

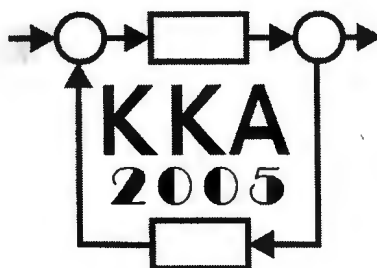
XV Krajowa Konferencja Automatyki

Tom II



**Redaktorzy:
Zdzisław Bubnicki
Roman Kulikowski
Janusz Kacprzyk**

XV Krajowa Konferencja Automatyki Tom II



Redaktorzy:
Zdzisław BUBNICKI
Roman KULIKOWSKI
Janusz KACPRZYK

ORGANIZATOR

Komitet Automatyki i Robotyki Polskiej Akademii Nauk
Instytut Badań Systemowych Polskiej Akademii Nauk

WSPÓŁORGANIZATORZY

Politechnika Warszawska

Przemysłowy Instytut Automatyki i Pomiarów

Polskie Stowarzyszenie Pomiarów, Automatyki i Robotyki

ORGANIZATOR

Komitet Automatyki i Robotyki Polskiej Akademii Nauk
Instytut Badań Systemowych Polskiej Akademii Nauk

WSPÓLORGANIZATORZY

Politechnika Warszawska
Przemysłowy Instytut Automatyki i Pomiarów
Polskie Stowarzyszenie Pomiarów, Automatyki i Robotyki

KOMITET PROGRAMOWY

Przewodniczący	Zdzisław BUBNICKI
Zastępca Przewodniczącego	Roman KULIKOWSKI

CZŁONKOWIE

Stanisław BAŃKA	Michał BIAŁKO
Mikołaj BUSŁOWICZ	Władysław FINDEISEN
Ryszard GESSING	Henryk GÓRECKI
Jakub GUTENBAUM	Jerzy JÓZEFczyk
Stanisław KACZANOWSKI	Tadeusz KACZOREK
Janusz KACPRZYK	Jerzy KLAMKA
Józef KORBICZ	Zbigniew KOWALSKI
Krzysztof KOZŁOWSKI	Juliusz L. KULIKOWSKI
Krzysztof KUŹMIŃSKI	Kazimierz MALANOWSKI
Krzysztof MALINOWSKI	Wojciech MITKOWSKI
Antoni NIEDERLIŃSKI	Władysław PEŁCZEWSKI
Tadeusz PUCHAŁKA	Leszek RUTKOWSKI
Stanisław SKOCZOWSKI	Roman SŁOWIŃSKI
Jerzy ŚWIĄTEK	Andrzej ŚWIERNIAK
Ryszard TADEUSIEWICZ	Piotr TATJEWSKI
Krzysztof TCHOŃ	Leszek TRYBUS
Jan WĘGLARZ	Andrzej P. WIERZBICKI

KOMITET ORGANIZACYJNY

Przewodniczący	Roman KULIKOWSKI
Zastępcy Przewodniczącego	Janusz KACPRZYK
	Stanisław KACZANOWSKI
	Tadeusz KACZOREK
	Krzysztof MALINOWSKI
Członkowie	Roman OSTROWSKI
	Tadeusz PUCHAŁKA
	Dariusz WAGNER
Sekretarze naukowci	Jan STUDZIŃSKI
	Jan W. OWSIŃSKI

ISBN 83-89475-01-4

Copyright © Instytut Badań Systemowych Polskiej Akademii Nauk
All rights reserved

Druk: ARGRAF, Warszawa

STEROWANIE
I TECHNIKA KOMPUTEROWA

STEROWANIE RUCHEM W PAKIETOWYCH SIECIACH KOMPUTEROWYCH DLA ZAPEWNIENIA JAKOŚCI USŁUG

Adam GRZECH

Instytut Informatyki Technicznej, Wydział Informatyki i Zarządzania, Politechnika Wroclawska
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław e-mail: adam.grzech@pwr.wroc.pl

Streszczenie: Jakość usług w sieciach z komutacją pakietów (w ogólnym przypadku jednostek danych) oznacza możliwość różnicowania obsługi ruchu wprowadzanego do sieci. W zależności od przyjętej koncepcji różnicowanie obsługi może dotyczyć albo poszczególnych przepływów w sieciach, albo klas ruchu agregujących poszczególne przepływy. Niezależnie od realizowanej w sieci koncepcji dostarczania jakości usług, każda z nich wymaga wyposażenia węzłów sieci pakietowej w mechanizmy sterowania ruchem i specyfikacji przepływów, klasyfikacji i szeregowania pakietów oraz przeciwdziałania przeciążeniom. Celem pracy jest przedstawienie wybranych mechanizmów sterowania stosowanych w realizacji różnych koncepcji dostarczania jakości usług w sieciach pakietowych.

Słowa kluczowe: sterowanie ruchem, sterowanie przepływem, przeciwdziałanie przeciążeniom, jakość usług, sieci z komutacją pakietów.

1. WPROWADZENIE

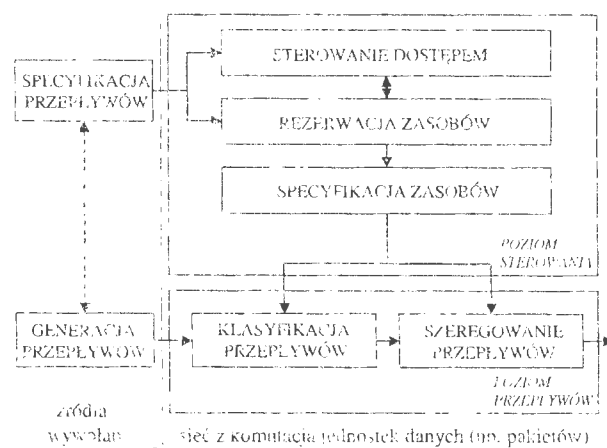
Różnicowanie jakości usług w sieciach pakietowych wymaga spełniania wielu nowych wymagań lokalizowanych w dwóch warstwach modelu sieci: sterowania (ang. *control plane*) i przepływów (transferu) danych (ang. *data path plane*) (Rys. 1).

