konkretni ludzie okazali się krótkowzroczni, oszczędności doprowadziły do awarii, a to z kolei do złości wśród klientów. Mógłbym to naprawić, gdyby tylko wszyscy się zamknęli i pozwolili mi robić to, do czego zostałem zatrudniony.

Czy nie brzmi to znajomo?

Zdziwiło mnie, że nie uwierzono w moje rekomendacje, a kadra zarządzająca nie rzuciła mi się do stóp, nie szczędząc prezentów (lubię Ferrari). Po tym, gdy skończyłem swoje tyrady, jeden z dyrektorów odciągnął mnie na bok i wyjaśnił, że nie tylko zrobiłem z siebie kompletnego idiotę, ale również zaprzepaściłem kredyt zaufania, jaki mogłem mieć u każdej z osób w tym pokoju.

Od tamtego czasu wiele się nauczyłem. Na szczęście na swoich błędach mogę uczyć innych. Co zrobiłem źle? **Wszystko**. Po pierwsze i najważniejsze: nie okazałem szacunku. Nie mam na myśli tego wynikającego z pozycji, w to nie wierzę. Chodzi o szacunek względem drugiego człowieka. Jeśli ktoś kiedykolwiek mówiłby do mnie w sposób, w jaki ja przemawiałem w pokoju pełnym dyrektorów, wstałbym i wyszedł. Czułem się jednak niczym ich ojciec, a tym samym upoważniony do takiej formy przekazu. Poniosły mnie również nieco emocje. Spojrzałem na grupę i zacząłem swoją tyradę. W świecie biznesu taka droga do niczego nie prowadzi.

Zatem jak zwracać się do kadry zarządzającej? Jak zamienić wściekłość rozsadzającą inżynierów w użyteczny język biznesowy, który będzie w stanie przynieść rezultaty? Oto kilka porad:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Dokumentuj swoje pomysły

Jeśli Twoje pomysły są trafne, Twój szef zechce mieć coś, co będzie mógł pokazać swoim przełożonym. Powinien być to krótki dokument przedstawiający propozycję wraz z korzyściami jej wdrożenia. Dobrze, by rozpoczynał się od streszczenia zawierającego minimum żargonu technicznego, a dalej zawierał kilka stron treściwych technicznych wyjaśnień opisujących również szczegóły. Umieść tam terminy oraz budżet — nawet jeśli mają to być jedynie oszacowania. Być może napisanie czegoś takiego stanowi dla Ciebie kłopot, poproś wtedy o pomoc kogoś, kto to potrafi. Alternatywą mogą być slajdy Power Pointa. Dlaczego też nie przygotować zarówno dokumentów, jak i slajdów?

Bądź poprawny gramatycznie

Produkując dokumentację, zadbaj o to, aby wszystko, co piszesz, było gramatycznie poprawne. Stosuj sprawdzanie pisowni oraz przeczytaj dokument **co najmniej dwukrotnie**. Następnie niech ktoś inny dokona korekty, po czym ponownie przeczytaj dokument, nanosząc ewentualne poprawki. Będziesz zdumiony, ile błędów znajdziesz nawet przy trzecim czytaniu.

Miałem przyjemność pracy z bardzo mądrymi ludźmi. Osobami, które posiadały tytuły CCIE, tytuły profesorskie, były nauczycielami, prawnikami, doktorami, naukowcami, takimi, które były dużo bardziej wykształcone niż ja. Zdumiewa mnie, jak nieczytelną dokumentację tworzą ludzie posiadający te wszystkie stopnie. Jeśli chcesz być odbierany poważnie, upewnij się, że Twoja dokumentacja jest poprawna gramatycznie i wygląda profesjonalnie.

Pozbądź się emocji

Każda emocja towarzysząca przekazywanej przez Ciebie wiadomości zwróci się przeciwko Tobie. Jedynym możliwym wyjątkiem od tej reguły będzie pasja. Uczucie pasji w stosunku do swoich pomysłów jest czymś właściwym. Miej ją jednak na oku, ponieważ może bardzo łatwo zamienić się w gniew albo frustrację.

Jeśli jesteś podobny do mnie, czujesz pasję w stosunku do rzeczy, w które wierzysz. Inżynierowi łatwo przychodzi wiara w pewien koncept lub ideę. Wszyscy wiemy, że naśladowanie pewnych dobrych praktyk ułatwia życie, jednak spotkałem inżynierów, którzy zostali wytrąceni z równowagi, ponieważ ktoś z kadry menedżerskiej zatwierdził tylko jeden z dwóch serwerów DNS. Spokojny inżynier byłby w stanie omówić zalety i wady redundantnego serwera DNS i być może przekonać menedżera do zmiany zdania. Nadpobudliwy inżynier zostanie zwyczajnie zignorowany. Kiedy angażujesz emocje w prezentację, tracisz wiarygodność.

Lubię się spierać. Nie znaczy to, że krzyczę na swoją żonę, podczas gdy ona ciska we mnie naczyniami. Lubię debatować z oponentem na temat propozycji, która różni się od mojej. Jeśli spieram się z Tobą, a Ty zaczynasz okazywać emocje, praktycznie wygrałem. Czemu? Ponieważ kiedy pęcznieją emocje, rozum ustępuje. Im bardziej emocjonalny jesteś, tym mniej rozsądny.

Kiedy dajesz się ponieść emocjom podczas prezentacji, wszyscy dookoła uświadamiają sobie, że nie jesteś osobą racjonalną. Irracjonalni ludzie w najlepszym przypadku są ignorowani. Osoby z pasją bywają czczone. Gdzie zatem przebiega ta granica? Ekscytacja własnym pomysłem jest czymś dobrym, ale agresja w obliczu kontrargumentów już nie.

Bądź uprzejmy

Przede wszystkim bądź uprzejmy. Jeśli taki nie jesteś, oznacza to prawdopodobnie, że pozwoliłeś sobie na emocje. Niezależnie od powodu brak uprzejmości z Twojej strony

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

oznacza brak szacunku ze strony innych. Ludzie nie będą traktować Cię z respektem, dopóki Ty ich nie będziesz nim darzył. Uprzejmość jest formą szacunku, który powinien cechować stosunki z każdym. W przypadku przełożonych brak szacunku, szczególnie podczas przedstawiania nowego pomysłu, na pewno nie poskutkuje upragnionymi rezultatami.

Bądź zwięzły

Menedżerowie zwykle nie mają zbyt wiele czasu, dlatego czas, który Ci poświęcą, powinien być wykorzystany rozważnie. Wyczekiwać będą podsumowania Twojego pomysłu oraz informacji o zyskach, jakie on wnosi. Jeśli ich to zainteresuje, będziesz mógł przejść do szczegółów, zaczynając jednak od wzrostu sprzedaży. W przypadku gdy Twój pomysł wzbudzi zainteresowanie, będziesz miał mnóstwo czasu, aby wdawać się w szczegóły techniczne.

Znajdź wady swojego pomysłu

Żaden pomysł nie jest idealny. Jeśli nie znasz możliwych argumentów przeciwko Twojej propozycji, ktoś inny je znajdzie. Mogą zostać wykorzystane przeciw Tobie podczas prezentacji. Profesjonaliści znają ograniczenia oraz wady swoich rozwiązań, mają przy tym przygotowane odpowiedzi na pytania o nie.

Punkt ten jest tak często pomijany, że możesz zrobić wrażenie na swoich dyrektorach, uprzedzając ich uwagi.

Zaakceptuj „nie" jako odpowiedź

Nie zawódź i nie narzekaj, kiedy otrzymasz odmowną odpowiedź. Będzie to nieprofesjonalne i zniechęci zarządzających do rozważania Twoich propozycji w przyszłości. Czasem nie ma wystarczająco dużo środków albo czas jest nieodpowiedni. Czasem też odpowiedź może brzmieć „nie" z powodów, o których nie można Ci powiedzieć. Niezależnie od przyczyny na godnym zaakceptowaniu klęski zyskają Twoje przyszłe prośby. Oczywiście dopuszczalne jest zapytanie, dlaczego pomysł został odrzucany, ale nie należy naciskać na odpowiedź.

Omiń bezpośredniego przełożonego

Czasami gdy proponujesz pewien pomysł, zostajesz spławiony. Może Twój szef Cię nie lubi, a może jest przekonany, że każdy Twój pomysł zawiedzie. Niezależnie od tego możesz uważać, że Twój pomysł powinien zostać wysłuchany i powinieneś mieć szansę, aby go przedstawić. W tym przypadku właściwą drogą jest poproszenie menedżera o szansę zaprezentowania swojego pomysłu jego przełożonemu. Odpowiedź może Cię zaskoczyć — menedżer może być zadowolony z przekierowania Cię do swojego szefa. Może być jednak tak, że się nie zgodzi, a praca z menedżerem, który boi się jakichkolwiek zmian, może nie być niczym przyjemnym. Skierowanie się od razu do bezpośredniego przełożonego może z kolei nie być politycznie najmądrzejsze, ale jeśli czujesz, że Twoje pomysły są dławione, stanowi to jedyną drogę.

Bez niespodzianek

Nie ogłaszaj na spotkaniu, że za 18 minut sieć stanie się niedostępna. Powiedz o tym swojemu szefowi z wyprzedzeniem. Właściwie powiadom swojego przełożonego i pozwól na zebraniu ogłosić to właśnie jemu. Jednym z pewnych sposobów na zdobycie wrogów w świecie biznesu jest krótkowzroczność podczas wystąpień przed innymi.

Co menedżerowie, dyrektorzy i inni członkowie kadry zarządzającej potrzebują wiedzieć? Niewiele ponad to, ile czasu musisz spędzić nad rozwiązaniem, jak długo będzie trwało wdrożenie oraz jaki jest szacowany koszt. Jeśli planujesz wydawać środki, które nie zostały ujęte w budżecie, lepiej żebyś miał solidne uzasadnienie, w innym przypadku nie dostaniesz ani grosza.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Najwyższa kadra menedżerska nie chce słuchać o tym, że nie możesz wykonywać swojej pracy, ponieważ nie masz przy swoim biurku przełącznika Nexus 7018. Zależy im, aby ich szefowie byli zadowoleni (najczęściej udziałowcy), chcą wiedzieć, co robisz, aby wzmocnić ich pozycję, oraz że nie sprawiasz zbyt wielu problemów.

Miałem kiedyś szefa, który wobec inżynierów stosował m.in. następującą regułę: „Jeśli coś jest fajne, nie możesz tego mieć". Jej (zazwyczaj nietrafnym) uzasadnieniem było to, że jeśli uznajemy coś za „fajne", to nie ma to realnej wartości biznesowej. Niestety po dziś dzień stosuję tę zasadę w stosunku do siebie.

Kiedy pytam inżynierów, po co im rzecz, o którą proszą, ich odpowiedź jest albo bardzo emocjonalna, albo mocno techniczna. Wyższa kadra zarządzająca nie dba o to, czy uszczęśliwią Cię nowe routery za 400 tysięcy dolarów, niezależnie od tego, jakie problemy one rozwiązują. Jednak kiedy udowodnisz, jak uczynią one sieć bardziej odporną i że dzięki temu firma zaoszczędzi wiele na niezrealizowanych umowach SLA, widownia Cię pokocha.

Zatem, co chce usłyszeć Twój szef? Zastanów się nad następującymi pytaniami:

- Jakie oszczędności przyniesie Twój pomysł? Poniżej znajdują się przykłady:

  Wydatki inwestycyjne (ang. *capital expenditure*)

  Zastąpi on wiele innych elementów, które i tak musiałyby zostać zastąpione.

  Wydatki operacyjne (ang. *operating expenditure)*

  Zredukowane zostaną koszty utrzymania łączy oraz koszty personalne.

  Zwrot z inwestycji (ang. *return on investment*)

  Wydanie środków teraz zwróci się na przestrzeni X miesięcy.
- Jak Twój pomysł zwiększy przychody?
- Jak Twój pomysł zwiększy zyski?
- Jakie korzyści Twój pomysł przyniesie wizerunkowi zespołu (a tym samym jego szefowi)?

Właściwe odpowiedzi na kilka z tych pytań spowodują, że Twoja prezentacja przyciągnie uwagę.

Kiedy i dlaczego aktualizować

Podnoszenie wersji oprogramowania routerów i przełączników nie jest niczym przyjemnym, ale nadchodzi moment, że trzeba to zrobić. Niektórzy lubią mieć najnowsze wersje oprogramowania i systemów operacyjnych. W przypadku sprzętu sieciowego nie jest to dobry pomysł, co postaram się tutaj udowodnić.

