

# ***Sposoby realizacji gwarancji jakości usług QoS w sieciach TCP/IP***

## **Wstęp**

W założeniu twórców TCP/IP stos protokołów miał umożliwiać transmisje w każdych warunkach, ale bez gwarancji jej jakości. W obecnych czasach, ze względu na wymagania stawiane usługom multimedialnym, takie podejście jest niewystarczające. W związku z tym na przestrzeni ostatnich lat opracowano szereg mechanizmów i koncepcji poprawiające realizację gwarancji jakości usług QoS (*Quality of Service*).

## **Jakość usług multimedialnych**

Usługi multimedialne mogą charakteryzować się dużym zapotrzebowaniem na pasmo. W usługach czasu rzeczywistego istotna jest wrażliwość na fluktuacje opóźnień. Inne wymagania są stawiane w przypadku transmisji danych, a inne w przypadku przekazów dźwięku. Największe wymagania mają transmisje multimedialne przy przesyłaniu obrazu ruchomego w czasie rzeczywistym (np. wideokonferencja). W zaleceniach ITU-T G.1010 [5] określono wymagania QoS (*Quality of service*) dla poszczególnych przekazów multimedialnych. Pojęcie QoS można definiować dwojako. W ujęciu sieciowym odnosi się do poziomu jakości usług oferowanych przez sieć aplikacjom, ocenianych na podstawie takich parametrów jak: opóźnienie, niezawodność (poziom utraty informacji), czy przepustowość. W ujęciu aplikacyjnym pojęcie QoS odnosi się do jakości postrzeganej przez użytkownika [1].

Użytkownik ocenia jakość przekazu na podstawie takich parametrów jak [5]:

- opóźnienie (*Delay*) - czas potrzebny do realizacji usługi liczony od momentu jej zainicjowania przez użytkownika do otrzymania pożądanej informacji, gdy usługa jest już aktywna. W zależności od typu realizowanej usługi multimedialnej wartość opóźnienia może w różnym stopniu wpływać na jakość przekazu,
- fluktuacje opóźnienia (*Delay Variation*) – charakteryzujące zmienność czasów nadejścia kolejnych pakietów co jest istotne w przypadku transmisji pakietowej. Często, dla uniknięcia ich wpływu niezbędne jest zastosowanie buforowania pakietów,
- poziom utraty informacji (*Information Loss*) - wielkość, która określa stopień degradacji transmitowanej informacji, mająca bezpośredni wpływ na jakość odbieranego przekazu. Zależy on od rodzaju transmisji multimedialnej.

W zależności od typu realizowanej usługi zmieniają się wymagania związane z jej realizacją. Podstawowe klasy usług związane są z dostarczaniem treści audio, wideo i danych. W tabeli 1. zawarte są wymagania stawiane przez różne usługi.

Tab.1. Wymagania QoS wybranych usług audio, wideo i transmisji danych [5]

Typ usługi	Aplikacja	Podstawowe parametry i wartości		
		Opóźnienie w jedną stronę [ms]	Fluktuacja opóźnienia [ms]	Utrata informacji [%]
Audio	Konwersacja	< 150 (maksymalnie 400)	< 1	< 3
	Wiadomości głosowe	< 1000 (odtwarzanie) < 2000 (nagrywanie)	< 1	< 3
	Usługi strumieniowe	< 10000	<< 1	< 1
Wideo	Wideofon	< 150 (maksymalnie 400)	Brak ograniczeń	< 1
	Usługi strumieniowe	< 10000	Brak ograniczeń	< 1
Dane	Gry interaktywne	< 200	Brak ograniczeń	0
	Telnet	< 200	Brak ograniczeń	0
	E-mail (dostęp do serwera)	Pref.. < 2000 Akceptow. < 4000	Brak ograniczeń	0
	E-mail (między serwerami)	Do kilkunastu min.	Brak ograniczeń	0
	Faks	< 30000/stronę	Brak ograniczeń	< 0,001

Podstawowym parametrem jest przepustowość łącza, zaś w przypadku usług realizowanych w czasie rzeczywistym istotne jest również opóźnienie i jego zmienność. W przypadku transmisji danych najistotniejszym parametrem jest utrata informacji, czyli niezbędne jest zastosowanie niezawodnych systemów zapewniających korekcję w przypadku pojawienia się błędów i retransmisję w przypadku utraty części informacji.