Różnicowanie obsługi indywidualnych przepływów daje możliwość bezwzględnej gwarancji jakości usług i najczęściej związane jest ze stosowaniem różnych schematów rezerwacji zasobów sieci. Ograniczeniem koncepcji różnicowania jakości obsługi elementarnych przepływów jest skalowalność rozwiązania; rozmiar zadania sterowania jest zależny od szybkości i liczby aktywnych przepływów. Różnicowanie jakości usług dla klas ruchu, tzn. dla ograniczonej staję lub wolnozmiennej liczby klas ruchu uniezależnia rozmiar zadań sterowania w sieci pakietowej od liczby aktualnie aktywnych i obsługiwanych przepływów.

Nowe mechanizmy w warstwie transferu danych są niezbędne do wymuszania różnicowania usług sieci dla różnych klas ruchu. Są to mechanizmy klasyfikowania i przydzielania pakietów do odpowiednich klas ruchu

oraz zarządzania ilością zasobów udostępnianych różnym usługom; ich zasadniczym zadaniem jest sterowanie dostępem pakietów do zasobów sieci. Nowe mechanizmy w warstwie sterowania są niezbędne do negocjacji warunków kontraktów sieciowych, weryfikacji uprawnień pakietów do zasobów oraz alokacji zasobów do zdefiniowanych usług [1,3,7].

Zasobami sieci, które muszą być zarządzane w celu uzyskania możliwości różnicowania jakości usług sieci są pamięci buforowe i przepustowość kanałów transmisyjnych. Odpowiednie dla tych celów są różne schematy zarządzania pamięciami buforowymi i różne algorytmy szeregowania. Schematy zarządzania pamięcią buforową decydują o tym, które z pakietów mogą być przechowywane w oczekiwaniu na transmisję, podczas gdy mechanizmy szeregowania sterują transmisją pakietów. Omawiane mechanizmy są ze sobą ściśle związane, tzn. szeregowanie może wpływać na ograniczenie pojemności pamięci buforowej i odwrotnie, ograniczona pojemność pamięci buforowej wymusza stosowanie innych algorytmów szeregowania [2,3].



Rys.1. Model funkcjonalny sieci realizującej koncepcję usług zintegrowanych

Istnieje i jest stosowanych wiele różnych schematów buforowania i szeregowania pakietów w węzłach sieci. Większość ze stosowanych schematów można sklasyfikować korzystając z ogólnych wskaźników jakości

sterowania ruchem i rozdziału zasobów takich jak sprawiedliwość (ang. *fairness*) – dostęp do ograniczonych zasobów, separowalność (ang. *isolation*) – zabezpieczenie przed dodatkowym ruchem od nieuprawnionych użytkowników, efektywność (ang. *efficiency*) – liczba przepływów możliwych do obsłużenia na danym poziomie jakości, złożoność (ang. *complexity*) – zarówno w sensie implementacji, jak i nadmiaru wnoszonego przez schematy zarządzania pamięcią buforową i algorytmami szeregowania oraz skalowalność (ang. *scalability*) – zdolność mechanizmu do obsługi ruchu o dużym zakresie zmienności [3].

Przykładem mogą być różne metody zarządzania pamięcią buforową, w których do podejmowania decyzji o przyjęciu lub odrzuceniu nowego pakietu wykorzystywane są różne jakościowo i ilościowo dane. Większa granulacja posiadanej i zbieranej informacji o stanie i sposobie zajęcia pamięci buforowej pozwala na uzyskanie większej sprawiedliwości i efektywności, ale za cenę większej złożoności obliczeniowej.

Podobne, jak w warstwie transferu danych, wskaźniki jakości mechanizmów są stosowane dla oceny mechanizmów użytych w warstwie sterowania. Odnoszą się one zarówno do klas dostępnych usług sieciowych, jak i sposobów w jaki usługi te są wywoływane. Przykładem może być gwarancja jakości usług; gwarancja jakości może mieć charakter jakościowy albo ilościowy; w tym ostatnim przypadku może być opisana wartościami jednego lub kilku parametrów.

W ogólnym przypadku do oceny mechanizmów stosowanych w warstwie sterowania wykorzystywane są: złożoność obliczeniowa procedur niezbędnych do realizacji usługi oraz ilość i jakość informacji potrzebnych do podejmowania decyzji. Ponadto występuje wzajemna zależność pomiędzy wskaźnikami oceny mechanizmów stosowanych jednocześnie w warstwie transferu danych i sterowania. Przykładowo, duża granulacja ruchu w warstwie transferu danych wymusza obsługę pojedynczych żądań dostępu w warstwie sterowania, a z kolei duża granulacja w warstwie transferu danych jest zbędna, jeżeli decyzje o dostępie do sieci podejmowane są na podstawie wartości parametrów ruchu zagregowanego.

2. STEROWANIE W RÓŻNYCH KONCEPCJACH DOSTARCZANIA JAKOŚCI USŁUG

Wzrost liczby zastosowań sieci pakietowych do obsługi ruchu o różnych wymaganiach powoduje konieczność zmian w infrastrukturze sieci. Wielość jednocześnie obsługiwanych aplikacji o różniących się wymaganiach dotyczących przepustowości sieci oraz opóźnień i strat pakietów wymusza stosowanie rozwiązań sieciowych zdolnych różnicować jakość usług.

Sieć teleinformatyczna jest zbiorem współdzielonych zasobów przeznaczonych do jednoczesnej obsługi ruchu elastycznego i strumieniowego generowanego przez aplikacje (użytkowników) współzawodniczące o zasoby sieci [1,7].