Może wydawać się oczywiste, że kiedy pojawi się nowa funkcjonalność, powinieneś jak najszybciej dokonać aktualizacji, tak aby z niej skorzystać. W rzeczywistości jest to jednak rzadkość. Zdarza się to tylko w małych sieciach, gdzie nie występuje kontrola zmian oraz brak jest konsekwencji związanych z niedostępnościami.

Do aktualizacji należy podejść poważnie. System działający na Twoim przełączniku lub routerze to tylko oprogramowanie, a ono może zawierać błędy, włączając w to luki w bezpieczeństwie. Mimo mechanizmów kontroli wersji oraz beta testów wykrywanych jest mnóstwo błędów, czego dowodzą ich zestawienia publikowane na stronach dostawców.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

A co z modernizacją pamięci? Dość łatwe, prawda? Może być to tylko wymiana kości DRAM SIMM lub też dodanie nowej, ale komplikacji nigdy nie można wykluczyć.

Zagrożenia aktualizacji

Oto niektóre przyczyny, w związku z którymi należy z ostrożnością podchodzić do kwestii aktualizacji:

Pojawienie się nowych błędów (ang. *bugs*)

Upgrade może pomóc w wyeliminowaniu dręczących Cię błędów, jednak nowe oprogramowanie może zawierać kolejne błędy. Najlepszym sposobem uniknięcia tego problemu jest poddanie nowego oprogramowania wyczerpującej analizie. Firma Cisco podaje listę błędów wraz z odsyłaczami oraz zalecenia, która wersja oprogramowania powinna być zainstalowana na danym sprzęcie. Jeżeli problemem okazuje się doprowadzenie do zgodności z zaleceniami Cisco, dobrze jest przynajmniej poinformować o tym e-mailem swojego szefa, przesyłając mu rekomendację producenta zalecającą zmianę. Może stwierdzi: „Wow, powinienem się temu przyjrzeć!".

Problemy ze sprzętem

Modernizacja sieciowego urządzenia zazwyczaj nie jest niczym skomplikowanym, najczęściej wiąże się jednak z restartem. Restartów nie przeprowadza się zbyt często (na szczęście), a podczas nich zawsze może pojawić się błąd lub mieć miejsce jakaś nieoczekiwana katastrofa.

Obserwowałem kiedyś prostą modernizację pamięci jednej z dużych stron typu *e-commerce*, która bardzo źle działała. W modularnym przełączniku 6506 wymieniano starą pamięć SIMM na nową. Kontrola zmian zaaprobowała działania. Czas niedostępności oszacowano na 15 minut. Okno przewidziane na zmianę (czas na wykonanie przez nas roboty): 4 godziny. Rzeczywisty czas niedostępności: **7 godzin**. Najprawdopodobniej odrobina kurzu wylądowała na gnieździe pamięci podczas wyciągania starej karty SIMM. Spowodowało to sytuację, w której nie wszystkie piny stykały. Ponowne umieszczenie karty nie pomogło. Wycofanie zmiany również nie wchodziło w grę, ponieważ stara karta używała tego samego gniazda. Wniosek z tego zdarzenia jest taki, że każda zmiana związana ze sprzętem powinna być traktowana jako znacząca zmiana (ang. *major change*), której nie należy lekceważyć.

Błędy ludzkie

Jeżeli zaangażowani są ludzie, błędy pojawiają się dużo częściej. Ludzkie błędy, niezależnie od tego, czy wynikają z literówek, ze zmęczenia, bezmyślności, czy niezdarności, są w stanie wyrządzić więcej szkód niż cokolwiek innego. Pewnego razu męczyłem się już od 3 godzin, wykonując zmianę zaplanowaną na 4 godziny. Zbliżał się czas, aby ją wycofać, niczego jednak nie udało mi się zrobić. Instalowaliśmy łącze T1, co nie było skomplikowanym zadaniem, mimo to wszystko szło źle. Okazało się, że problem wynikał z błędu ludzkiego. Moje instrukcje wskazywały na konfigurację portu oznaczonego numerem 4 jako portu dla T1 na 8-portowym panelu. Dokumentacja jednak nie mówiła o tym, że istnieją dwa takie panele. Przez cały ten czas próbowałem konfiguracji na niewłaściwym. Poleciłem wyłączyć cały system telefoniczny i przez 3 godziny cała firma pozbawiona była możliwości prowadzenia rozmów telefonicznych. Można to nazwać prawdziwą praktyczną lekcją. Mam nadzieję, że mój przypadek Ci wystarczy, choć wiem, że niektórzy uczą się jedynie na własnych błędach.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zmiana domyślnego zachowania w oprogramowaniu

Przez lata Cisco doprowadziło do tego (i słusznie), że pewne komendy powinny stanowić domyślne zachowanie routerów Cisco — na przykład: `ip subnet-zero` lub `ip class less`. Jeżeli zaprojektujesz swoją sieć tak, że funkcjonalności te będą wyłączone, to w jakiś sposób zależny od tak skonfigurowanych funkcjonalności po dokonaniu aktualizacji niektóre sieci mogą nagle przestać działać.

Powody do aktualizacji

Poniżej przedstawiłem powody uprawniające do dokonywania czynności aktualizacji:

Rozwiązanie problemu

Istnieją duże szanse, że router, który należy zrestartować po każdym wydaniu polecenia `show arp`, zacznie lepiej działać po dokonaniu aktualizacji. Określenie, czy na pewno jest to kwestia błędu oprogramowania, leży po stronie technicznego wsparcia, choć listy błędów dostępne są dla każdego w sieci internet.

Wycofanie oprogramowania lub sprzętu

Dostawcy wspierają swoje rozwiązania tylko do pewnego momentu. Jeżeli Twoje oprogramowanie lub sprzęt osiągnęły status *end-of-life*, przestanie ono być wspierane. Gdy wystąpią jakieś problemy, trudno będzie je rozwiązać. Brak wsparcia dla sprzętu generuje ryzyko dla operacji obsługiwanych przez sieć. W takich sytuacjach należy dokonać modernizacji.

Potrzebna jest nowa funkcjonalność

Zauważ, że nie napisałem: „Chcielibyśmy nowej funkcjonalności". Czasami nowe możliwości są niezbędne. Pracowałem kiedyś z dużą siecią z interfejsami DS3 opinającymi glob. Mieliśmy ich prawie 40. Kiedy jedna z tras migotała kilka razy w ciągu kilku minut, siało to spustoszenie w całym środowisku VoIP. Sieć wspierała ponad 300 połączeń na sekundę, więc migotanie stanowiło poważny problem. Rozwiązaniem okazało się tłumienie tras (ang. *IP route dampening*) dostępne w ostatniej wersji kodu (był to protokół EIGRP, a nie BGP, stąd brak tłumienia, którego EIGRP nie ma). Normalnie nigdy nie zaleciłbym wgrania najnowszej wersji systemu, jednak w tym przypadku, gdy było to niezbędne dla zachowania ciągłości operacyjnej, postanowiliśmy skorzystać z takiego właśnie systemu. Na szczęście nowa funkcjonalność rozwiązała nasz kłopot, nie przysparzając innych problemów.

Zmiana technologii

Jeżeli korzystałeś z T1, a teraz posiadasz DS3, prawdopodobnie należy dodać interfejs DS3, aby móc z niego korzystać.

Zwiększenie wydajności

Czasami firmy rozwijają się tak szybko, że pojawia się konieczność zakupienia wydajniejszego sprzętu. Widziałem strony WWW, które przechodziły z PIX 515 do 520, następnie do 525 i dalej do modułu Firewall Services Module tylko w związku ze wzrostem wymagań co do pasma, którego słabsze urządzenia nie były w stanie obsłużyć.

Uproszczenie

Jestem wielkim fanem kart WIC z wbudowanym interfejsem CSU/DSU. Jeśli posiadasz łącze T1, które nadal korzysta z CSU/DSU podłączonego kablem V.35, możesz wyeliminować zbędne połączenia, integrując interfejs CSU/DSU z routerem. Istnieje wiele podobnych sytuacji, a to jest tylko jeden z przykładów.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Zwiększenie niezawodności

Jeżeli masz tylko jeden firewall, dodanie kolejnego, skonfigurowanego jako redundantny, zwiększy niezawodność całej sieci. Wysoka dostępność jest zawsze mile widziana i powinno się o nią zadbać, jeśli pozwala na to budżet.

Obniżenie kosztów

Jeżeli potrafisz dowieść, że aktualizacja sprzętu lub oprogramowania obniży koszty, wtedy jest ona jak najbardziej opłacalna. Przykładem jest połączenie przełącznika i routera w jeden wielowarstwowy switch 3750. Połączenie dwóch urządzeń powoduje, że potrzebujesz kontraktu na wsparcie tylko jednego urządzenia, co z kolei obniża roczne koszty utrzymania.

Dlaczego kontrola zmian jest Twoim przyjacielem

Jeżeli pracowałeś w małej firmie, która przekształciła się w dużą, lub jeśli przeszedłeś z małego do dużego przedsiębiorstwa, zapewne spotkałeś się z procesem kontroli zmian. **Kontrola zmian** oznacza, że przedsiębiorstwo ogranicza zmiany związane z siecią (lub czymkolwiek innym), dopóki nie zostaną w pełni zrozumiane oraz zaplanowane. Wykonanie modernizacji na routerze, który jest objęty kontrolą zmian, generuje konieczność zgłoszenia żądania zmiany. Żądanie takie musi zostać zaaprobowane i dopiero wtedy ustala się termin wykonania prac. W niektórych firmach zmiany akceptuje lub odrzuca jedna osoba, w innych robią to specjalne komitety. Spotkałem się z organizacjami, w których komitety te zbierały się raz w tygodniu, np. we wtorki. W przypadku kiedy Twoja zmiana została odrzucona, musiałeś wyczekiwać kolejnego wtorku, aby komitet rozważył Twoją alternatywną propozycję.

Planowanie zmian dla inżynierów jest często czymś denerwującym. Dlaczego musisz czekać na to, aż grupa ludzi, którzy nie mają pojęcia o Twojej zmianie, pozwoli Ci ją przeprowadzić? Po co marnować czas, jeśli wiesz, że możesz pozbyć się problemu, wydając jedną komendę, która do niczego złego nie doprowadzi? Dla inżynierów kontrola zmian wydaje się marnowaniem czasu i energii, stojąc w sprzeczności z ich zdolnością do wykonywania dobrej roboty.

Przez lata swojej pracy piastowałem wiele stanowisk, poczynając od młodszego inżyniera do dyrektora dużej konsultingowej firmy oraz szefa własnego przedsiębiorstwa. Podczas tych lat nauczyłem się kilku rzeczy na temat tego, jak działa biznes i jak myślą inżynierowie. Prowadzenie biznesu i bycie jednocześnie inżynierem pozwoliło mi posiąść pewne unikalne spojrzenie.

Kontrola zmian jest istotna z prostego powodu: niepowodzenie generuje koszty. Weźmy na przykład dużą firmę typu *e-commerce*. Jeżeli inżynier wykonuje zmianę wymagającą niedostępności strony, firma traci w tym czasie przychód. Co więcej, może ucierpieć reputacja strony w związku z jej niedostępnością dla klientów. Większość użytkowników zapewne nie będzie czekała dłużej niż 10 sekund na załadowanie się strony, jeśli będą mogli dokonać zakupów gdziekolwiek indziej. Wyobraź sobie konsekwencje niedostępności strony przez godziny w związku z jakimś prostym błędem!

Kontrola zmian ma za zadanie zapobiegać temu typowi niedostępności poprzez ostrożne planowanie i metody akceptacji zmian. Niemożliwe jest osiągnięcie wskaźników dostępności na poziomie 99,999% bez odpowiedniego procesu kontroli zmian.

Błędy ludzkie są obecnie najczęstszą przyczyną awarii. Kontrola zmian na wiele sposobów ogranicza właśnie tego typu błędy. Radzenie sobie z kontrolą zmian uczyni Cię lepszym inżynierem, który będzie potrafił skuteczniej porozumiewać się z kadrą menedżerską. Spójrzmy na niektóre z zalet kontroli zmian:

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Kontrola zmian uczy przewidywania

Odpowiedni proces kontroli zmian zobowiąże Cię do opisania w szczegółach każdego kroku planowanej pracy, włączając w to wszystkie komendy, jakie wprowadzisz na każdym z urządzeń, którego dotkniesz. Wyszczególnianie czynności wymusza zastanowienie się z wyprzedzeniem nad tym, co zamierzasz zrobić. Stanowi to również doskonały sposób na zrozumienie, jak działają switch/router/firewall, ponieważ zobowiązany jesteś określić wpływ zmiany na pozostałe komponenty.

