

Wykład 12

Ruting dynamiczny

Sieci Komputerowe 2017



System Autonomiczny

AS (ang. Autonomous System) to sieć lub zbiór sieci zarządzanych przez jedną organizację (np. UW) i realizujący spójną politykę routingu.

AS jest identyfikowany za pomocą numeru 16 bitowego ASN (ang. Autonomous System Number).

1-64511 – numery publiczne przydzielane przez RIR

64512-65534 – prywatne, nieprzydzielane

0 oraz 65535 zarezerwowane

Numery AS są przydzielane przez RIR (dla Europy RIPE: <http://www.ripe.net>).

Rutowanie dynamiczne

Zachodzi wtedy, gdy routery informują się wzajemnie o dostępnych sieciach.

Komunikacja następuje za pomocą protokołów routowania.

Rutowanie dynamiczne nie zmienia obsługi procesu routowania w warstwie IP przez jądro systemu operacyjnego.

Na routerze jest uruchamiany proces, który wprowadza wpisy do tablicy tras automatycznie.

Dwie klasy protokołów routowania

IGP (Interior Gateway Protocols)

używany wewnątrz AS

OSPF (Open Shortest Path
First)

(i)BGPv4

EGP (Exterior Gateway Protocols)

używany do komunikacji ruterów
z różnych AS

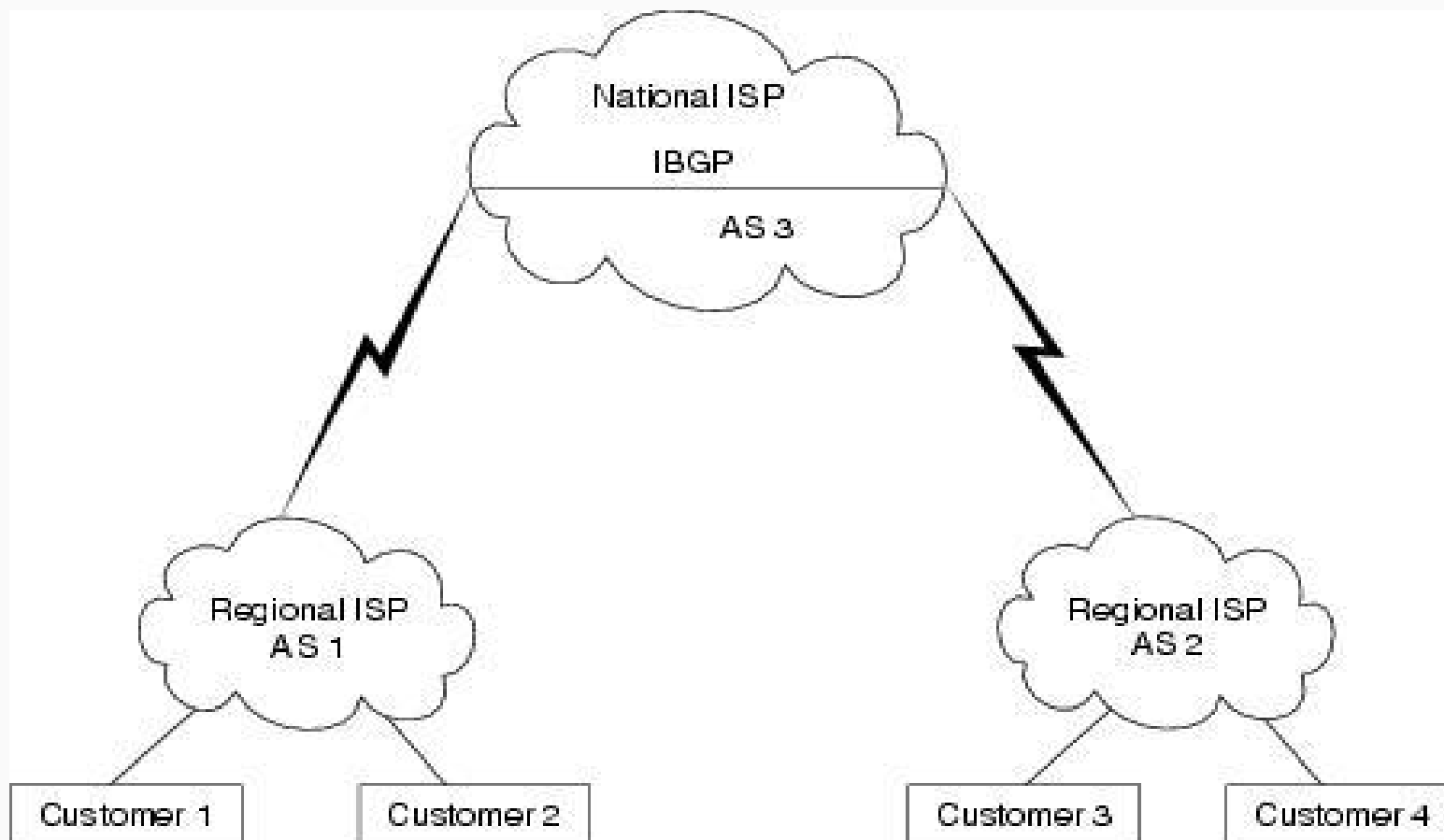
(e)BGPv4 (Border Gateway
Protocol), RFC 4271