Realizacja zadania dostarczania różnej jakości usług sieci polega na implementacji w sieci mechanizmów sterowania zdolnych różnicować:

- zasady dostępu do zasobów (ang. *admission control*),
- zasady przyznawania (alokacji) zasobów (ang. *resource allocation*),
- ilości zasobów przyznawanych użytkownikom.

Przykładami różnych architektur różnicowania jakości usług sieciowych są koncepcje usług:

- dostarczanych bez gwarancji jakości (ang. *best effort*), zgodnie z którą alokacja zasobów jest wynikiem współzawodnictwa przepływów,
- zintegrowanych (ang. *Integrated Services*), zgodnie z którą alokacja zasobów odbywa się przez ich rezerwację dla poszczególnych przepływów generowanych przez aplikacje,
- zróżnicowanych (ang. *Differentiated Services*), zgodnie z którą alokacja zasobów odbywa się przez ich udostępnianie zróżnicowanym klasom ruchu; efektywność alokacji zasobów uzyskuje się przez zarządzanie ilością ruchu wprowadzanego do sieci, albo ruchu jako całości, albo w obrębie wyróżnionych klas ruchu.

Wyróżnione trzy koncepcje dostarczania jakości usług sieciowych nie są jednoznaczne i wykluczające się.

Koncepcja obsługi bez gwarancji jakości usług wynika wprost z filozofii sieci z komutacją wiadomości i opiera się na założeniu, że jednakową (w sensie średnim) jakość usług można uzyskać przez reglamentację ilości ruchu (w zasadzie elastycznego) wprowadzanego do sieci. Z tego punktu widzenia koncepcja ta ma wiele wspólnego z usługą sterowanego obciążenia w koncepcji usług zintegrowanych i koncepcją usług zróżnicowanych.

Koncepcje usług zintegrowanych i zróżnicowanych różni od koncepcji sieci bez gwarancji jakości usług założenie o konieczności i możliwości różnicowania jakości usług sieci. Koncepcje usług zintegrowanych i zróżnicowanych różni między sobą z kolei filozofia dostarczania jakości usług.

Koncepcja usług zintegrowanych opiera się na założeniu, że efektywnym sposobem dostarczania różnej jakości usług jest sterowanie poszczególnymi przepływami; ograniczeniem zastosowań tej koncepcji jest skalowalność.

Koncepcja usług zróżnicowanych zakłada natomiast, że skutecznym sposobem dostarczania jakości usług jest różnicowanie ich jakości dla wyróżnianych klas ruchu, tzn. dla przepływów zagregowanych.

3. PRZECIWDZIAŁANIE PRZECIĄŻENIOM

Możliwość powstawania przeciążeń w sieciach pakietowych jest prostą konsekwencją zasad organizacji obsługi ruchu w takich sieciach. Sieć z komutacją pakietów jest systemem z opóźnieniami (ang. *delay system*), tzn. ilość ruchu wprowadzanego w danej chwili do sieci może być większa od ilości zasobów niezbędnych do obsłużenia wprowadzanego ruchu [5].

Podstawowym kryterium podziału metod sterowania ruchem w sieciach na prewencyjne i reakcyjne jest rodzaj danych, na podstawie których podejmowane są decyzje i oczekiwany czas reakcji na te decyzje. W metodach prewencyjnych decyzje podejmowane są w zasadzie na podstawie danych *a priori*, dotyczących danego strumienia pakietów i dostępnych zasobów. W metodach reakcyjnych zazwyczaj są uwzględniane dane

a posteriori, opisujące skutki wcześniej podjętych decyzji o przyjęciu wywołań i przydziale zasobów.

Zadaniem technik sterowania prewencyjnego jest zapobieganie przeciążeniom w wyniku podejmowania działań uprzedzających – tzn. zmierzających do zapobiegania zdarzeniom, których następstwem są przeciążenia. Sterowanie prewencyjne nie eliminuje przeciążeń i nie jest wystarczające do zapobiegania przeciążeniom, a jeżeli już przeciążenie powstanie, niezbędne są działania zmierzające do jego usunięcia. Dzieje się tak dlatego, że efektywność metod prewencyjnych zwiększa się wtedy, gdy zasoby sieci są rozdzielane deterministycznie, a to z kolei powoduje niewielką efektywność wykorzystania zasobów.

Zadaniem metod reakcyjnych jest odtworzenie stanu sieci poprzedzającego powstanie stanu przeciążenia (rozładowanie przeciążenia) i odbywa się zwykle zgodnie z pewnymi ogólnymi zasadami. Detekcja stanu przeciążenia oznacza uruchomienie wcześniej zdefiniowanych procedur, których celem jest na ogół spowolnienie lub wstrzymanie napływu nowych żądań obsługi aż do chwili likwidacji stanu przeciążenia.

Głównym ograniczeniem skuteczności mechanizmów reakcyjnych jest wartość ilorazu czasu propagacji do czasu transmisji pakietów. Wzrost jego oznacza zwiększenie prawdopodobieństwa, że podjęta akcja może być spóźniona. Zjawisko to jest określane mianem bezwładności sieci, a jego miarą jest wartość iloczynu czasu propagacji jednostki danych i obciążenia sieci.

Innym kryterium wyróżniania kategorii mechanizmów sterowania ruchem w sieci są różne długości przedziału czasu, upływającego od chwili podjęcia decyzji do chwili pojawienia się skutków podjętej decyzji. Mechanizmy prewencyjne praktycznie nie mają takich ograniczeń; są zwykle stosowane w punktach dostępu do usług sieci. Skuteczność mechanizmów reakcyjnych zależy od szybkości zmian procesów, których dotyczą. Mechanizmy te działają (są skuteczne) w czasie nie krótszym od czasu propagacji [1].