Kontrola zmian poprawia umiejętności tworzenia dokumentacji

Program kontroli zmian przeważnie zawiera etap dokumentowania proponowanej zmiany, zanim zostanie ona wdrożona. Wielu inżynierów nie lubi tworzenia dokumentacji i za wszelką cenę stara się go unikać. Najlepszym sposobem na poradzenie sobie z tym jest regularne dokumentowanie. Zespół ludzi, począwszy od inżynierów aż po menedżerów i dyrektorów, prawdopodobnie przejrzy Twoją dokumentację związaną ze zmianą. Jeśli nie jest ona zwięźle i jasno napisana, będziesz musiał odpowiedzieć na wiele pytań i docelowo również poprawić dokumentację.

Kontrola zmian poprawia umiejętność komunikacji

Kontrola zmian z reguły obejmuje proces przeglądu żądania zmiany, podczas którego zmiana jest oceniana, komentowana, a później akceptowana lub odrzucana. Musisz przedstawić techniczne aspekty osobom, które mogą nie posiadać odpowiedniego merytorycznego zaplecza. Porażka w tym obszarze może skutkować odrzuceniem zmiany.

Kontrola zmian pozwala zauważyć błędy, zanim pojawi się ich skutek

Kontrola zmian projektowana jest tak, aby błędy w założeniach lub procesie konfiguracji zostały zauważone, zanim wykonana zostanie sama zmiana. Zazwyczaj tylko pierwsze parę razy ktoś znajdzie błąd w żądaniu zmiany, zanim nauczysz się większej ostrożności.

Największą zaletą kontroli zmian jest to, że powala Ci dobrze wypaść. Jeżeli dokonujesz pewnych zmian „w locie" i generują one godziny lub dni niedostępności, jak sądzisz, co Twoi wyżsi przełożeni o Tobie pomyślą? Myślisz, że będziesz pierwszy na liście związanej z awansem lub podwyżkami? Mało prawdopodobne.

Teraz pomyśl, jak odbiorą Cię zarządzający, jeśli masz reputację osoby, która zawsze postępuje zgodnie z procedurami kontroli zmian, przedstawia dobrze udokumentowane żądania zmian i zawsze przeprowadza zmiany zakończone sukcesem.

Kontrola zmian stanowi ochronę interesów firmy. Niezależnie od Twojego zdania na ten temat, interes organizacji jest ważniejszy od Twojej niechęci do procesu kontroli zmian. Wniosek z tego wszystkiego taki, że powinieneś polubić ten proces. Będziesz szczęśliwszy, jeśli podejdziesz do tego jak do pewnego wyzwania i zrozumiesz, że każda duża organizacja wykorzystuje jakąś odmianę kontroli zmian.

Jak nie być komputerowym głupkiem

Rynek informatyczny znany jest ze swojej atrakcyjności dla pewnych określonych indywidualności: osób, które uważają się za mądrzejsze i zwracają się do innych jak do idiotów, ponieważ nie wiedzą oni, do czego służy utrwalacz termiczny w drukarce. Powiedzmy sobie szczerze, większość osób w tym biznesie nie zachowuje się w ten sposób — i nie wszyscy z tych, którzy się tak jednak zachowują, to profesjonaliści!

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Co sprawia jednak, że komputerowcy są tacy, jacy są, i dlaczego są oni atrakcyjni dla rynku informatycznego? Mimo że nie mam zaplecza w postaci studiów, przez lata poczyniłem pewne obserwacje dotyczące zarówno mnie, jak i innych. Przyznaję, że w przeszłości byłem jedną z tych denerwujących osób. Podzielę się z Tobą tym, jak przeobraziłem się w profesjonalistę odnoszącego sukcesy.

Istnieje kilka rzeczy, które przyczyniają się do powstania fenomenu **komputerowego głupka** (ang. *computer jerk*). Niektóre z nich są wrodzone, inne nabyte. Komputerowe głupki pielęgnowane są czasem (zazwyczaj nieświadomie) przez rówieśników lub liderów, ale też właśnie te osoby są w stanie ich zmienić. Rozpoznanie pewnych cech i ich wpływu przyczyni się, mam nadzieję, do pomocy komuś, kogo znasz — może nawet Tobie — by stał się osobą bardziej zrównoważoną i użyteczną, a nie taką, którą określić można jako komputerowego głupka.

Moim zdaniem podstawowe determinanty zawierają się w trzech głównych kategoriach.

Zachowanie

Wszystko, co robimy, opiera się na pewnych nawykach, włączając w to pewne schematy interakcji. Jeżeli przywykłeś radzić sobie z ludźmi w pewien określony sposób, będziesz miał tendencję do utrzymywania tego schematu. Nie jest powiedziane, że te wzorce zachowań nie mogą zostać przełamane. Obcujemy z wieloma typami osób w różny sposób. Na przykład inaczej traktujemy kobiety, a inaczej mężczyzn, różnie odnosimy się do ludzi w zależności od ich pozycji (inaczej do przełożonych, a inaczej do współpracowników). Wielu ludzi twierdzi, że wszystkich traktuje w ten sam sposób, prawda jest jednak inna. Samo to, że masz przyjaciół, wskazuje, iż traktujesz ich nieco inaczej aniżeli pozostałych ludzi.

Jeden ze sposobów, w jaki komputerowe głupki funkcjonują, opiera się na udowadnianiu innym, że są od nich mądrzejsi. Mimo że są przeważnie bardzo inteligentni, z jakichś powodów muszą to udowadniać wszystkim dookoła. Niestety zabierają się za to w niewłaściwy sposób.

Są dwa sposoby na to, aby być uważanym za bystrzejszego od innych:

Być bystrzejszym od ludzi wokół
: Wiedza jest czymś innym niż inteligencja. Posiadanie wiedzy na temat różnych rzeczy nie oznacza umiejętności szybkiego rozwiązywania problemów. Nauczenie się na pamięć książki o anatomii nie uczyni z Ciebie chirurga. Ludzie naturalnie inteligentniejsi od innych — „nieobnoszący się" z tym — są powszechnie uważani za inteligentnych. Najbystrzejsze osoby najczęściej są skromne, ponieważ nie muszą nikomu niczego udowadniać.

Czynić ludzi dookoła głupszymi
: Oto sposób, którym posługuje się często komputerowy głupek. Wierzy, że jeśli sprawi, iż inni będą wyglądali na niemądrych, on wyda się inteligentny. Niestety zapomina o tym, że robienie z innych idiotów powoduje, iż traci on ich sympatię. Gdyby komputerowe głupki przestały tak robić, wiele w ich życiu uległoby zmianie.

Każdy, kogo spotykasz, jest w czymś dobry. Przez te wszystkie lata nauczyłem się tego na własnej skórze. Pamiętaj, że zawsze znajdzie się ktoś bystrzejszy i lepszy od Ciebie.

Najbystrzejsza osoba w departamencie cierpi zawsze na coś, co określam mianem **syndromu alfa-geek**. Osobniki takie potrzebują, aby ludzie wokół wiedzieli, że są one najinteligentniejsze i najlepsze. Często osoby te są najlepsze w swoim małym kręgu i nie uwzględniają w tym szerokiego świata specjalistów IT. Kiedy przybywa konsultant lub zatrudniana jest nowa osoba ze

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

sporymi umiejętnościami, alfa-geek będzie starał się ją zdyskredytować, tak aby odzyskać swoją pozycję. Jeśli jednak stanie naprzeciw pewnego siebie, inteligentnego przeciwnika, alfa-geek poniesie porażkę, próbując zdyskredytować intruza, i skończy jako błazen.

Innym częstym błędem informatyków jest wygłaszanie osądów. Nie jesteś najmądrzejszą osobą na Ziemi i każdy powinien być traktowany z szacunkiem — jeśli to zapamiętasz, będziesz daleki od popełniania takich błędów. Ludzie nie lubią słuchać o swoich potknięciach (szczególnie w obecności innych) i nie potrzebują, aby im mówić, że zrobili źle.

Powiedz innym, że **zrobiłeś** coś źle, i wskaż dlaczego, będą Cię za to szanować.

Powiedz innym, że **oni zrobili** coś źle, i wskaż dlaczego, a uznają Cię za osobę nieprzyjemną.

Dzielenie się wiedzą jest czymś właściwym, utrzymywanie jej w tajemnicy nie. Ludzie często sądzą, że nie powinni się dzielić tym, co wiedzą, ponieważ może to doprowadzić ich do utraty pracy. Dopóki nie pracujesz dla jakichś despotów, jest to dalekie od prawdy. Zespoły pracują lepiej razem, a pewien wewnętrzny szacunek formuje się, jeśli członkowie potrafią dzielić się informacjami. Rozwiązywanie problemów jest skuteczniejsze, jeśli więcej niż jedna osoba wie, jak dana rzecz działa. Pomyśl, kogo w branży informatycznej podziwiasz. Prawdopodobnie Twój podziw wynika z tego, że nauczyłeś się czegoś od tej osoby.

Ostatnia rzecz. Jeżeli zauważyłeś coś katastrofalnego w projekcie, nie czekaj na spotkanie, na którym wszystkim to uzmysłowisz. Taktyka ta jest cechą charakterystyczną dla komputerowych głupków. Przedstawienie złych wiadomości na spotkaniu wzmacnia siłę przekazu i zapewnia zwrócenie uwagi zebranych. Co więcej, komputerowy głupek wierzy, że w ten sposób pokazuje swoją mądrość. W szczególności menedżerowie nie lubią niespodzianek, zwłaszcza podczas zebrań. Właściwym sposobem na przekazanie złych wiadomości jest ciche poinformowanie o tym swojego menedżera, który będzie w stanie odpowiednio to zakomunikować. W takim przypadku spotkanie nie będzie stanowiło reakcji na ogłoszony **problem**, a raczej dyskusję na temat jego rozwiązania.

Środowisko

Wierzę w to, że komputerowe głupki są tacy, jacy są, ponieważ jest to akceptowane, a w niektórych przypadkach są oni zachęcani do takiego zachowania. Jednym z głównych czynników mających na to wpływ jest brak poczucia własnej wartości. Ludzie chcą być lubiani, nawet więcej, podziwiani. Problemem komputerowych głupków jest to, że uważają, iż sposobem na zdobycie podziwu innych jest pokazywanie, jak bardzo są od nich mądrzejsi.

Pracowałem kiedyś w firmie, która wzmocniła osobowość typu alfa-geek tak, że pewien pracownik tej firmy prezentował typowe tendencje takiego modelu. Problem polegał na tym, że człowiek **był** dobry, ale nie tak dobry, jak uważali wszyscy. Doprowadziło to do niekończących się błędów w projektach i problemów, jakich doświadczały zespoły zarządzające wsparciem. Wszystko to napędzane było właśnie cechami osobowości tego człowieka.

Kiedy usiadłem z nim i wytłumaczyłem, że jego pomysły są cenne, ale trzeba je nieco skorygować, aby były lepsze, wydał się tym zainteresowany. Gdy wyjaśniłem, że ogranicza go jego własna osobowość, również był zainteresowany. Po tygodniach wspólnej pracy zmienił się na lepsze. Moje przewodnictwo oraz mentoring okazały się silniejsze niż wpływ środowiska.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Przywództwo i mentoring

Z mojego doświadczenia wynika, że ludzi zachowujących się jak komputerowe głupki można zmienić. Zazwyczaj jest w stanie zrobić to osoba posiadająca ich szacunek oraz ciesząca się autorytetem. W moim przypadku było to tak, że pracowałem w fabryce chemicznej, myśląc, iż jestem królem świata. Miałem minikomputer, zarządzałem wszystkimi komputerami PC, a nawet siecią Santa Clara Systems w wersji 1.1 z okablowaniem miedzianym 10Base-2 (dla tych, którzy nie znają Santa Clara Systems, połączyli się oni w 1986 roku z firmą Novell).