Internet to BGP

iBGP i eBGP

Exterior i Interior BGP

sesje połączeń między ruterami w ramach tego samego AS i różnych AS



Sesja BGP

BGP wymaga utrzymywania połączenia (tzw. peeringu), TCP port 179.

Rutery z różnych AS powinny być w tej samej sieci fizycznej.

Komunikat BGP typu open zawiera:

- wersja protokołu
- lokalny ASN
- hold-time (czas po którym w przypadku braku odebranych pakietów, peer powinien uznać połączenia za uszkodzone)
- identyfikator BGP – adres IP interfejsu

Sesja BGP

Po rozpoczęciu sesji peery wymieniają się informacjami takimi jak:

- dostępne prefiksy
- ścieżki (sekwencje numerów AS) wraz z atrybutami
- niedostępne prefiksy

Poza tym okresowo wysyłają między sobą pakiety keep-alive, które także służą do wykrywania zerwanego połączenia.

Wymiana informacji między AS

Aby sieci w dwóch różnych AS mogły się komunikować:

- Pierwszy AS musi rozgłosić swoje prefiksy do drugiego.
- Drugi musi zaakceptować rozgłaszane prefiksy.
- Drugi AS musi rozgłosić swoje prefiksy do pierwszego.
- Pierwszy musi zaakceptować rozgłaszane prefiksy.
 - Rozgłaszany prefiks to np. 193.0.96.0/24

BGP - atrybuty

Atrybuty opisują charakterystykę podsieci (prefiksu) otrzymanego lub rozgłaszanego za pośrednictwem BGP.

Pozwalają na wybór optymalnej trasy, kontrolę rozgłaszanych informacji i mogą być określane przez administratora, tak aby realizować określoną politykę rutowania.

Pozwolę sobie ich nie wypisywać, bo jest ich dużo. Żeby to zobrazować, algorytm wyboru ścieżki jest na kolejnych slajdach...

Oto dlaczego nie opisuję atrybutów, czyli algorytm wyboru ścieżki (ciekawostka)

BGP wybiera tylko jedną trasę, umieszcza ją w tablicy rutowania IP i (ewentualnie) rozgłasza.

Uwzględniane kryteria wyboru trasy w następującej kolejności:

- Nie uwzględniaj ścieżki, dla której next-hop jest nieosiągalny.
- Preferuj trasę z większą wagą.
- Jeśli wagi są takie same, uwzględnij większą wartość local preference.
- Preferuj trasę z krótszym AS_path.
- Jeśli wszystkie ścieżki mają taki sam AS_path, uwzględnij tę z mniejszym atrybutem origin (IGP<EGP<incomplete).
- Jeśli atrybuty origin są takie same, uwzględnij najmniejszą wartość MED.
- Jeśli wartość MED jest taka sama, preferuj ścieżkę zewnętrzną.
- Jeśli ścieżki są takie same wybierz ścieżkę przez ruter bliższy w sensie IGP.
- W końcu uwzględnij najmniejszy leksykograficznie adres IP, odczytany z komunikatu BGP.

Kiedy **nie** potrzebujemy BGP

Jeden provider i jedno łącze

- wystarczy default route...