4. STEROWANIE STRUMIENIEM PAKIETÓW

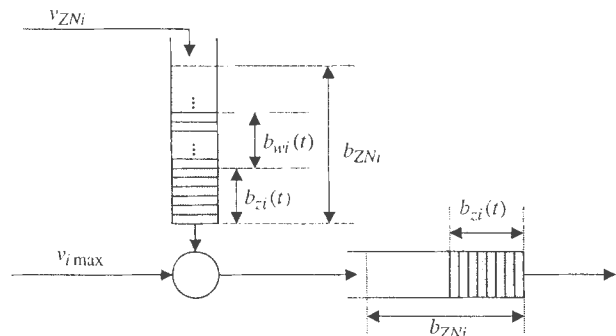
Jednym z podstawowych algorytmów sterowania strumieniami pakietów wprowadzanych do sieci z gwarancją jakości jest algorytm ciekącego wiadra (ang. *leaky bucket*). Sterownik z tym algorytmem (Rys.2) może być użyty do szacowania wartości parametrów charakteryzujących stan wykorzystania zasobów, w tym średniego i maksymalnego opóźnienia wnoszonego przez sterownik oraz zaległości (ang. *backlogged*) w obsłudze przepływu, wynikających z proporcji ilości przydzielonych i żądanych zasobów. Wartości parametrów algorytmu pozwalają zarówno na specyfikację ruchu wprowadzanego do sieci (kontrakt ruchowy), jak i na sprawdzanie zgodności rzeczywiście wprowadzanego ruchu z kontraktem ruchowym [2-4].

Jeżeli $a_i(\tau, t)$ oznacza ilość ruchu dla i -tego przepływu, wyprowadzaną ze sterownika do pamięci buforowej pakietów oczekujących na obsługę w sieci, w przedziale czasu $(\tau, t]$, to (Rys. 3):

$$a_i(\tau, t) \leq \min\{(t - \tau)C_i, b_{ZNi} + v_{ZNi}(t - \tau)\},$$

a i -ty przepływ jest tzw. przepływem (b_{ZNi}, v_{ZNi}, C_i) - zgodnym.

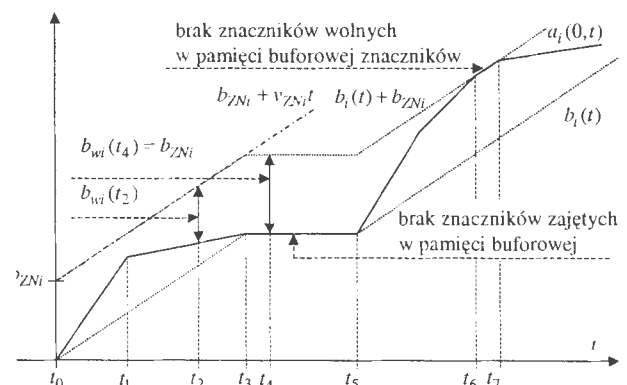
Funkcja $a_i(0, t)$ jest ciągłą i niemalejącą funkcją czasu t . Założenie, że $a_i(\tau, t)$ jest ilością ruchu, oznacza przyjęcie przepływowego modelu ruchu, tzn. wartość $a_i(0, t)$ jest ciągłą aproksymacją dyskretniej liczby pakietów i -tego przepływu, wyprowadzanych ze sterownika w przedziale czasu $(0, t]$.



Rys.2. Sterowanie szybkością wprowadzania pakietów do sieci - algorytm ciekącego wiadra

Ograniczenia szybkości wyprowadzania pakietów i -tego przepływu ze sterownika do sieci, dotyczą:

- średniej szybkości pakietów; ograniczeniem jest szybkość napływu znaczników (v_{ZNi}),
- szybkości szczytowej przepływu; ograniczeniem jest stała szybkość obsługi pakietów (C_i),
- wybuchowości źródła przepływu; ograniczeniem jest pojemność pamięci buforowej znaczników (b_{ZNi}).



Rys.3. Zależności pomiędzy liczbą znaczników wolnych, całkowitą liczbą znaczników przyjętych i ilością ruchu wyprowadzonego ze sterownika

Ilość ruchu i -tego przepływu, wyprowadzaną ze sterownika do sieci w przedziale $(\tau, t]$, można wyrazić:

- liczbą znaczników wprowadzonych do pamięci buforowej znaczników w przedziale czasu $(\tau, t]$,
- liczbą wolnych znaczników w pamięci buforowej na początku i końcu przedziału czasu $(\tau, t]$.

Jeżeli $b_i(t)$ oznacza całkowitą liczbę znaczników przyjętych do pamięci buforowej znaczników w przedziale czasu $(0, t]$ i liczba ta nie zawiera znaczników znajdujących się w pamięci buforowej w chwili początkowej i znaczników nieprzyjętych z powodu przepełnienia pamięci buforowej, to:

$$b_i(t) = \min \{a_i(0, \tau) + v_{ZNi}(t - \tau)\},$$

$$0 \leq \tau \leq t$$

co oznacza, że:

$$b_i(t) - b_i(\tau) \leq v_{ZNi}(t - \tau).$$

Liczba wszystkich znaczników w pamięci buforowej, w dowolnej chwili t dla i -tego przepływu jest sumą liczb znaczników wolnych $b_{wi}(t)$ i znaczników zajętych $b_{zi}(t)$:

$$(b_{wi}(t) + b_{zi}(t) \leq b_{ZNi}).$$

Liczba znaczników zajętych jest równa liczbie pakietów przyjętych do obsługi i oczekujących w pamięci buforowej na obsługę w pamięci buforowej pakietów.