Pewnego dnia, kiedy czułem się nadzwyczaj zadowolony z siebie, obraziłem swojego szefa swoim protekcjonalnym zachowaniem, gdy prosił mnie o pomoc. Był to człowiek, którego raczej lubiłem i bardzo podziwiałem, wiele się od niego nauczyłem. Mimo to w wieku 22 lat uważałem, że jestem od niego mądrzejszy.

Szef zaprosił mnie do swojego biura i zapytał, czy wiem, co to jest toluol, metakrylan metylu oraz proces dyspersji. Kiedy odpowiedziałem, że nie znam żadnego z tych pojęć, wyjaśnił mi nieco ich znaczenie i zapytał, dlaczego ktoś taki jak ja, trzy razy młodszy od niego, uważa, że ma prawo zwracać się do niego w ten sposób. Wiedziałem, że ma rację. Byłem kretynem.

Morał tej historii jest taki, że potrzebowałem czasu, aby ktoś mi powiedział, iż jestem kretynem, i zrobił to w taki sposób, by to do mnie dotarło. Fakt, iż szanowałem tę osobę, bardzo pomógł, ale wszystko zmieniło to, że zostałem odciągnięty na bok i tam wytłumaczono mi, że zachowuję się niestosownie.

Uważam, że większość ludzi zwyczajowo już piętnujących innych w celu udowodnienia swojej wyższości po prostu potrzebuje lekcji od lidera lub mentora, który wskaże im, jak należy się zmienić.

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Skorowidz

10 Gigabit Ethernet, 34
1000Base-T, 34
100Base-T, 34
10Base-2, 19
10Base-5, 20
10Base-T, 20, 34

A

ABR, area border router, 142
ACD, Automatic Call Distributor, 336, 511
ACE, Access Control Entries, 403
ACE, Application Control Engines, 28
ACTS, Automated Computer Time Service, 665
adaptery portów, port adapters, 248
add/drop, 336
adres anycast, anycast address, 646
adres lokalny IPv6, 653
adres loopback, 649
adres MAC, 25, 75, 324
adres mechanizmu przełączania awaryjnego, 479
adres nieokreślony, 649
adres VIP, 183
adresacja, 647
adresowanie klasowe, 627
adres-sieciowy/długość-prefiksu, 116
adresy globalne, global unicast, 649
adresy lokalne dla łącza, link-local unicast, 649
AES-CCMP, 499
agregacja portów, Link Aggregation, 73
agregacja prędkości, 74
aktualizacja
  destrukcyjna, disruptive, 309
  niedestrukcyjna, nondisruptive, 309
aktualizacja niewyzwalana, nontriggered updates, 135
aktualizacja wyzwalana, triggered update, 135
aktualizacje, 691–693
  nowe błędy, bugs, 692
  problemy ze sprzętem, 692
  zagrożenia, 692
aktywne połączenia, 494
alarm czerwony, 356
alarm niebieski, 358
alarm żółty, 357
algorytm
  back-off, 109
  Cisco, 74
  Dijkstry, 130
  haszujący, 74
  host-dependent, zależny od hosta, 190
  round-robin, karuzela, 190
  rozłożenia obciążenia na pakiet, 118
  weighted load balancing, ważone rozłożenie obciążenia, 190
  WPA, 499
    enterprise, 499
    personal, 499
alokacja adresów IP, 625
AMI, Alternate Mark Inversion, 349
analiza łączy, 594
analizator przełączanego portu, 327
analizator składniowy, parser, 199
ANI, Automatic Number Identifier, 511
anycast, anycast address, 646
aplikacja SSH, 169
ARAP, AppleTalk Remote Access Protocol, 440
architektura
  backplane, 265
  chassis-based, 28
  dwuwarstwowa, 616
  jednowarstwowa, 618
  trójwarstwowa, 615
    rdzeń, core level, 615
    warstwa dostępowa, access level, 615
    warstwa dystrybucyjna, distribution level, 615
argument
  back, 252
  global, 296
  gre, 179
  match, 164, 167

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

argument
permit, 163
route-map, 163
show, 276
trace, 317
arkusz IP, 610
arkusz układu portów, 607
arkusz VLAN, 610
arkusz z podsieciami, 641
ARP, Address Resolution Protocol, 25
ARPANET, 335
ASA, Adaptive Security Appliance, 455, 459, 478
ASBR, autonomous system border router, 142
ASDM, Adaptive Security Device Manager, 455
ASIC, Application-Specific Integrated Circuits, 284
ASN, autonomous system number, 127
ATA, Analog Telephone Adapter, 510
atak typu denial-of-service, 39
atak typu double tagging, 39
atak typu spoofing, 39
ATM, asynchronous transfer mode, 15
automatyczna konfiguracja interfejsów, 658
autonegocjacja, 33–34, 37
autonomous system external (ASE) LSA, 143
AutoQoS, 332
AVF, Active Virtual Forwarder, 189
AVG, Active Virtual Gateway, 189
awaria łącza internetowego, 185
awaria łącza szeregowego, 185
awaria pośredniego linku, indirect link failure, 108
awarie, 669, 673
kontrola warstwy fizycznej, 674
rozwiązywanie problemów, 672

B

B8ZS, Binary Eight Zero Substitution, 349
BackboneFast, 107
backplane, 265
bank kanałów, channel bank, 338
baza danych OSPF, 114
baza danych sieci VLAN, 43
baza danych synchronizacji, config-sync database, 302
baza informacji routingu, 114
baza RIB, 116
B-channels, 340
BDR, backup designated router, 143
bearer channels, kanały B, 340, 349
BECN, Backward-Explicit Congestion Notfication, 385
BERT, Bit Error Rate Test, 338
bezpieczeństwo, 445, 620
bezpieczeństwo portów, 325
ochrona, protect, 325
ograniczenie, restrict, 325
zamknięcie, shutdown, 325
bezpieczeństwo sieci bezprzewodowej, 497
AES-CCMP, 499
EAP, 498
filtrowanie adresów MAC, 498
TKIP, 498
WEP, 498
WPA, 499
WPA2, 499
BGP, Border Gateway Protocol, 126, 147
bit lokalnie administrowany, 650
biuro centralne, central office, 338
blade enclosure, 613
blokowanie portów, blocking, 97, 99
blokowanie ruchu, 447
błąd, 176
błąd Duplicate NAT entry, 487
błąd ludzki, human error, 669, 692
błąd ramkowania, framing error, 352
błędy CRC, 591
błędy końca ramki, 370
błędy ramek, framing errors, 591
błędy wejścia, input errors, 591
bond, 74
BPDU Guard, 105
BPDU, bridge protocol data units, 97
brama domyślna, default gateway, 114
brama, gateway, 113
BRI, Basic Rate Interface, 341
Bridge Assurance (BA), 111
broadcast network, 64, 145
bufor fluktuacji opóźnienia, jitter buffer, 515
burza broadcastowa, 23, 91
burza multicastowa, 320
burza rozgłoszeniowa, 320
burza unicastowa, 320
BVP, bipolar violation, 354

C

CAM, content-addressable memory table, 25
CAN, campus area network, 15
Catalyst, 24
Catalyst 3750, 313
Catalyst 6500 VSS, 84
CatOS, 24, 43, 57, 249
CBWFQ, class-based weighted fair queuing, 597
CCIE, 658
CEF, Cisco Express Forwarding, 208, 214, 218
centrum wsparcia technicznego, 675
channel, 73
channel ID, 79

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

CHAP, 432
chmura MPLS, 397
chwilowa przepływność, burst rate, 380
CIDR, Classless Internet Domain Routing, 137, 630
  agregacja tras, 630
cienki Ethernet, thin-net, 19
CIR, Commited Information Rate, 380, 587
Cisco Adaptive Security Appliances, 405
Cisco Nexus, 263
Cisco Technical Assistance Center, 219
Cisco Unified Communications Manager, 513
Cisco Unified Communications Manager Express, 513
Cisco Unity, 513
Cisco Unity Connection, 513
Cisco Unity Express, 513
CLI, command-line interpreter, 43
CME
  przekazywanie typu blind, 524
  przekazywanie typu consult, 524
CMM, Communication Media Module, 248
cofanie zmian w konfiguracji, 304
configuration rollback, 304
CoS, Class of Service, 330
CPE, customer premises equipment, 339
CRC, Cyclic Redundancy Check, 351
CRC6, 355
cross-stack EtherChannel, 84
CSM, Content Switching Modules, 231, 245
CSU, Channel Service Unit, 338
CSU/DSU, 358, 363, 365
cykliczne sprawdzanie redundancji, 350
czas starzenia, 326
czas UTC, 663
częstotliwość przepływu danych, data rate, 337

D

DACCS, Digital Access Cross-Connect System, 339
data channel, kanał D, 340, 349
data port/DTE, 359
dCEF, distributed Cisco Express Forwarding, 235
D-channel, 340, 349
DE, Discard Eligible, 380
debugowanie procedury uwierzytelnienia, 432
debugowanie utoQoS, 332
definicja sekundy, 662
demarc, 338, 339
detekcja równoległa, parallel detection, 34
diagram sieci, 613
DLCI, Data Link Connection Identifier, 378, 392
DMZ, 489
dodawanie vPC, 86
dokumentacja, 689
domena rozgłoszeniowa, broadcast domain, 21
domena VTP, 62, 67
domeny kolizyjne, collision domains, 21
domeny, domains, 127
dopasowanie ostatniego oktetu, 624
dopełnienie tabulatorem, 275
dostęp do portu przełącznika, 324
dostęp zdalny, 490
DR, designated routers, 143–144
drzewo binarne, binary tree, 211
drzewo trie, 214
drzewo wielodrożne, multiway tree, 213
DS, digital signal, 340
DS0, 373
DS1, 348
DS3, 565
  AMI, 371
  B3ZS, 371
  bezkanałowy sygnał, 367
  bity C, 369
  HDB3, 371
  kanałowe łącze T3, 367
  konfiguracja, 371
  M13, 368
DS3, Digital Signal 3, 367
DSCP, Differentiated Services Code Point, 560
DSU, Data Service Unit, 338
DTP, Dynamic Trunking Protocol, 54
dual-active, 258– 259
dupleks, 33
dynamic secure MAC addresses, 324
dynamiczny routing, 93
dyrektywa privilege, 428
dyrektywa secret, 428
dystans administracyjny, 115
dystans administracyjny, administrative distance, 115, 128, 131
dystrybucja obciążenia, 191
dzienniki pracy systemu, 446
dzień słoneczny, 662

E

EAP, Extensible Authentication Protocol, 498
EFS, Extended Super Frame, 353
EGP, external gateway protocol, 133
EIGRP, Enhanced Internal Gateway Routing Protocol, 93, 126, 138
enhanced PAgP, 258
enkapsulacja ietf, 386
enkapsulowanie trunku, 56
Ephone, 530
Ephone-DN, 527
Errored Second, 355

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

ESF, Extended Super Frame, 349
EtherChannel, 73
  cross-stack, 84
  konfiguracja, 77
  Multichassis, 84
  protokoły, 78
  rozkład obciążenia, 74
  zarządzanie, 77
EtherChannels, 618
Ethernet, 273, 335
EtherSVI, 224
express forwarding, 155

F

farmy serwerów, 617
Fast EtherChannel (FEC), 73
Fast Ethernet (100 Mb/s), 73
FCoE, Fibre Channel over Ethernet, 263
FCS, frame check sequence, 53
FDDI, fiber distributed data interface, 15
FEBE, Far-End Block Errors, 369
FECN, Forward-Excplicit Congestion Notification, 385
FEOOF, Far-End Out-Of-Frame, 369
FEX, 287
  przypinanie poczwórne, pinning four, 290
  przypinanie podwójne, pinning two, 290
  przypinanie pojedyncze, pinning one, 290
  vPC, 290
FIFO, 557
filtrowanie adresów MAC, 504
filtrowanie tras redystrybuowanych, 168
filtry dynamiczne, 418
firewall PIX, 179, 244
flex links, 319
FlexWAN, 28, 248
fluktuacja opóźnienia, 515
fluktuacja opóźnienia międzypakietowego, 515
Frame Relay, 377
  BECN, 385
  CIR, 381
  DE, 381
  FECN, 385
  konfiguracja, 385
  nadsubskrypcja, 383
  niedopasowanie prędkości, 583
  podinterfejsy, 390
  połączenia punkt-punkt, 382
  projektowanie sieci, 381
  przeciążenia, 385
  sieci siatkowe, 382
  zamawianie usługi, 380
full mesh, 381
funkcja
  add/drop, 336
  autonegocjacji, 36
  BackboneFast, 102, 107
  BPDU Guard, 105, 288
  config-sync, 89
  dCEF, 244
  ETRN, 463
  fixup, 463
  hot-swap, 233
  plinkFast, 102
  PoE, 244
  PortFast, 104
  QoS, 332
  smartport macros, 316
  UplinkFast, 105
  wykrywania równoległego, 34
FWSM, Firewall Services Module, 28, 231, 244
FXS, Foreign eXchange Service, 248