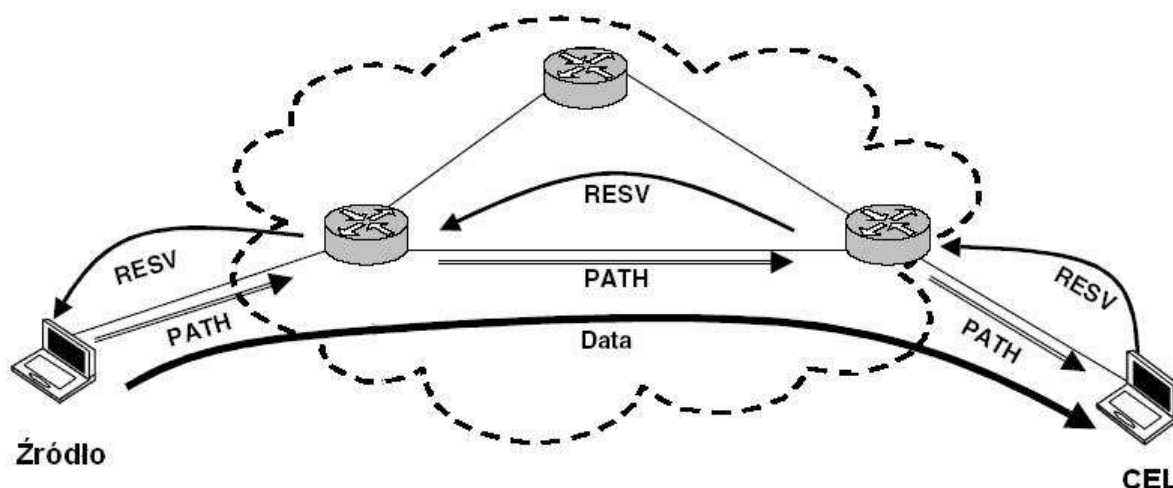
### Mechanizmy QoS

Różne aplikacje generują różne rodzaje ruchu w sieci i tym samym wymagają innego traktowania. Z tego względu zostały stworzone różne mechanizmy pozwalające klasyfikować ruch i zapewniające każdej z klas odpowiednie traktowanie.

Jednym z najistotniejszych mechanizmów QoS jest protokół rezerwacji zasobów RSVP (*Resource Reservation Protocol*). Jest to protokół sygnalizacyjny pozwalający na

skoordynowanie urządzeń sieciowych w celu rezerwacji zasobów na czas transmisji strumienia danych. RSVP pozwala zarezerwować zasoby jedynie w jedną stronę, od źródła do celu. Proces rezerwacji zasobów rozpoczyna źródło, które wysyła do celu wiadomość *Path* zawierającą wymagania do realizacji transmisji. Pakiet ten przechodząc przez poszczególne węzły w sieci informuje urządzenia o potrzebie rezerwacji zasobów na czas transmisji. Jednak na tym etapie żadne zasoby nie są jeszcze rezerwowane, gdyż zestawiana jest jedynie trasa pomiędzy źródłem, a celem. Jeśli odbiornik odbierze wiadomość *Path* i uzna, iż jest w stanie zrealizować połączenie przy zachowaniu wymaganych poziomów jakości wysyła tą samą trasą odpowiedź zwrotną *Resv*, która jest jednocześnie prośbą o rozpoczęcie transmisji o ustalonych parametrach. Wiadomość *Resv* przechodząc przez poszczególne routery rezerwuje zasoby na czas transmisji. Kiedy węzeł uzna, iż nie jest możliwa realizacja takiego połączenia odrzuca wiadomość *Resv* i informuje o tym odbiorcę. Dopiero gdy wszystkie urządzenia sieciowe na ścieżce źródło – cel uznają, że są w stanie zapewnić wymagane zasoby wiadomość *Resv* dociera do źródła, które przystępuje do transmisji strumienia. Zasoby w sieci rezerwowane są jedynie na krótkie okresu czasu ze względu na dużą dynamikę zmian w topologii sieci. Dlatego okresowo co pewien czas źródło ponownie wysyła wiadomość *Path* i czeka na wiadomość *Resv*. Oczywiście proces zachodzi zanim minie czas poprzedniej rezerwacji, co pozwala uniknąć utraty płynności transmisji. Aby sieć mogła rezerwować zasoby wszystkie jej elementy muszą mieć zaimplementowany protokół RSVP [3].