Liczba znaczników wolnych w pamięci buforowej znaczników dla i -tego przepływu w dowolnej chwili t jest równa sumie liczby znaczników w pamięci buforowej znaczników w chwili początkowej (dla uproszczenia liczba znaczników w pamięci buforowej w chwili początkowej jest równa pojemności pamięci b_{ZNi}) i całkowitej liczby znaczników przyjętych do pamięci buforowej znaczników w przedziale czasu $(0, t]$ pomniejszonej o liczbę znaczników wykorzystanych w przedziale czasu $(0, t]$ (ilość ruchu wyprowadzonego przez sterownik do sieci w przedziale czasu $(0, t]$):

$$b_{wi}(t) = b_{ZNi} + b_i(t) - a_i(0, t)$$

Z zależności tej wynika, że ilość ruchu i -tego przepływu, wyprowadzonego ze sterownika w przedziale czasu $(\tau, t]$ można wyrazić nierównością:

$$a_i(\tau, t) = a_i(0, t) - a_i(0, \tau) = [b_{ZNi} + b_i(t) - b_{wi}(t)] -$$

$$-[b_{ZNi} + b_i(\tau) - b_{wi}(\tau)] \leq$$

$$\leq b_{wi}(\tau) + v_{ZNi}(t - \tau) - b_{wi}(t).$$

Generowanie pakietów i -tego przepływu rozpoczyna się w chwili t_0 , kiedy liczba znaczników wolnych jest równa pojemności pamięci buforowej znaczników b_{ZNi} .

W przedziale (t_0, t_1) szybkość generowania pakietów i -tego przepływu jest większa od szybkości generowania znaczników – liczba wolnych znaczników maleje.

W przedziale (t_1, t_3) szybkość generowania pakietów i -tego przepływu jest mniejsza od szybkości generowania znaczników – liczba wolnych znaczników rośnie.

W przedziale (t_3, t_5) źródło nie generuje nowych pakietów; liczba znaczników wolnych jest równa pojemności pamięci buforowej znaczników i nowe znaczniki są odrzucane.

W chwili t_5 źródło ponownie rozpoczyna generowanie pakietów – liczba znaczników wolnych zmniejsza się. W przedziale (t_6, t_7) szybkość napływu pakietów jest ograniczona szybkością napływu znaczników.

Maksymalna ilość ruchu, wyprowadzana przez sterownik w przedziale $(0, t]$, jest nie większa od sumy pojemności pamięci buforowej znaczników b_{ZNi} i liczby znaczników generowanych w tym przedziale $v_{ZNi}t$, tzn.:

$$a_i(0, t) \leq b_{ZNi} + v_{ZNi}t, \text{ gdy } tC_i \geq b_{ZNi} + v_{ZNi}t$$

W takim przypadku średnia szybkość generowania pakietów i -tego przepływu przez źródło:

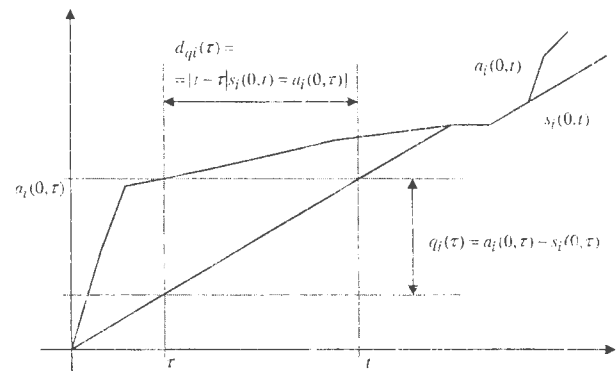
$$\bar{v}_i = v_{ZNi} + b_{ZNi}/t.$$

Opóźnienia pakietów i -tego przepływu w chwili czasu t , wnoszone przez działanie sterownika, zależą od różnicy pomiędzy ilością ruchu wyprowadzonego przez sterownik do chwili czasu t i ilości ruchu obsłużonego przez sieć do chwili czasu t .

Jeżeli przez $s_i(0, t)$ oznaczyć ilość ruchu i -tego przepływu, wyprowadzonego ze sterownika i obsłużonego przez sieć w przedziale czasu $(0, t]$, to możliwe jest zdefiniowanie dwóch miar jakości obsługi generowanego ruchu: zaległości w obsłudze ruchu i -tego przepływu i opóźnienia wnoszonego przez sieć.

Zaległość w obsłudze ruchu i -tego przepływu w chwili τ , oznaczona przez $q_i(\tau)$, jest definiowana jako różnica ilości ruchu wyprowadzonego ze sterownika $a_i(0, \tau]$ i ilości ruchu obsłużonego $s_i(0, \tau]$ w przedziale $(0, \tau]$ (Rys. 4):

$$q_i(\tau) = a_i(0, \tau) - s_i(0, \tau).$$



Rys. 4. Zaległości i opóźnienia w obsłudze ruchu

Maksymalna wartość zaległości w obsłudze ruchu i -tego przepływu jest ograniczona z góry maksymalną długością kolejki pakietów oczekujących na obsługę w sieci, tj. pojemnością pamięci buforowej znaczników:

$$q_{i \max}(\tau) = \max[a_i(0, \tau) - s_i(0, \tau)] = b_{ZNi}.$$

Występowanie zaległości w obsłudze ruchu i -tego przepływu w chwili τ oznacza konieczność oczekiwania na obsługę przez sieć, tzn. powstanie zaległości oznacza opóźnienie ruchu i -tego przepływu. Opóźnienie ruchu i -tego przepływu w chwili τ , oznaczone przez $d_{qi}(\tau)$, jest definiowane jako długość przedziału czasu potrzebnego do likwidacji powstałej zaległości:

$$d_{qi}(\tau) = \inf\{t \geq \tau : s_i(0, t) = a_i(0, \tau)\} - \tau.$$

Miarą opóźnienia obsługi ruchu i -tego przepływu $d_{qi}(\tau)$, w chwili τ jest odległość w poziomie wykresów funkcji $a_i(0, t)$ i $s_i(0, t)$ dla wartości rzędnej $a_i(0, \tau)$ (Rys. 4). Długość tego przedziału czasu jest równa, z dokładnością do różnicy pomiędzy modelem dyskretnym i przepływowym ruchu i -tego przepływu, czasowi oczekiwania na obsługę ostatniego pakietu i -tego przepływu wyprowadzonego ze sterownika i oczekującego na obsługę przez sieć.