G

G.711, 510
G.722, 510
G.729, 510
GGP, Gateway-to-Gateway Protocol, 113
Gigabit EtherChannel (GEC), 73
Gigabit Ethernet, 33, 580
GLBP, Gateway Load Balancing Protocol, 181, 189
  AVF, 189
  AVG, 189
  konfiguracja, 190
  śledzenie obiektu, object tracking, 193
główny adres MAC, root MAC address, 98
GPS, Global Positioning System, 665
GRE, Generic Routing Encapsulation, 169
gruby Ethernet, thick-net, 19
grupa administracyjna, admin group, 79
grupa kanałów, channel group, 82
grupa podstawowa, 484
grupa zapasowa, 484
grupowanie, teaming, 74
gwarantowana szybkość transmisji, 380

H

H.225, 509
H.245, 509
H.323, 509
hairpinning, 512
hasła, 425, 429
hasła szyfrowane, 426
hasło VTP, 69
hierarchia podnośnej E, E-carrier, 340

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

hierarchia podnośnej J, J-carrier, 341
hierarchia sygnału cyfrowego, 340
histogram ASCII, 93
HSRP, Hot Standby Router Protocol, 181
  konfiguracja, 183
  router aktywny, active, 182
  router rezerwowy, standby, 182
HSRP group, 181
HSRP interface tracking, 184
hub, 19, 23
hub and spoke, 381

I

IANA, Internet Assigned Numbers Authority, 126, 625
ID główne, root ID, 98
ID kanału, 79
ID mostka, bridge ID, 98
ID routera OSPF, 144
identyfikator DLCI, 378
identyfikator OSPF, 144
identyfikator SSID, 502
ID-proces, process-ID, 127
IDS, Intrusion Detection Systems, 327
IDSM, Intrusion Detection System Module, 28, 246
IEEE 802.11a, 496
IEEE 802.11b, 496
IEEE 802.11g, 496
IEEE 802.11n, 496
IEEE 802.1Q, 52
IEEE 803.ad, 78
IETF, 651
IETF, Internet Engineering Task Force, 645
IGP, internal gateway protocol, 132
IGRP, nternal Gateway Routing Protocol, 126
ilość danych wchodzących, input rate, 92
ilość danych wychodzących, output rate, 92
infrastrukturalna bezprzewodowa sieć LAN, 495
inspekcja SMTP, 463
inspekcje, 463
instalacja VSS, 252
instalowanie obrazu, 308
interfejs
  CLI, 43
  Ethernet, 623
  fizyczny, 74
  HSRP, 184
  pętli zwrotnej, 170
  WAN, 669
  wirtualny, 74
  zapory sieciowej ASA, 456
internal router, 142
Internet Protocol, 113
interwał, 321
IOS, Internetwork Operating System, 24, 45, 55, 249, 272
IP precedence, 198
ip_input, 209
IPv6, 645
  adresacja, 647
  konfiguracja routera, 652
  maski podsieci, 648
  NAT, 651
  podsieci, 650
  typy adresów, 648
ISDN, Integrated Services Digital Network, 340
ISL, Inter-Switch Link, 52–53
ITU, International Telecommunications Union, 509
IXC, interexchange carrier, 341
IXP, Internet eXchange Point, 16

J

Janitor Principle, 675

K

kable kategorii 5e, 20
kanał danych, data channel, 508
kanał transportowy, bearer channel, 508
kanał trunkowy, trunk, 51
kanał vPC, 86
kanały głosowe, 349
karta AIM-CUE, 524
karta dystrybucji danych, distributed forwarding card, 234
karta MSFC, 222, 226
karta PFC, 238
keepalive link, 85
klasa HTTP, 596
klasa Voice-RTP, 598
klaster VSS, VSS cluster, 84, 291
klasyfikator, classifier, 467
klauzula deny, zabraniająca, 199
klauzula permit, zezwalająca, 199
klauzule, clauses, 198
klient NTP, 664
klient VTP, 62
kodek, 510
kodowanie, encoding, 349
  AMI, 350
  B8ZS, 351
kodowanie liniowe, 371
kolejka
  bez priorytetu, 603
  błędna konfiguracja, 599
  domyślna, 604

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

kolejka
LLQ, 568
niski priorytet, 601
priorytetowa, 602
kolizja, collision, 21
komunikat błędu, 176
komunikat LMI, 393
Active, 395
Deleted, 396
Inactive, 396
Static, 396
koncentrator VPN, 169
konfiguracja
autonegocjacji, 37
bezkanałowego łącza DS3, 372
CSU/DSU, 362
DHCP, 501
dla map klas, 596
dla trunku, 296
DS3 podzielonego na kanały, 373
dual-home, 291
EtherChannel, 77, 80
FEX, 290
firewalla wielokontekstowego, 467
Frame Relay, 385, 388
GLBP, 190
hasła, 433
HSRP, 183, 186
interfejsów SVI, 223, 227
interfejsu radiowego, 501
kolejki, 599
kontekstów, 471
kontroli burz, 321
linii, 528
list ACL portu, 415
list ACL routera, 416
lokalnych kont użytkowników, 427
łączy DS3, 371
map VLAN-ów, 416
mapy polityk, 596
mechanizmu przełączania awaryjnego, 477, 479
mechanizmu SIP, 541
modułu IDS, 247
NTP, 664
OSPF, 127
PAT, 487
przełączania awaryjnego
aktywny-aktywny, 482
aktywny-oczekujący, 478
punktu dostępowego, 500
refleksyjnych list dostępu, 421
routera w IPv6, 652
routerów, 575, 579
routingu z poziomu interfejsów, 276
serwera zabezpieczeń, 437
sieci konwergentnej, 595
SPAN, 327
T1, 362
telefonów, 526
trunków, 55
urządzania ASA, 455
urządzeń FEX, 287, 296
VLAN-ów, 43
VLAN-ów za pomocą bazy danych, 45
vPC, 85
VTP, 66
zapory sieciowej, 448
kontekst administracyjny, 466
kontekst standardowy, 466
kontekst systemowy, 466, 474
konteksty, 456
konteksty kaskadowe, 469
kontrola burz, storm control, 320
kontrola zmian, 694
kontrole scalenia, 300
konwencja nazewnicza dla urządzeń, 614
konwergencja, 134
koszt linku, 130
koszt ścieżki, path cost, 98
koszty konsultacji, 682
koszty stałe, 682
koszty wyposażenia, 682
koszty zasobów ludzkich, 682
kształtowanie ruchu, 584
Bc, 584
Be, 584
CIR, commited information rate, 584
Tc, 584
KVM, Keyboard Video Monitor, 622

L

LACP, Link Aggregation Control Protocol, 78
LAG, Link Aggregation, 73
LAN, local area network, 15
LAP, Lightweight Access Point, 500
LATA, local access and transport area, 341
latencja, latency, 342
LEC, local exchange carrier, 343
liczba odrzuconych pakietów, 601
liczba przeskoków, hop count, 129
liczba wpisów dla sieci, network entries, 631
liczenie do nieskończoności, counting to infinity, 129
line, 360
linia odbiorcza RX, 35
linia POTS, 551
linia transmisyjna TX, 35

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

linie dostępowe, 425
link awaryjny, link-failover, 188
link jednokierunkowy, unidirectional, 110
link podstawowy, 188
linki flex, 319
lista metod uwierzytelniania, 440, 443
listy ACL, 403, 406, 414, 577
listy ACL portu, port ACL, 414
listy ACL routera, router ACL, 415
listy dostępu, access lists, 197, 406, 407, 419, 461
listy kontroli dostępu, access control list, 403
listy refleksyjne, 406
listy Turbo ACL, 411
LLQ, low-latency queuing, 223, 571
LMI, Local Management Interface, 384
local, 359
logiczne plany numeracji, 525
logiczny interfejs SVI, 224
lokalny interfejs zarządzania, 384
LOS, loss of signal, 354, 364
LSA, link state advertisements, 143

Ł

łańcuch katastrof, 671
łańcuch wypadków, 671
łącza ISDN, 399
łącza OC3, 398
łącze DS3, *Patrz* DS3
łącze logiczne, 73
łącze OC3, 579
łącze równorzędne, peer link, 86
łącze T1, *Patrz* T1
łącze trunkowe SIP, 540
łącze wysycone, 590
łączenie VLAN-ów, 39
łączenie w stos, stacking, 27, 313

M

MAC address table, 25
MAE, Metropolitan Area Ethernet, 188
makra, 316
makro globalne, 316
makro interfejsu, 316
makro Workstation, 319
maksymy GAD-a, 677
MAN, metropolitan area network, 15
mapa polityk, 577
  class-default, 596
  HTTP, 596
  Voice Control, 596
  Voice-RTP, 596
mapy klas, 576
mapy routingu, route maps, 163–164, 197
mapy VLAN-ów, VLAN map, 415
MARS, Monitoring Analysis and Response System, 246
maski o zmiennej długości, 119, 627
maski odwrotne, 404
maski podsieci, 648
maski wieloznaczne, wildcard masks, 140, 404
match, 199
MEC, 84
mechanizm
  AAA, 430
  CHAP, 430
  CoS, 562
  grup obiektów, 460
  MWI, 543
media-konwertery, media converters, 114
merge checks, 300
metoda rozkładania obciążenia, 77, 219
metoda ważenia, weighting, 191
metody przydzielania adresów IP, 634
  metoda dzielenia na pół, 634
  metoda odwrotna binarna, reverse binary, 635
  metoda sekwencyjna, 634
metryka, metric, 128
midplane, 265
miejsca docelowych pakietów, per-destination basis, 218
model ISO/OSI, 114
moduł
  ACE, 243
  CMM, 248
  CSM, 231, 245
  Ethernet, 244
  FlexWAN, 248
  FWSM, 231, 243–244
  FXS, 248
  IDSM, 246
  NAM, 231, 243, 246
  SFM, 233, 235
  SRST, 249
moduł zarządzający, 239
moduł zarządzający Supervisor-720, 234
moduł zarządzający, supervisor, 238
moduły, 240
moduły fabric-only, 240
moduły nonfabric-enabled, 240
moduły systemu, 272
monitorowanie przełączania awaryjnego, 480
mostek główny, 112
mostek główny STP, 101
mostek główny, root bridge, 97
mostek wyznaczony, designated bridge, 99

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

mostek, bridge, 91
MPLS, Multiprotocol Label Switching, 397, 399, 584
obsługa QoS, 401
sieci wirtualne, 400
MRTG, Multi Router Traffic Grapher, 348, 591
MSFC, Multilayer Switch Feature Card, 58, 222, 238, 670
MTA, Mail Transfer Agent, 464
MTU, Maximum Transmission Unit, 413
multicast, 649
Multichassis EtherChannel, 84
multipleksacja M13, 368
multipleksowanie, multiplexing, 344
mutual-exclusion (mutex), 300
MWI, Message Waiting Indicator, 533

N

nadsubskrypcja, oversubscription, 383
nadziewanie bitami, bit stuffing, 369
NAM, Network Analysis Module, 28, 231, 246, 571
narzędzie MRTG, 594
NAS, Network Attached Storage, 77
NASI, NetWare Asynchronous Services Interface, 440
NAT overload, 486
NAT w telefonii SIP, 544
NAT, Network Address Translation, 486, 626
NBMA, nonbroadcast multiaccess, 146
negocjacja trunków, 54
network LSA, 143
Nexus, 59, 189, 227, 263
architektura, 277
ikony, 277
Nexus 1000, 270
Nexus 2000, 269
Nexus 5000, 266
Nexus 7000, 264
niedopasowanie prędkości, 583
niespójność tablicy adresów MAC, 96
niezgodność dupleksu, 108
NIST, National Institiute of Standards and Technology, 665
NRO, Number Resource Organization, 645
NSSA, 143
NSSA LSA, 143
NTP, Network Time Protocol, 661, 663
NX-OS, 27, 48, 59, 271
aktualizacja, 307
pliki obrazów, 307
NX-OS Ethernet, 273