Jeden provider, dwa łącza z inną adresacją, tylko dla redundancji

- wystarczy skonfigurować policy routing

Kiedy BGP jest potrzebne

2 ISP

- redundancja
- możliwość stosowania własnej polityki routingu
- rozkładanie obciążenia

Tranzyt przez nasz AS do innych AS

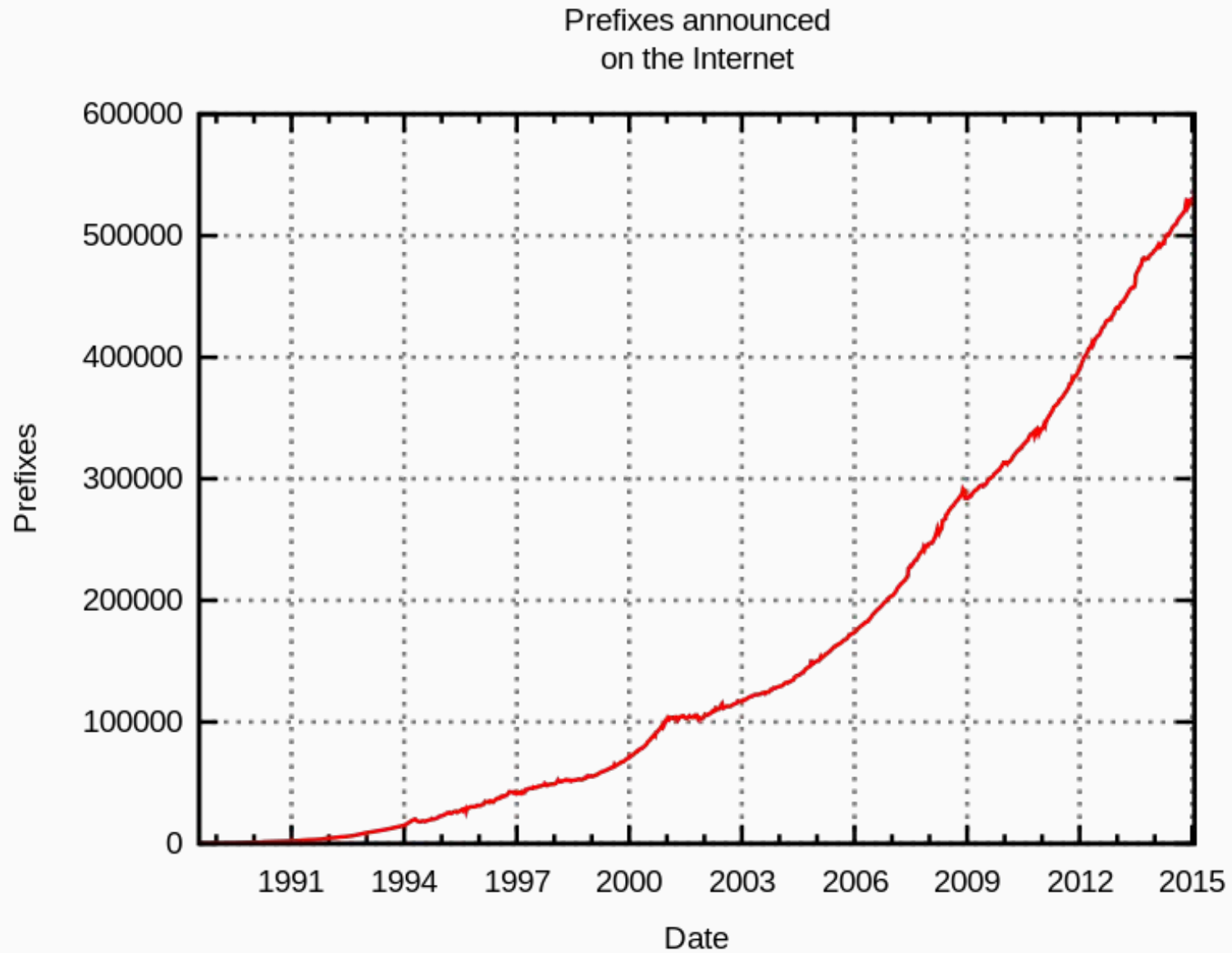
Rozgłaszanie prefiksów nieprzydzielonych lub należących do klas prywatnych:
1/8, 2/8, 10/8

- stosowanie filtrowania (tzw. bogon filters)
- np. spowodowały kiedyś problemy z Neostradą
- a w dzisiejszym nieidealnym świecie IPv6 można się natknąć na cykle co najmniej raz w roku...

W skrócie: w BGP nie ma wbudowanych mechanizmów ufania ścieżkom, każdy może próbować rozgłaszać co chce.