Jeżeli maksymalna wartość zaległości w obsłudze ruchu i -tego przepływu jest równa pojemności pamięci buforowej znaczników b_{ZNi} , a pakiety oczekujące

w kolejce są obsługiwane ze stałą szybkością C_i , to maksymalna wartość opóźnienia jest równa:

$$d_{qi\max} = b_{zNi} / C_i.$$

5. MECHANIZMY SZEREGOWANIA

Rodzaj gwarancji jakości usług sieciowych zależy od stosowanych mechanizmów szeregowania, które pełnią funkcje arbitra dla pakietów oczekujących na obsługę (transmisję) [1,3].

Zastosowanie jednego z możliwych mechanizmów szeregowania zależy od wymagań co do jakości usług, tzn. od wybranych kryteriów jakości obsługi pakietów.

Do częściej stosowanych kryteriów mechanizmów szeregowania należą [2,3]:

- granulacja miar jakości przepływów (szeregowanie na poziomie przepływów lub klas ruchu),
- efektywność (poziom wykorzystania zasobów dla założonej gwarancji jakości usług),
- elastyczność (możliwość oddzielnego sterowania wartościami różnych kryteriów jakości usług),
- sprawiedliwość (elastyczność i równomierność rozdziału dostępnych zasobów).

5.1. Metody szeregowania

W większości systemów obsługi pakietów przepływów stanowisko obsługi jest wolne tylko wtedy, gdy kolejka pakietów oczekujących na obsługę (transmisję) jest pusta. Algorytmy takie są określane mianem algorytmów szeregowania zachowujących pracę (ang. *work conserving*). Większość algorytmów szeregowania i regulaminów obsługi należy do tej grupy algorytmów.

Inną grupą algorytmów szeregowania są algorytmy niezachowujące pracy (ang. *non-work conserving*); pakiety są obsługiwane (transmitowane) w węzłach wtedy, gdy zostaną wybrane do obsługi, tzn. gdy wymuszona zostanie alokacja zasobów. Jeżeli zaden z pakietów oczekujących na obsługę w kolejce nie zostanie wybrany do obsługi, to stanowisko obsługi pozostaje wolne, mimo że w kolejce znajdują się pakiety oczekujące na obsługę.

Algorytmy szeregowania dzielone są na grupy:

- sprawiedliwego szeregowania (ang. *fair queueing algorithms*),
- z gwarantowanym czasem obsługi (ang. *deadline based algorithms*),
- warunkowego szeregowania (ang. *rate based algorithms*).

5.1.1. Algorytmy sprawiedliwego szeregowania

Działanie algorytmów sprawiedliwego szeregowania pakietów opiera się na założeniu, że część zasobów (np. pojemności łącza) przydzielanych danemu przepływowi jest reprezentowana przez liczbę rzeczywistą oznaczającą wartość współczynnika wagowego (ang. *weight coefficient*).

Algorytm sprawiedliwego szeregowania przydziela zasoby przepływowi aktywnym proporcjonalnie do wartości ich wag. Każdy przepływ jest upoważniony do dostępu do zasobów proporcjonalnie do przypisanej mu

wagi. Ilość przydzielonych przepływowi zasobów, w porównaniu z zapotrzebowaniem na nie, jest podstawą kwalifikacji przepływów do dwóch grup przepływów:

- bez zaległości (ang. *non-backlogged flow*), tzn. takich, którym zostały przydzielone zasoby większe lub równe aktualnemu zapotrzebowaniu; kolejka pakietów oczekujących na obsługę jest pusta,
- z zaległościami (ang. *backlogged flow*), tzn. przepływów, którym przydzielone zostały zasoby mniejsze od aktualnego zapotrzebowania; skutkiem tego długość kolejki pakietów oczekujących na obsługę jest różna od zera.

Jeżeli przepływ bez zaległości nie wykorzystuje w pełni przydzielonych mu zasobów, to niewykorzystane (nadmiarowe) zasoby są przydzielane przepływowi z zaległościami, proporcjonalnie do wag wszystkich przepływów z zaległościami. Zgodnie z algorytmem sprawiedliwego szeregowania każdy z przepływów upoważnionych do współdziałania w korzystaniu z zasobów (w procesie rezerwacji) proporcjonalnie do przypisanej mu wagi, ma gwarancję otrzymania zarezerwowanych zasobów i możliwość uzyskania dodatkowych. Gwarancja dostępu do zarezerwowanych zasobów w każdym węzle, znajdującym się na trasie od źródła do ujścia, jest równoznaczna z możliwością gwarantowania ograniczonych z góry opóźnień pakietów na trasach, tzn. możliwością gwarantowania podstawowej właściwości jakości usług w koncepcji usług zintegrowanych.

5.1.2. Algorytmy szeregowania z nieprzekraczalnym czasem obsługi

W algorytmach szeregowania z nieprzekraczalnym czasem obsługi kolejność obsługi pakietów jest zgodna z uporządkowaniem pakietów według niemalejących wartości nieprzekraczalnego czasu obsługi EDF (ang. *Earliest Deadline First*), tzn. jako pierwszy jest obsługiwany pakiet z najmniejszą wartością nieprzekraczalnego czasu obsługi. Do działania algorytmu wymagane jest przypisanie każdemu pakietowi jednoznacznej wartości nieprzekraczalnego (ograniczonego z góry) czasu obsługi (ang. *deadline*).