O

obszar MAE, 188
obszar nieszkieletowy, nonbackbone
not so stubby area (NSSA), 144
NSSA totally stub area, 144
obszar normalny, normal area, 143
obszar pnia, stub area, 143
totally stubby area (TSA), 144
obszar szkieletowy, backbone
obszar zerowy, 143
obwód PVC, 392, 395
obwód VC, 377
PVC, 377
SVC, 377
obwód wirtualny, 377
odporność na problemy, fault-tolerance, 181
odrzucane pakiety, abort errors, 591
OFDM, Orthogonal Frequency-Division Multiplexing, 496
ogłoszenia stanu łącza, 143
oktety podsieci, 639
OOF, out-of-frame, 354
opis makra, 318
opisowe zdarzenia telefoniczne, named telephone events, 545
opóźnienie, 514
opóźnienie propagacji, 342
opóźnienie przetwarzania, 342
organizacja IANA, 126, 626
OSPF, Open Shortest Path First, 114, 125, 142
ostatnia mila, last mile, 344

P

PAgP, Port Aggregation Control Protocol, 78, 253, 258
pakiet, packet, 24
pakiet ARP, 25
pakiet ICMP
Echo reply, 413
ICMP unreachable, 413
parameter-problem, 414
source-quench, 414
Time exceeded, 413
pakiet IP, 560
pakiet kontrolny, 155
pakiety głosowe RTP, 574
pakiety multicast, multicast packets, 126
pakiety RTP, 560
pamięć podręczna trasy, route cache, 211
parametr global, 87
parametr spanning tree, 89
partial mesh, 381
pasmo, band, 337

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

PAT, Port Address Translation, 418, 486
payload, 359
PBX, Private Branch Exchange, 344, 511
peer keepalive link, 85, 292
peer link, 293
pełny dupleks, full-duplex, 33, 36
per-destination, 219
per-destination basis, 218
per-packet, 219
per-packet basis, 219
pętla, 97, 110
pętla lokalna, local loop, 344
pętla redystrybucji, 160
pętla warstwy drugiej, 92
pętla zwrotna, 170
pętla zwrotna bloku danych, 358
pętla zwrotna linii, 358
PFC, policy feature card, 238
PIX, 455, 459
PLAR, 512
platforma 6500, 231
plik List.xml, 551
pliki PST, 573
pływająca trasa statyczna, floating static route, 125
podinterfejs, subinterface, 382, 391
podnośna, 340
podnośna T, T-carrier, 346
podsieci, 650
podsieci VLSM, 629
podsieci w adresacji klasowej, 628
podsieciowanie, IP subnetting, 637
podsieć, Subnet, 117, 119, 135
PoE, Power over Ethernet, 511
point-to-multipoint, 145
point-to-point, 145
poison reverse, 134
pole Bytes Delayed, 582
pole TOS protokołu IP, 561
polecenia NAT, 486
polecenia vPC, 86
polecenie
  aaa new-model, 437
  authentication shared, 501
  auto qos voip, 332
  auto-cost reference-bandwidth, 142
  bridge irb, 500
  buffer-delete, 299
  buffer-move, 299
  clear trunk, 59
  clear vtp pruneeligible, 72
  clear xlate, 493
  commit, 300
  copy run start, 491
  copy running startup, 274
  create cnf-files, 523
  debug auto qos, 332
  debug frame-relay lmi, 394
  debug ip policy, 204
  debug voip dialpeer all, 551
  default-metric, 153, 156
    MTU, Maximum Transmission Unit, 156
    niezawodność, reliability, 156
    obciążenie, loading, 156
    opóźnienie, delay, 156
    przepustowość, bandwidth, 156
  description, 82
  eigrp log-neighbor-changes, 141
  encapsulation frame-relay ietf, 386
  ETRN deque, 463
  failover, 479
  failover group, 484
  failover lan unit primary, 476
  feature, 272
  feature glbp, 191
  feature hsrp, 188
  feature vpc, 85
  frame-relay interface-dlci, 391
  glbp weighting, 194
  global, 486
  how clock details, 666
  interface, 275
  interface range, 48
  interface-vlan, 224
  ip load-sharing per-destination, 219
  ip load-sharing per-packet, 219
  ip ospf priority, 144
  ip reflexive-list timeout liczba_sekund, 423
  ip route-cash cef, 218
  ipv6 unicast-routing, 657
  join-failover-group grupa, 484
  konfiguracji przycisku, 530
  limit-resource, 282
  mac-address use-virtual, 255
  mac-address-table synchronize, 258
  macro description, 318
  name nazwa-vlan, 48
  nat, 486
  network, 136
  no auto qos voip, 334
  no auto-summary, 140
  no ip route-cache, 216
  no login, 426
  no mls qos, 334
  no redistribute igrp system-autonomiczny, 156
  no shutdown, 224
  no switchport, 222

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

polecenie
no vtp pruning, 70
ntp peer ip-address, 667
object-group, 461
passive-interface, 136, 137
ping, 123, 279
priority, 578
redistribute, 152
redistribute ospf, 157
redistribute static, 154
redundancy force-switchover, 260
redundancy reload peer, 260
reload in, 670
rollback running-config checkpoint, 305
router-id, 144
routing-context, 124
routing-context vrf default, 124, 280
security-level, 457
service password-encryption, 426
service-module t1 linecode, 362
set port channel, 73
set spantree uplinkfast enable, 106
set trunk, 57
set vlan, 44
set vtp domain, 67
set vtp mode, 68
set vtp passwd, 69
set vtp pruning enable, 71
shape peak bps, 585
show, 45
show access-list, 411
show arp, 275
show auto qos, 334
show channel, 73
show channel info, 79
show checkpoint summary, 305
show controllers, 376
show etherchannel, 81
show etherchannel summary, 80
show fabric status, 242
show fabric switching-mode, 242
show fabric utilization, 242
show failover, 480
show feature, 272
show frame-relay map, 396
show frame-relay pvc, 386, 395
show glbp, 191, 194
show glbp brief, 193
show int brief, 286
show interface, 80, 82, 273, 393, 458, 592
show interface brief, 83, 274
show interface capabilities, 52
show interface interfejs_nr, 250
show interface status, 250
show interface switchport backup, 320
show interface trunk, 56, 57, 60
show ip arp, 275
show ip cef, 219
show ip eigrp neighbors, 141
show ip eigrp topology, 173
show ip interface brief, 83, 223, 279, 428
show ip ospf database, 162
show ip protocols, 154
show ip route, 117, 118
show mac-address-table, 96
show module switch all, 261
show ntp peers, 666
show ntp peer-status, 666
show ntp status, 666
show parser macro brief, 318
show parser macro name nazwa-makro, 318
show policy-map interface nazwa_interfejsu, 581, 586
show port, 249
show port capabilities, 52
show port channel, 73, 79
show port trunk, 58
show port-security interface, 326
show processes cpu history, 93
show processes cpu sorted, 217
show redundancy states, 253
show route-map, 204
show running-config switch-profile, 302
show service-policy global, 464
show sip-ua register status, 542
show spanning-tree, 101
show spanning-tree root, 103
show spanning-tree summary, 102
show spantree summary, 103
show switch virtual dual-active summary, 259
show switch virtual link, 257
show switch virtual link detail, 257
show switch virtual role, 256
show switch-profile nazwa-profilu buffer, 299
show trunk, 59
show vdc, 282
show version, 475
show vlan, 46, 47
show vpc consistency-parameters, 87, 296
show vrf, 123
show vtp password, 69
show vtp status, 68
show xlate, 493
spanning-tree port type edge trunk, 297, 301
spanning-tree portfast, 104
standby preempt, 183

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

standby priority, 183
standby track Serial0/0 10, 185
static, 486
switch switch_nr preempt, 254
switchport, 49, 55, 222
switchport access, 47
switchport backup interface, 319
switchport broadcast, 321
switchport mode fex-fabric, 289
switchport mode trunk, 59
switchport multicast, 321
switchport port-security aging type, 326
switchport priority extend, 331
switchport trunk allowed, 56
switchport trunk allowed vlan, 59
switchport trunk allowed vlan numer-vlan, 88
switchport trunk pruning vlan, 71
switchport unicast, 321
switch-profile, 302
switchto vdc vdc_nazwa, 283
sync-peers, 300
sync-peers destination, 301
tacacs-server, 438, 439
traceroute, 410
trunk encapsulation, 56
username, 429
vdc combined-hostname, 283
vlan, 45
vlan database, 45
vlan filter, 417
vlan nr_vlan, 48
vpc role priority, 292
vrf context nazwa_vrf, 122
vrf member nazwa_vrf, 123
vtp domain, 67
vtp password, 69
witchport port-security, 324
write mem, 474
write memory, 474
write standby, 483
write-erase, 304
polityka routingu, policy routing, 198, 200
polityki usług, 578
połączenia Frame Relay, 398
połączenia MPLS, 401
połączenia PPP, 430
połączenia SIP, 538
połączenia wychodzące, 537
połączenie, 15
połączenie SSH, 490
połączenie telnet, 490
poprawka Bridge Assurance, 111
port aktywny, listening, 100
Port Channel, 73
port channel interface, 73
port docelowy SPAN, 324
port FXO, 512
port FXS, 512
port główny, root port, 99
port konsoli, console port, 268
port trunkowy, 41
port w stanie error-disabled, 325
port w stanie inicjalizacji, 100
port wyłączony, disabled, 100
port wyznaczony, designated port, 99
port zablokowany, blocking, 100
PortFast, 104
włączanie, 104
wyłączanie, 104
porty fabric, fabric ports, 270
porty głosowe, 526
porty hosta, host ports, 111, 270
porty normalne, 111
porty przełącznika, 518
porty sieciowe, 111
porty Spanning Tree, 111
host, 111
network, 111
normal, 111
POTS, plain-old telephone service, 344, 534
powiadamianie o przeciążeniu w kierunku transmisji, 385
poziom bezpieczeństwa, 457, 458
poziom uprawnień, privilege level, 428
poziom ważności komunikatów, 492
półdupleks, half duplex, 33, 35
PPTP, Point-to-Point Tunneling Protocol, 179
prefiks, 147
prędkość transmisji, 33
prędkość transmisji, bursting speed, 586
PRI, Primary Rate Interface, 341, 348
priorytet główny, root priority, 98
priorytet kolejki, 601
problem mikrozbitek, 566
problem w działaniu systemów VoIP i SIP, 549
problem ze sprzętem, 692
problemy Frame Relay, 393
problemy punktu dostępowego, 504
profil przełącznika, switch profile, 298, 301
projektowanie NTP, 662
projektowanie QoS, 571
projektowanie sieci, 605, 678
arkusz IP, 610
arkusz układu portów, 607
arkusz VLAN, 610
diagram sieci, 613

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

dokumentacja, 605
schematy szaf dystrybucyjnych, 611
specyfikacja wymagań, 606
wymagania dotyczące zasilania i chłodzenia, 611
projektowanie sieci IP, 625
projekty sieci, 615
protokoły EtherChannel, 78
protokoły sterujące połączeniami, 603
protokół
802.1Q, 53
ARP, 379
bramy wewnętrznej, 132
bramy zewnętrznej, 133
CDP, 521
CHAP, 432
EIGRP, 120
Frame Relay, 377
FTP, 420, 557
GLBP, 181, 189
GRE, 169
HSRP, 181
HTTP, 557
hybrydowy, 130
ICMP, 412
Inverse ARP, 379
ISL, 53
OSPF, 142
PAP, 431
PPTP, 179
RIP, 129, 134
RIPv2, 137
routingu bezklasowego, 134
routingu klasowego, 133
routingu, routing protocol, 114, 125
SIP, 507, 537, 544
Spanning Tree, 291, 319
SSH, 557
stanu łącza, 130
STP, 91
TCP, 557
TCP/IP, 24
telnet, 557
UDP, 508, 557
VRRP, 181
VTP, 54, 61, 65
wektora odległości, 129
WPA, 499
próg opadania, falling threshold, 322
próg wznoszenia, rising threshold, 322
prywatne zakresy IP, 625
przeciążenie, 568
przeciążony NAT, 486
przekazywanie ramek, forwarding, 100
przekazywanie rozmów, 524
przekierowanie portów, 488
przełączanie, switching, 24, 207
przełączanie awaryjne, failover, 465, 475
aktywność, 476
aktywny-aktywny, 165, 482
aktywny-oczekujący, 482
konfiguracja, 477
monitorowanie, 480
oczekiwanie, 476
pakiet hello, 477
primary, 476
secondary, 476
w urządzeniach ASA, 478
zachowanie stanu, 476
przełączanie kontekstu
ekspresowe przekazywanie Cisco, 208
optymalne przełączanie, optimum switching, 208
szybkie przełączanie, fast switching, 208
przełączanie kontekstu z obsługą przerwań, interrupt context switching, 207, 210
przełączanie krzyżowe, crossbar, 234
przełączanie optymalne, optimum switching, 213
przełączanie pakietów w routerze, 208
przełączanie procesowe, process switching, 207, 209, 216
przełącznik, 24
cena, 27
chłodzenie, 30
ciężar, 27
elastyczność, 28
instalowanie, 30
moduły, 30
możliwość rozbudowy, 28
prędkość, 28
redundancja, 28
rozmiar, 27
zasilanie, 27, 29
przełącznik 1000, 270
przełącznik 2000, 270
przełącznik 3550, 93, 313
przełącznik 3750, 84, 223, 313
AutoQoS, 332
kontrola burz, 320
linki flex, 319
łączenie w stos, 313
makra, 316
SPAN, 327
VLAN głosowy, 330
zabezpieczanie portów, 324
zakresy interfejsów, 316
przełącznik 3Com, 54
przełącznik 5000, 266, 267