Dobre praktyki nakazują po obydwu stronach peeringu specyfikować przesyłane/akceptowane prefiksy, i zazwyczaj tak się robi. Ale bywają niechlubne wyjątki.

Liczba prefixów w tablicach FIB



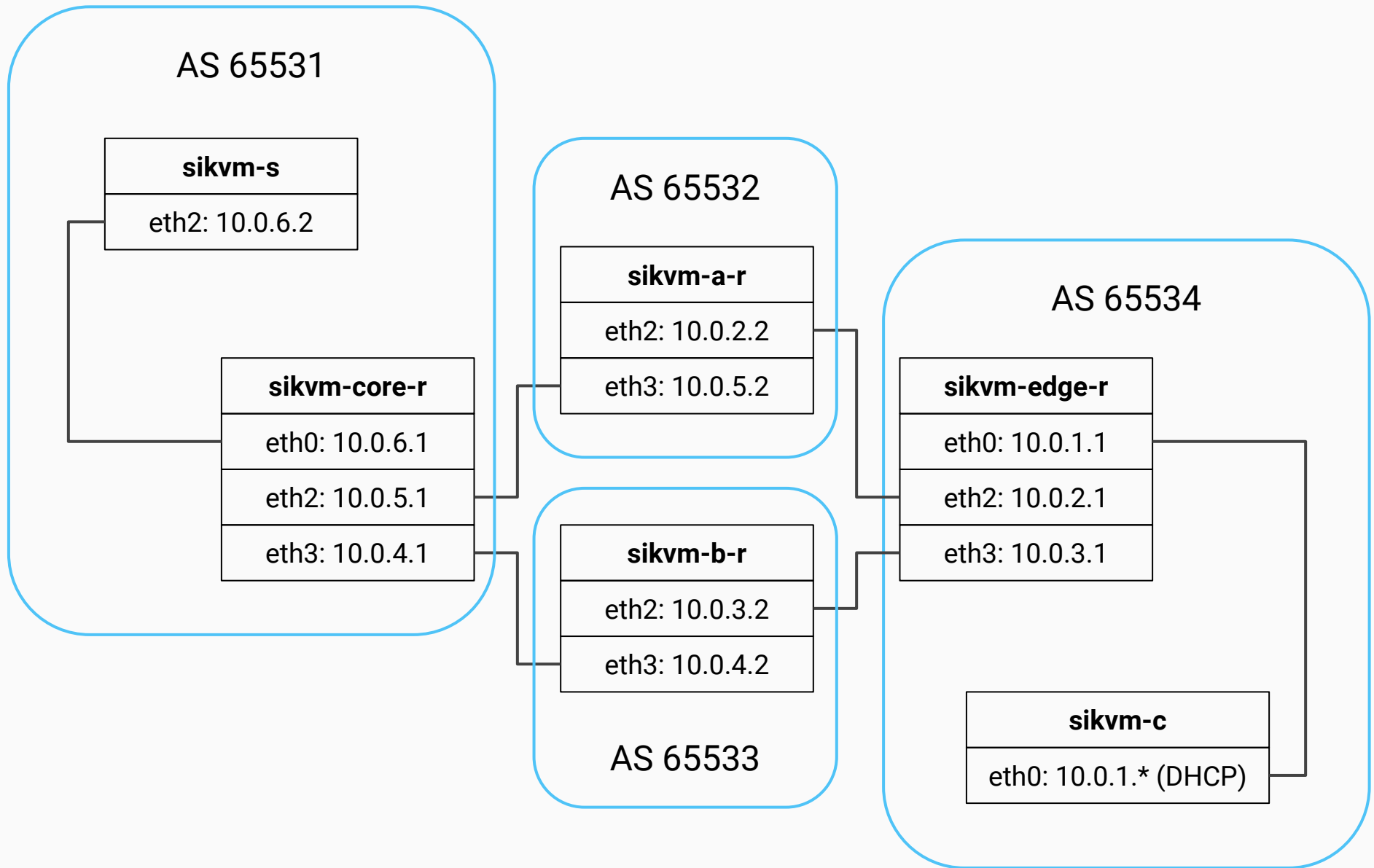
Źródło: Wikipedia

Quagga (<http://www.nongnu.org/quagga/>)

- wywodzi się z projektu Zebra
- implementuje zarówno BGPv4, jak i OSPF (oraz starsze protokoły)



Demo



Dziękuję

Szymon Acedański
WMIM UW
accek@mimuw.edu.pl