Cechą charakterystyczną działania omawianych algorytmów jest rozłączność górnego ograniczenia opóźnienia pakietów i alokacji zasobów, tzn. ilość zasobów przydzielonych dla przepływu nie determinuje jednoznacznie wartości górnego ograniczenia opóźnienia pakietów wnoszonego przez sieć. Cecha ta powoduje m.in. to, że omawiane algorytmy na ogół nie są algorytmami sprawiedliwego szeregowania jednostek danych, tzn. nadmiarowe zasoby nie są rozdzielane sprawiedliwie pomiędzy przepływy.

Możliwość przypisywania pakietom każdego przepływu wartości nieprzekraczalnego czasu obsługi, niezależnie od ilości przypisanych danemu przepływowi zasobów powoduje, że kolejności napływu i wypływu pakietów do i z węzła, należących do tych samych i różnych przepływów, mogą być różne.

W algorytmach, w których nie ma możliwości zmiany kolejności napływających i obsługiwanych pakietów, względna ilość zasobów (udział w całkowitej ilości zasobów) przyznanych przepływowi determinuje wartość górnego ograniczenia opóźnienia pakietów.

Zmniejszenie ilości zasobów powoduje wzrost wartości górnego ograniczenia opóźnienia.

Zastosowanie algorytmu z nieprzekraczalnym czasem obsługi, tzn. algorytmu szeregowania pakietów wnoszącego niezależność górnego ograniczenia opóźnienia od ilości przyznanego zasobów, wymaga względnie złożonych systemów sterowania dostępem. Złożoność ta wynika z konieczności implementacji dwukryterialnego sterowania dostępem, tzn. sterowania:

- zapobiegającego przydzielaniu przepływowi zasobów w ilościach większych od całkowitej ilości zasobów,
- nadzorującego realizowalność obsługi pakietów w nieprzekraczalnym czasie ich obsługi.

5.1.3. Algorytmy szeregowania warunkowego

Algorytmy szeregowania warunkowego spełniają jednocześnie dwa zadania: sterowania przepływem pakietów i właściwego ich szeregowania.

Sterownik służy do wyznaczania warunków czasowych dla napływających pakietów przepływu, które następnie są wybierane do obsługi zgodnie z implementowanym algorytmem szeregowania. W ten sposób sterownik realizuje różne zadania kształtowania strumienia pakietów wprowadzanych do obsługi, w tym m.in.:

- kompensowanie różnic szybkości napływu i wypływu,
- spowalniania wyprowadzania pakietów w tych przedziałach czasu, w których pakiety napływają z maksymalną szybkością,
- ograniczenie zakresu zmienności (fluktuacji) opóźnienia wprowadzanych pakietów, pozwalające na ograniczenie różnic opóźnień pakietów wnoszonych przez te elementy trasy przepływu, które poprzedzają dany węzeł na trasie.

Złożenie różnych algorytmów sterowania przepływem oraz szeregowania pakietów wyprowadzanych ze sterownika daje możliwość budowania algorytmów szeregowania warunkowego o różnych właściwościach, dla szerokiego zakresu zmienności wartości parametrów źródeł przepływów i jakości usług.

5.2. Sprawiedliwy rozdział zasobów

Wynikiem działania algorytmów sprawiedliwego szeregowania FQ (ang. *Fair Queueing*) w węzle jest równomierny rozdział zasobów, w ilościach odwrotnie proporcjonalnych do liczby aktywnych przepływów.

W sieciach, w których możliwa jest zróżnicowana jakość usług dla różnych przepływów, podstawowym zadaniem dyscyplin szeregowania stosowanych w węzłach sieci jest przydział różnych ilości zasobów różnym przepływom, umożliwiających zarówno realizację żądanej jakości usług, jak i różnicowanie jakości usług dla różnych przepływów w sieci.

Niezależnie od różnego rozumienia gwarancji jakości usług w sieciach, naturalne jest wymaganie, aby jakość usług sieci była przewidywalna; przewidywalność dotyczy np. znajomości górnego ograniczenia wartości opóźnienia wnoszonego przez sieć w czasie transferu pakietów na trasie od źródła do ujścia. Do najczęściej stosowanych dyscyplin sprawiedliwego szeregowania,

dających możliwość przewidywania jakości usług dla ruchu w sieciach, w tym dla ruchu w czasie rzeczywistym, zalicza się dyscypliny: z uogólnionym podziałem czasu procesora GPS (ang. *Generalized Processor Sharing*) oraz z nieprzekraczalnym czasem obsługi EDF (ang. *Earliest Deadline First*) [6]. Działanie GPS jest oparte na modelu przepływowym zakładającym możliwość nieskończonej granulacji przepływu. Ze względu na to, że przepływ pakietów nie spełnia takich wymagań, GPS jest dyscypliną odniesienia dla jej wersji pakietowych PGPS (ang. *Packetized GPS*).

W zadaniach alokacji stałych pojemności dla przepływów i gwarantowania górnego ograniczenia opóźnienia pakietów stosowane są różne algorytmy ważonego sprawiedliwego szeregowania (kolejkowania) WFQ (ang. *Weighted Fair Queueing*), które umożliwiają przydział różnych ilości zasobów dla różnych aktywnych przepływów.

Algorytmy WFQ są praktycznymi realizacjami dyscypliny szeregowania z uogólnionym podziałem czasu procesora GPS. Praktyczna realizacja dyscypliny szeregowania GPS przez algorytm WFQ polega na tym, że WFQ emuluje dyscyplinę szeregowania GPS, ale do obsługi wybierane są jednostki danych (pakiety). Większość różnych wariantów WFQ jest porównywana z dyscypliną GPS będącą modelem odniesienia.