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

przełącznik 5010, 266
przełącznik 5020, 266
przełącznik 5548, 266
przełącznik 6500, 28, 84, 228, 231
architektura, 233
CSM, 231
FWSM, 231
NAM, 231
obudowy enhanced chassis, 237
SFM, 233
szyna, 234
przełącznik 6500 typu chassis, 84
przełącznik 6509, 231
przełącznik 6509-V-E, 237
przełącznik 7000, 265
przełącznik 7018, 264
przełącznik chassis-based, 28, 29, 222, 238
przełącznik Cisco, 24, 27, 100
przełącznik Cisco Catalyst, 249
przełącznik ethernetowy, 24
przełącznik kasetowy, blade, 53
przełącznik modularny, 27
przełącznik NEBS chassis, 237
przełącznik Nexus, 122
przełącznik Nexus 7000, 265, 280
przełącznik o stałej konfiguracji, 27
przełącznik rdzenia sieci, 609
przełącznik typu chassis, 221
przełącznik typu chassis-based, 228
przełącznik VSS, 260
przełącznik w funkcji routera, 223
przełącznik warstwy trzeciej, Layer-3 switches, 24, 43, 221
przełącznik wielowarstwowy, multilayer switches, 24, 221, 227, 231
przełącznik z zerową domeną VTP, 64
przepustowość, throughput, 234, 337
przerwa międzycyfrowa, interdigit timer, 535
przerwanie odbioru, receive interrupt, 209
przestrzeń adresowa w sieciach IP, 631
przetaktowywanie, overclocking, 369
przetwarzanie list dostępu, 407
przydział adresów IP, 183
przydzielanie podsieci IP, 633
przyrostek metryczny, 322
przywoływanie, 529
przywracanie do trybu online, 194
PSK, PreShared Key, 499
PTT, Push To Talk, 532
publiczne adresy IP, 488
publiczne zakresy IP, 625
punkty dostępowe, 495
PVC, Permanent Virtual Circuit, 134, 377
PVC, private virtual circuits, 583
PVDM, Packet Voice Digital Signal Procesor, 512
PVST, Per-VLAN Spanning Tree, 100
PVST+, Per-VLAN Spanning Tree Plus, 100

Q

QoS, Quality of Service, 198, 332, 401, 555, 558, 597
CBWFQ, Class-Based Weighted Fair Queuing, 563
harmonogramowanie, 560
kolejkowanie priorytetowe, 564
kształtowanie ruchu, 564
LLQ, Low Latency Queuing, 564
mechanizmy, 558
monitorowanie, 597
nakładanie polityki, 560
oznaczanie pakietów, 560
priorytety, 563, 572
szerokość pasma, 569
WFQ, 563
własne kolejkowanie, 564
wymagania, 574

R

radio FM, 337
RADIUS, Remote Authentication Dial-In User Service, 438
ramka BPDU, 97
ramka, frame, 24
ramkowanie, framing, 352
DS1, 353
głosowe, 352
M13, 369
typu C-bit, 369
RBOC, Regional Bell Operating Company, 344
redundancja, 112
redystrybucja, redistribution, 127, 151
EIGRP, 156
OSPF, 158
RIP, 153
redystrybucja tras IGRP, 139
redystrybucja wzajemna, mutual redistribution, 159, 166
refleksyjne listy dostępu, 418, 420
regenerator, repeater, 19, 21
reguła dla GRE, 179
reguły filtrowania pakietów, 412, 446
reguły połączeń przychodzących, 537
reguły połączeń wychodzących, 537
reguły VoIP, 536
reguły wybierania, dial peer, 534, 551
rejestracja zdarzeń, 491
rejestrowanie aktywności, accounting, 437

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

relacja przylegania, neighbor adjacency, 141
remote line, 360
remote payload, 360
rezerwowy adres IP, standby IP address, 183
RFC 1129, 661
RFC 1232, 354, 356
RFC 1334, 430
RFC 1701, 169
RFC 1819, 646
RFC 1918, 625
RFC 2281, 184
RFC 2460, 645, 646
RFC 2474, 561
RFC 2784, 169
RFC 3330, 626
RFC 4291, 648
RFC 5095, 645
RFC 5722, 645
RFC 5871, 645
RFC 791, 113
RIB, routing information base, 114, 207
RIP, Routing Information Protocol, 114, 125, 134
RIPv2, 136, 137
RJ-45, 19
rodzaje modułów, 243
rodzaje tras, 132
  podsieć, 118
  sieć nadrzędna, 118
  supersieć, 118
  trasa domyślna, 118
  trasa hosta, 118
  trasa skonsolidowana, 118
role, 430
  network-admin, 430
  network-operator, 430
  vdc-admin, 430
  vdc-operator, 430
root link query, 108
rotacja haseł, 446
route poisoning, 134
router, 125
router CME, 519
router DR, 146
router EIGRP, 126
router granicy obszaru, 142
router granicy systemu autonomicznego, 142
router LSA, 143
router na patyku, router-on-a-stick, 42
router OSPF, 143
router wewnętrzny, 142
router wyznaczony, 143
routery Cisco, 24
routery szkieletowe, backbone routers, 142
routing, 113, 207
routing dynamiczny, 125
routing między VLAN-ami, 221
routing pomiędzy warstwami, 621
routing rekurencyjny, 175
rozgłoszenia podzbioru, subset advertisements, 63
rozgłoszenia zbiorcze, summary advertisements, 63
równoważenie obciążenia, load balancers, 181
RPS, Redundant Power System, 228
RPVST, 228
RSPAN, Remote Switched Port Analyzer, 327
RTP, Real-Time Protocol, 155, 508

S

SAN, Storage Area Network, 263
sąsiad, neighbor, 141
SCCP, Skinny Call Control Protocol, 510
scheduler, 602
schemat
  6500 VSS, 292
  Frame Relay, 378
  GLBP, 189
  HSRP, 182
  MPLS, 399
  Spanning Tree, 291
  vPC, 292
  VSS, 253
schematy szaf dystrybucyjnych, 611
segment sieci, 19
serwer
  DHCP, 501
  DNS, 449
  lokalizacji, 538
  NTP, 524, 665, 667
  pocztowy, 449
  pośredniczący, 538
  przekierowania, 538
  RADIUS, 505
  rejestracji, 538
  syslog, 492
  TFTP, 307
  typu rack, 624
  VTP, 62
  WWW, 449
serwery kasetowe, blade servers, 296
serwery TACACS+, 438
sesja monitor, 327
set vtp pruneeligible, 72
SFM, Switch Fabric Module, 233
sieci ad hoc, 495
sieci bezprzewodowe, 495
sieci korporacyjne, 615
sieci packet radio, 495

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

sieć
ATM, 15
CAN, 16, 17
chroniona, 447
DS3, 15
Ethernet, 15
firmowa, 452
Frame Relay, 15, 377, 382
główna, major, 630
IPv6, 653, 656
klasowa, classful, 630
konwergentna, 595
konfiguracja, 595
LAN, 16, 17
lokalna, 114
MAN, 16, 17
mostkowana, bridged network, 15
MPLS, 223
nadrzędna, Major network, 117, 121
nierozgłoszeniowa wielodostępowa, 146
OSPF, 145
pakietowa, 15
przeciążona, congested network, 589
punkt-punkt, point-to-point network, 15, 145
punkt-wielopunkt, 145
rozgłoszeniowa, 145
Ethernet, 145
FDDI, 145
Token Ring, 145
rutowalna, routed network, 15
rutowana, 621
siatkowa, 377, 382
światłowodowa FDDI, 15
Token Ring, 15
VLAN, 517
VoIP, 559
w standardzie T1, 15
WAN, 16, 17
wewnętrzna, 448
zbieżna, 134
zewnętrzna, 449
Simple NTP, 663
SIP INVITE, 546
SIP, Session Initiation Protocol, 509, 537, 552
skok, hop, 118
skrętka, 19
sloty, 235
sloty GBIC, 313
słowo kluczowe
detail, 494
dual-line, 527
host, 438
password, 427
radius, 439
range, 409
tacacs+, 439
tiemout, 528
smartport macro, 319
SMDS, Switched Multimegabit Data Service, 146
sneakernet, 15
SONET, synchronous optical network, 346
SPAN destination, 324
SPAN, Switched Port Analyzer, 327
Spanning Tree, 91, 112, 619
BackboneFast, 107
BPDU Guard, 105
PortFast, 104
sposób działania, 97
stany portu, 99
UplinkFast, 105
spedytor, forwarder, 189
split horizon, 134
split-brain, 258
SRST, Survivable Remote Site Telephony, 249, 512
SSH, 169
SSID, Service Set ID, 500
stacking GBIC, 313
stakowanie, stack, 27
stan łącza, link-state, 130
stan poza ramką, 354
standardowe polecenia IOS, 46
standardy Ethernetu, 34
starzenie się adresów MAC, 326
static secure MAC addresses, 324
statyczna translacja NAT, 488
sterowanie połączeniem, 508
sticky secure MAC addresses, 324
STP, Spanning Tree Protocol, 91, 319
strategia kolejkowania, strategy queueing, 592, 597
stratum, 662
strefa DMZ, 447, 448, 450
strona e-commerce, 619, 622, 672
struktura, fabric, 234
struktury przełączania, 235
strumień danych głosowych, 508
strumień DS3, 367, *Patrz* DS3
strumień RTP, 509
summary LSA for ABR, 143
summary LSA for ASBR, 143
superframe, 352
supersieć, Supernet, 117, 121
Supervisor-1A, 239
Supervisor-2, 239
Supervisor-32, 239
Supervisor-720, 238, 239
Supervisor-720-10G VSS, 239

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

SVC, Switched Virtual Circuit, 377
SVI, switched virtual interfaces, 221
switch, *Patrz* przełącznik
switch virtual domain, 254
sygnalizacja poza pasmem, out-of-band signaling, 349, 352
sygnalizacja w paśmie, in-band signaling, 352
sygnał analogowy, 336
sygnał cyfrowy, 336
synchronizacja konfiguracji, config-sync, 296
system VSS, 254
systemy autonomiczne, autonomous systems, 127, 132
szafa rackowa, 29
szerokopasmowa transmisja audio, 510
szerokość pasma (ang. bandwidth, 337
szybkie przełączanie, fast switching, 211, 217
szybkość przesyłania bitów, 566
szyna
  C bus, 234
  crossbar fabric bus, 234
  D bus, 234
  R bus, 234
szyna crossbar switching bus, 241
szyna danych, data bus, 234
szyna sterująca, control bus, 234

Ś

ścieżka do mostka głównego, 98
ścieżka przełączania, 216
ścieżki, paths, 147, 207
śledzenie interfejsu HSRP, 184
śledzenie obiektu w GLBP, 193
śledzenie obiektu, object tracking, 191