Na rys.1. zilustrowano proces rezerwacji zasobów w sieci.



Rys.1. Rezerwacja zasobów w protokole RSVP [4]

Kolejnym mechanizmem QoS jest kontrola dostępu (*Admission Control*). Mechanizm ten pozwala w węzłach na podejmowanie decyzji związanej z przyjęciem strumienia danych do sieci. Pozwala to kontrolować obciążenie w sieci i uniknąć zatorów. W przypadku transmisji bezprzewodowych musimy wziąć pod uwagę mobilność systemu. Bezprzewodowe sieci powinny przed przeprowadzeniem przełączenia do innego AP uzyskać zezwolenie na przyjęcie strumienia przez urządzenie połączone bezpośrednio z AP.

Istnieją dwie koncepcje realizacji kontroli przyjmowania strumieni. W pierwszej każde z urządzeń sieciowych podejmuje samodzielnie decyzję dotyczącą obsługi strumienia bazując na informacji o obciążeniach na bezpośrednio podłączonych łączach. W drugim podejściu w sieci istnieje urządzenie (*Bandwidth Broker*) zbierające informacje o obciążeniu z wszystkich urządzeń sieciowych i na podstawie tych danych podejmuje decyzję dotyczącą obsługi strumienia danych.

Następnym mechanizmem QoS jest klasyfikacja i znakowanie pakietów. Mechanizm pozwala na podstawie wybranego kryterium przydzielić dany pakiet do konkretnej klasy. Przydział może odbyć się na podstawie m.in. nr portu, rodzaju protokołu warstwy transportowej lub aplikacji, adresu IP źródła bądź celu. Na podstawie klasy do której został przydzielony pakiet zostaje odpowiednio zmienione pole ToS w nagłówku protokołu IPv4. Gdy oznakowany pakiet trafia do nowej sieci może ulec przeklasyfikowaniu lub nowa sieć może zaakceptować dotychczasową klasyfikację.

Klasyfikacja nieodzownie łączy się z kolejnym mechanizmem QoS, regułami kolejowania (*Queuing Disciplines*). Reguły te ustalają kolejność wysyłanych pakietów na podstawie klasy pakietu.

Najprostszą regułą kolejowania jest FIFO (First In First Out) polegająca na traktowaniu wszystkich pakietów tak samo. Jedynym parametrem jest tu czas przybycia pakietu. W regule FIFO wcześniejsze pakiety są wysyłane przed pozostałymi.

Inną regułą kolejowania jest kolejowanie priorytetowe (*Priority Queuing*) polegające na klasyfikacji pakietów w buforze wyjściowym jedynie na podstawie klasy, do której został on przydzielony.

Bardzo interesującą grupą reguł kolejowania jest grupa tzw. sprawiedliwego kolejowania (*Fair Queuing*). Pozwala w procesie ustalania kolejności uwzględnić priorytety pakietów jednocześnie zapobiegając blokowaniu ruchu o niższych klasach. Miejsce w buforze wyjściowym zostało podzielone na obszary, w których kolejność pakietów ustalana jest a podstawie priorytetu (wartości w polu ToS nagłówka protokołu IPv4).

Do mechanizmów QoS należy również zaliczyć kształtowanie ruchu (*Traffic Shaping*). Polega ono na kontroli szybkości ruchu przepływającego przez router. Jest też często nazywany miękką kontrolą dostępu. Istnieją dwa mechanizmy kształtowania ruchu.

Pierwszy nazywa się ciekącym wiadrem (*Leaky Bucket*) i jest bardzo podobne do mechanizmu kolejowania FIFO. Jediną różnicą są stałe przerwy pomiędzy pakietami czekającymi do wysłania. Przerwy te pozwalają na utrzymaniu prędkości transmisji poniżej granicznej wartości i w ten sposób zapobiegają zatorom.

Inną metodą kształtowania ruchu jest wiadro żetonów (*Token Bucket*). Mechanizm ten jest bardziej elastyczny niż poprzedni, gdyż pozwala na tymczasowe przekroczenie granicznej wartości prędkości transmisji. Jednak w pewnym odcinku czasu nie może zostać przekroczona średnia wartość szybkości transmisji. Metoda polega na zezwoleniu na transmisję pewnej liczbie pakietów w danym przedziale czasu. W zależności od potrzeb limit może być dynamicznie spożytkowany w danym przedziale czasu.