T

T1, 384, 508, 565, 574
  AMI, 350
  B8ZS, 351
  CRC6, 355
  CSU/DSU, 364
  czysty kanał T1, Clear-channel T1, 349
  D4/Superframe, 352
  dupleks, 348
  EFS, 353
  kanał T1, Channelized T1, 349
  konfiguracja, 362
  PRI, 349
  ramkowanie, 351
T1 PRI, 526
tabela przekazywania, forwarding table, 214
tabela przylegania, adjacency table, 214
tabela routingu, 114, 170
tabela routingu IP, 117
tablica adresów MAC, 25, 96
tablica CAM, 25, 96
tablica routingu, 204, 631
tablica topologii, 114
tablica topologii EIGRP, 163
TAC, 675
TACACS, Terminal Access Controller Access-Control System, 438
TACACS+, 438
tagi, 162
tagowanie tras, tag, 52, 162
T-Berd, T1 Bit Error Rate Detector, 346
TDM, time-division multiplexing, 346
technika MIMO, 497
telefon 7921G, 533
telefon 7970, 527
telefon IP, 518
telefon programowy, 512
telefonia komórkowa, 507
telekomunikacja, 335
Telnet, 62
test pętli zwrotnej, loopback test, 358
testowanie failover, 672
TFTP, 550
TKIP, Temporal Key Integrity Protocol, 498, 502
tokeny generujące hasła jednorazowe, 446
TOS, Type Of Service, 560
translacja adresów i portów, 486
translacja adresów sieciowych, 486
translacja połączeń głosowych, voice translation, 546
trasa domyślna, Default route, 117, 121
trasa EIGRP uzyskana poprzez tunel, 177
trasa hosta, Host route, 116, 119
trasa RIP, 131
trasa skonsolidowana, Summary, 117, 120
trasa wewnętrzna, internal, 127
trasa zewnętrzna, external, 128, 168
trasowanie, 113
trasy podłączone bezpośrednio, connected routes, 152
trasy pozyskiwane dynamicznie, 152
trasy statyczne, static routes, 152
trasy zewnętrzne OSPF, 158
troubleshooting, 672
trunki, 40, 55, 618
trunking, 41
tryb access, 55
tryb aktywny, 78
tryb config-if-range, 315
tryb config-sync, 301
tryb crossover, 287

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

tryb dedykowanej przepustowości, dedicated rate-mode, 284
tryb dupleks, duplex mode, 33, 79
tryb dynamic, 55
tryb dynamic auto, 55
tryb dynamic desirable., 55
tryb failover, 181, 233, 253
tryb hybrydowy, hybrid mode, 222, 225
tryb klienta VTP, 63
tryb konfiguracji routera, 136
tryb konfiguracji zakresu interfejsów, interface-range mode, 275
tryb konfiguracji, configuration mode, 46
tryb konfiguracyjny bazy danych VLAN-ów, 45
tryb kontekstowy, 465
tryb natywny, native mode, 222
tryb passthru, 287
tryb pasywny, 78
tryb portu dedykowanego, dedicated port mode, 284
tryb routingu, routed mode, 59, 478
tryb serwera VTP, 66
tryb simplex, 256
tryb testowania, 477
tryb transparentny, 63, 67, 478
tryb transparentny urządzeń ASA, 457
tryb trunk, 59
tryb uprzywilejowany, 428
tryb VSS, 257
tryb VTP, 63, 68
  client, 68
  off, 68
  server, 68
  transparent, 68
tryb wielokontekstowy, 456, 465
tryby switchport, 55
TTL, time-to-live, 182
tunel, tunnel, 169
  miejsce docelowe, destination, 172
  trasa EIGRP, 177
  źródło, source, 172
tunel GRE, 170, 174
tunel GRE w tunelu VPN, 178
tworzenie
  EtherChannel, 78
  EtherChannel „w poprzek”, 84
  list metod uwierzytelniania, 440
  listy, 421
  listy dostępu, 504
  podinterfejsu, 391
  podsieci, 627
  VRF, 122
typ interfejsu, interface type, 79
typy adresów w IPv6, 648
typy LSA
  LSA Not So Stubby Area, 143
  LSA routera, 143
  LSA sieci, 143
  łączne LSA dla routerów ABR, 143
  łączne LSA dla routerów ASBR, 143
  zewnętrzne LSA systemu autonomicznego, 143
typy routerów OSPF
  granicy obszaru, 142
  granicy systemu autonomicznego, 142
  szkieletowy, 142
  wewnętrzny, 142
  wyznaczony, 143
  wyznaczony zapasowy, 143

U

uczenie się adresów MAC, learning, 100
udziały, load-shares, 190
umowa SLA, 688
unified fabric, 263
UplinkFast, 105
uproszczone punkty dostępowe, 500
urządzenia ASA,405, 463
urządzenia FEX, 229
urządzenie podstawowe, primary, 181
urządzenie zapasowe, secondary, 181
usługa DHCP dla telefonów, 521
usługa RADIUS, 505
usługa telefoniczna, 522
usługa TFTP, 521
usługa zakończenia ethernetowego, 579
usuwanie list ACL, 411
utrata pakietów3, 515
utrata sygnału, loss of signal, 354
utylizacja łącza T1, 591
uwierzytelnianie, authentication, 437
uwierzytelnianie AAA, 437
uwierzytelnianie CHAP, 435
uwierzytelnianie linii, 441
  enable, 441
  group radius, 441
  group tacacs+, 441
  krb5, 441
  krb5-telnet, 441
  line, 441
  local, 441
  local-case, 441
  none, 441
  własna, 441
uwierzytelnianie PPP, 430, 443
  group radius, 443
  group tacacs+, 443
  if-needed, 443

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

uwierzytelnianie PPP
krb5, 443
local, 443
local-case, 443
none, 443
własna, 443
uwierzytelnianie standardowe, 425
uwierzytelnianie typu Login, 442
uwierzytelnianie użytkowników
ARAP, 440
Login, 440
NASI, 440
PPP, 440
uwierzytelnienie dwustronne, 431, 434
uwierzytelnienie jednostronne, 431, 433

V

VAD, Voice Activity Detection, 545
VC, Virtual Circuit, 377
VDC, Virtual Device Context, 122, 280–281, 429
Virtual Port Channel, 84, 291
Virtual Routing and Forwarding, 122, 279
VLAN, 22, 328, 619
VLAN 0, 330
VLAN 1, 226
VLAN 10, 22, 51, 226, 327, 330
VLAN 100, 57, 64, 330
VLAN 128, 228
VLAN 20, 22, 51, 226
VLAN 30, 51
VLAN 40, 51
VLAN database, 43
VLAN głosowy, voice VLAN, 330
VLAN, Virtual LAN, 39
VLAN-hopping, 39
VLAN-y w sieci VoIP, 517
VLSM, Variable Length Subnet Masks, 119, 137, 627
VoIP, Voice over IP, 507, 515, 557
vPC, virtual Port Channel, 84, 294
VPN, Virtual Private Networks, 169, 446, 686
point-to-point, punkt-punkt, 169
remote Access, zdalnego dostępu, 169
VPN concentrator, 169
VRF, Virtual Routing and Forwarding, 122
domyślny, default, 122, 279, 307
zarządzający, management, 122, 279
VRF Earth, 123
VRF Mars, 123
VRRP, Virtual Router Redundancy Protocol, 181
VSL, virtual switch link, 255
VSS detekcja dual-active, 258
VSS, Virtual Switching System, 84, 259
VTP clients, 62
VTP Pruning, 64, 65, 70, 71
VTP servers, 62
VTP transparent, 63
VTP, Virtual Trunking Protocol, 619
VTP, VLAN Trunking Protocol, 43, 54, 61
VTY, Virtual Teletype Terminal, 426

W

WAN Interface Card, 361, 669
WAN, wide area network, 15
WAP, Wireless Access Point, 495
warstwa aplikacji, 619
warstwa druga łącza danych, 52
warstwa fizyczna, 111
wartości oktetów podsieci, 639
warunek match, 199
wektor odległości, distance-vector, 129
WEP, Wired Equivalent Privacy, 498
WFQ, Weighted Fair Queuing, 559
WiFi protected Access, 499
wirtualna domena przełącznika, 254
wirtualne interfejsy przełączania, 221
wirtualne sieci lokalne, 39
wirtualne sieci prywatne, 169
wirtualny link przełącznika, 255
wirtualny spedytor, forwarder, 189
WLAN, Wireless LAN, 495
WPA2-Enterprise, 499
WPA-PSK, 502
wpis local6.debug, 493
wpis PAT Global, 493
wpisy kontroli dostępu, 403
współczynniki oktetów podsieci, 640
współdzielenie interfejsu wewnętrznego, 469
współdzielenie interfejsu zewnętrznego, 468
wsteczne powiadamianie o przeciążeniu, 385
wydatki inwestycyjne, 691
wydatki operacyjne, 691
wydzielone obszary kontroli, 127
wykres MRTG, 594
wykres T1, 591
wzajemne wykluczenie, 300

Z

zabezpieczenie portów, port security, 324, 622
zagnieżdżanie, 421
zakresy interfejsów, interfaces ranges, 315
zapasowy router wyznaczony, 143
zapisywanie informacji o zdarzeniach, 493
zapora sieciowa, firewall, 406, 409, 445–446
zapora sieciowa ASA, 471, 490
zapora sieciowa PIX, 463, 491

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

zasady kształtowania ruchu, 565
zatruwanie tras, 134
zewnętrzny interfejs zapory sieciowej, 451
złamanie kodowe, 354
złącze
  Amphenol, 244
  GBIC, 244
  RJ-45, 244
  SFP, 244
  small-form-factor GBIC, 244
  XENPAK, 244
zmiana nazwy użytkownika, 435
znacznik administratora, administrator tag, 164
znak %, 124
znak (@), 316
znakowanie tras, 162
zwrot z inwestycji, 691

Ż

żądanie rozgłoszenia, advertisement, 64

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

O autorze

Gary A. Donahue jest aktywnym zawodowo konsultantem, który pracuje w branży komputerowej od 27 lat. Gary zajmował już takie stanowiska jak: programista, administrator systemów mainframe, inżynier wsparcia technicznego, administrator sieciowy, architekt sieci oraz konsultant. Pracował jako dyrektor ds. infrastruktury sieciowej w ogólnokrajowej firmie konsultacyjnej i jest obecnie prezesem własnej firmy konsultingowej GAD Technologies z New Jersey.

Kolofon

Pies przedstawiony na okładce książki *Wojownik sieci* to dog niemiecki. Znany również pod nazwą Wielki Duńczyk, dog niemiecki jest imponującym, lecz eleganckim i uczuciowym psem, który waży z reguły 45 – 60 kg i mierzy 70 – 80 cm. Może być różnej maści: pręgowany, szaro-brązowy czy biało-czarny (arlekin). Przeciętna długość życia tego psa wynosi 7 – 10 lat.

Istnieją pewne kontrowersje na temat pochodzenia doga niemieckiego. Jedni twierdzą, że pochodzi on z Danii, inni, że z Niemiec. Tak czy inaczej, to niemieccy hodowcy doprowadzili tę rasę do jej obecnego kształtu.

Jego angielska nazwa, *German boarhound* („niemiecki pies tropiący dziki"), wzięła się od zdolności psów tej rasy do polowań na dziki, wilki czy jelenie. Królowie Danii i Anglii często uznawali psy myśliwskie za święte, a mówi się, że w pewnym okresie dogi niemieckie zamieszkiwały prawie wszystkie zamki w Niemczech.

Wizerunki dogów niemieckich można znaleźć na ścianach egipskich grobowców. W poemacie historycznym *Beowulf* dog niemiecki również pojawia się jako pies myśliwski. W czasach średniowiecza dogi były chowane razem z ich właścicielami, ponieważ uważano je za przewodników duchowych w zaświatach. Jednak ich duchowe personalizacje nie zawsze były mile widziane — dog był czasem uznawany za wysłannika piekieł zwanego czarnym włochatym demonem. Taki właśnie czarny pies-widmo był prawdopodobnie inspiracją dla powieści *Pies Baskervillów* autorstwa Artura Conana Doyle'a.

Obraz z okładki zaczerpnięty został z wielotomowej publikacji Richarda Lyddekera *Library of Natural History*

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com

PROGRAM PARTNERSKI

GRUPY WYDAWNICZEJ HELION

1. ZAREJESTRUJ SIĘ
2. PREZENTUJ KSIĄŻKI
3. ZBIERAJ PROWIZJĘ

Zmień swoją stronę WWW
w działający bankomat!

Dowiedz się więcej i dołącz już dzisiaj!

http://program-partnerski.helion.pl

GRUPA WYDAWNICZA
Helion SA

Ebookpoint.pl KOPIA DLA: Robert Arndt robertarndtt@gmail.com