Następnym mechanizmem QoS jest reguła postępowania z pakietami niskopriorytetowymi (*Policing*). Są one oznaczane i traktowane bez żadnych gwarancji transmisji (*best-effort*). W razie niebezpieczeństwa zatoru pakiety te zostają usunięte z kolejek.

Ostatnim mechanizmem QoS opisanym w niniejszej pracy jest inżynieria ruchu (*Traffic Engineering*). Zagadnienie dotyczy przede wszystkim zastosowanych w sieci protokołów routingu mających istotny wpływ na czas uzyskania w sieci zbieżności, czyli czasu w jakim wszystkie routery będą posiadały identyczne i aktualne informacje o topologii sieci. Istotna jest również reguła i kryteria decydujące o wyborze najlepszej trasy między dwoma punktami [3].

## **Architektura QoS w sieciach IP**

Różne mechanizmy są używane w sieciach dla zagwarantowania jakości połączeń. Aby usystematyzować i skoordynować używanie tych mechanizmów wprowadzono pojęcie architektury QoS w sieciach IP. IETF zaproponowała kilka takich architektur, z których najważniejszymi są: MPLS (*Multiprotocol Label Switching*), IntServ (*Integrated Services*) i DiffServ (*Differentiated Services*).

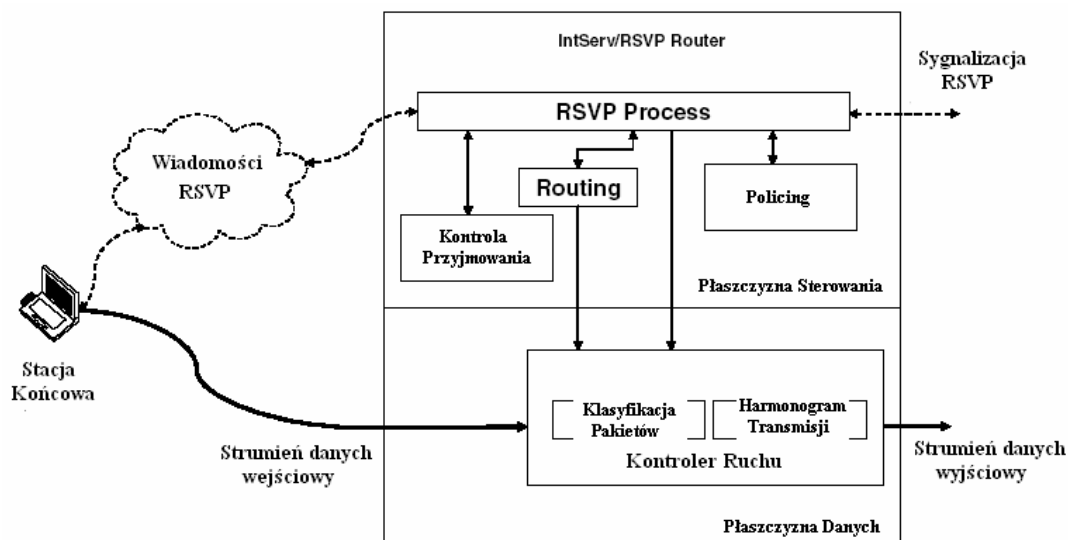
MPLS jest mechanizmem nadającym pakietom etykiety o stałej długości. Każdemu pakietowi wnikającemu do domeny MPLS zostaje dołączony nagłówek, który jest kapsułkowany pomiędzy warstwą sieciową, a warstwą łącza danych modelu sieciowego

ISO/OSI. Router z zaimplikowanym mechanizmem MPLS oznacza się skrótem LSR (*Label switching router*). Tego typu routery analizują etykietę jedynie dla pakietów wysyłanych. Pakiety opuszczające domenę MPLS są pozbawiane nagłówka MPLS. Protokół dystrybuje etykiety w domenie MPLS w celu ustanowienia stałych połączeń zwanych LSP (*Label Switched Paths*). Są one analogiczne do wirtualnych obwodów w sieciach ATM. Aby ustawić LSP używany jest protokół RSVP lub specjalnie zaprojektowany dla MPLSu protokół LDP (*Label Distribution Protocol*). Każdy z routerów LSR utrzymuje dodatkowo tablicę routingu dla etykiet MPLS. Na podstawie jej wpisów wybierana jest droga dla pakietów z etykietami. Zaletami MPLS jest szybka obsługa pakietów, a także efektywne ich tunelowanie [2].

Architektura IntServ została stworzona w 1994 roku, aby umożliwić sieciom realizację transmisji w czasie rzeczywistym z zachowaniem wymaganych parametrów QoS przy jednoczesnej transmisji danych. W architekturze tej każdy strumień jest traktowany oddzielnie. Aby zapewnić dostępność zasobów sieciowych na trasie wykorzystywany jest protokół RSVP. Architektura IntServ wprowadza dwie dodatkowe klasy ruchu (poza istniejącą best-effort) posiadające własne wymagania dotyczące QoS.

Pierwszą z klas jest klasa gwarantująca usługę (*Guaranteed service*). Została stworzona dla aplikacji wymagających pewnej stałej minimalnej wartości pasma i nie przekraczania granicznej wartości opóźnienia. Mechanizmy w tej klasie pozwalają kontrolować całkowite opóźnienie wynikające z kolejkowania w poszczególnych węzłach sieci. Fluktuacje opóźnienia są minimalizowane przez bufor u odbiorcy.

Druga klasa została nazwana klasą z kontrolowanym obciążeniem (*Controlled Load Service*). Została stworzona dla aplikacji wymagającej niezawodności ale mającej mniejsze wymagania dotyczące opóźnienia i jego fluktuacji. Głównym mechanizmem jest tu kontrola dostępu pozwalająca unikać zatorów i przeciążeń. Ruch nie jest kontrolowany przez dodatkowe mechanizmy (jak ma to miejsce w poprzednim przykładzie) na trasie, gdyż założono, iż w przypadku zaimplementowania mechanizmu zapobiegającego zatorom transmisja nie ulegnie zbyt dużym opóźnieniom. Również i w tym przypadku zasoby są rezerwowane przy pomocy RSVP [3]. Na rys.2. przedstawiono podstawowe komponenty architektury IntServ/RSVP.



Rys.2. Komponenty architektury IntServ/RSVP [4]

Głównym problemem architektury IntServ jest brak skalowalności tego typu sieci. Rezerwacja zasobów odbywa się dla każdego strumienia i musi być okresowo odświeżane. Wraz ze zwiększaniem się sieci rośnie ilość informacji sygnalizacyjnej związanej ze zwiększoną ilością strumieni w sieci. Takiego obciążenia nie są w stanie wytrzymać routery w rdzeniu sieci. Z tego względu IETF zaproponowała dla tego typu sieci architekturę DiffServ. W tym podejściu sieć nie rozpatruje każdego ze strumieni osobno lecz pojawiające się strumienie zostają przydzielone do którejś z istniejących klas. Tego typu agregacja pozwala uodpornić mechanizmy sieciowe na duży rozrost sieci. Bez względu na ilość strumieni w sieci każdy router obsługuje stałą liczbę klas. Zasady postępowania z pakietami należącymi do danej klasy należy wybrać i zaimplementować w routerach w sieci. Dodatkowe mechanizmy, jak kształtowanie ruchu, znakowanie pakietów, czy postępowanie z pakietami niskopriorytetowymi, stosuje się w routerach brzegowych. Architektura DiffServ wykorzystuje pole ToS w nagłówku protokołu IPv4 lub pole klasa ruchu w IPv6. Do znakowania pakietów używanych jest jedynie 6 bitów. Podobnie jak w architekturze IntServ routery brzegowe realizują znakowanie pakietów, kształtowanie ruchu, kontrolę dostępu i przypisanie do klas. Wewnętrzne routery jedynie agregują ruch należący do tych samych klasy. Realizacja agregacji jest zdefiniowana w każdym routerze przez PHB (*Per-Hop Behavior*) [2].

## **Literatura:**

- [1] A. & Z. Ganz, *Multimedia Wireless Networks – Technologies Standards And Qos*, Prentice Hall, USA 2003
- [2] T. Janevski, *Traffic Analysis and Design of Wireless IP Networks*, Artech House, USA 2003
- [3] D. Wong, *Wireless Internet Telecommunication*, Artech House, USA 2005
- [4] A. & N. Prasad ,*802.11 WLANs and IP Networking: Security, QoS, and Mobility*, Artech House, USA 2005
- [5] Zalecenia ITU-T G.1010, *End User Multimedia QoS Categories*, Genewa 2001

**Autor:** Piotr Horbatowski,  
Satlan ( [www.satlan.pl](http://www.satlan.pl) )