5.2.1. Uogólniony podział czasu procesora

Podstawowa zasada działania dyscypliny GPS polega na tym, że każdy i -ty przepływ, przechodzący przez serwer GPS, obsługujący jednocześnie n przepływów i działający ze stałą szybkością C , jest charakteryzowany dodatnią liczbą rzeczywistą ϕ_i , taką, że dla każdego przedziału czasu $(\tau, t]$, w którym i -ty przepływ jest przepływem z zaległościami, spełniony jest warunek:

$$\frac{s_i(\tau, t)}{s_j(\tau, t)} \geq \frac{\phi_i}{\phi_j} \text{ dla } j = 1, 2, \dots, n,$$

gdzie $s_j(\tau, t)$ jest ilością ruchu j -tego przepływu obsługiwanego przez serwer GPS w przedziale $(\tau, t]$.

Podstawowe właściwości dyscypliny GPS to:

- gwarancja minimalnej jakości obsługi dla każdego z obsługiwanych przepływów,
- sprawiedliwy rozdział zasobów pomiędzy przepływy,
- możliwość wyprowadzenia prostych reguł sterowania dostępem w przypadku deterministycznych ograniczeń całkowitego opóźnienia pakietów w sieci.

Ze względu na te właściwości dyscyplina GPS jest, jako model odniesienia, podstawą wielu odmian algorytmów szeregowania, różniących się stopniem przybliżania rzeczywistych procesów, jakością działania, złożonością obliczeniową, złożonością implementacji i stopniem wykorzystywania zasobów.

5.2.2. Sprawiedliwość dyscypliny szeregowania

Jedną z wielu definicji sprawiedliwości szeregowania (ang. *scheduling fairness*) jest następująca. Jeżeli:

- obsługiwane przepływy są nieskończenie podzielne, a ich obsługa jest obsługą dowolnie małych kwantów,

- i -ty i j -ty przepływy ($i, j = 1, 2, \dots, n; i \neq j$) są dwoma dowolnymi, które w danym przedziale czasu $(t_1, t_2]$ są przepływami z zaległościami,
- $s_i(t_1, t_2)$ i $s_j(t_1, t_2)$ są, odpowiednio, ilościami ruchu i -tego i j -tego przepływów obsługowanego przez serwer w przedziale czasu $(t_1, t_2]$,

to dyscyplina szeregowania realizowana przez serwer, której rezultatem działania jest przydział pojemności C_i i C_j , odpowiednio dla i -tego i j -tego przepływów, jest sprawiedliwa wtedy i tylko wtedy, gdy dla $\forall t_1, t_2$ i $\forall i, j$, takich, że i -ty i j -ty przepływy są aktywne w przedziale czasu $(t_1, t_2]$ spełniony jest warunek:

$$\left| \frac{s_i(t_1, t_2)}{C_i} - \frac{s_j(t_1, t_2)}{C_j} \right| = 0.$$

Interpretacja tego warunku jest następująca: stopień wykorzystania zasobów przydzielonych poszczególnym przepływom z zaległościami jest taki sam. Warto zwrócić uwagę, że wartości ilorazów $s_i(t_1, t_2)/C_i$ i $s_j(t_1, t_2)/C_j$ są jednakowe i mogą być mniejsze, równe lub większe od jedności. Wartości ilorazów są mniejsze od jedności, gdy wszystkie przepływy są przepływami z zaległościami, i są większe od jedności, gdy nie wszystkie przepływy są przepływami z zaległościami.

Podana definicja sprawiedliwości dyscypliny szeregowania jest spełniana tylko w przypadku idealnego, nieskończenie podzielonego przepływu. W rzeczywistych sieciach, w których obsługa ruchu polega na obsłudze dyskretnych jednostek danych o zmiennej długości, definicja sprawiedliwości jest modyfikowana do następującej: realizowana przez serwer dyscyplina szeregowania jest sprawiedliwa wtedy i tylko wtedy, gdy dla $\forall t_1, t_2$ i $\forall i, j$, takich, że i -ty i j -ty przepływy są aktywne w przedziale $(t_1, t_2]$, spełniony jest warunek:

$$\left| \frac{s_i(t_1, t_2)}{C_i} - \frac{s_j(t_1, t_2)}{C_j} \right| = \beta,$$

gdzie β jest dolnym ograniczeniem sprawiedliwości rozdziału zasobów, tzw. indeksem sprawiedliwości alokacji zasobów (ang. *fairness index*).

Im mniejsza wartość indeksu sprawiedliwości, tym bardziej sprawiedliwa jest dyscyplina szeregowania.

W sieciach z komutacją pakietów najmniejsza wartość indeksu sprawiedliwości dowolnej dyscypliny szeregowania nie może być mniejsza od czasu transmisji pakietu o największej długości. Jeżeli:

- i -ty oraz j -ty przepływy są rzeczywistymi przepływami z zaległościami, aktywnymi w przedziale czasu $(t_1, t_2]$,
- maksymalne długości pakietów tych przepływów są, odpowiednio, równe $l_{i \max}$ oraz $l_{j \max}$,

to dla $\forall t_1, t_2$ ograniczenie indeksu sprawiedliwości alokacji zasobów β jest równe maksymalnej różnicy sumy czasów obsługi pakietów o maksymalnych długościach w przepływach współdzielących zasoby:

$$\max \left| \frac{s_i(t_1, t_2)}{C_i} - \frac{s_j(t_1, t_2)}{C_j} \right| =$$

$$= \left| \frac{s_i(t_1, t_2) \pm l_{i \max}}{C_i} - \frac{s_j(t_1, t_2) \pm l_{j \max}}{C_j} \right| \leq \leq \frac{l_{i \max}}{C_i} + \frac{l_{j \max}}{C_j} = \beta.$$

5.2.3. Opóźnienia wnoszone przez dyscyplinę szeregowania.

Każdą dyscyplinę szeregowania, poza sprawiedliwością, charakteryzują gwarantowane wartości górnych ograniczeń opóźnień i fluktuacji opóźnień.

Do szacowania gwarantowanego, przez dyscyplinę szeregowania, górnego ograniczenia całkowitego opóźnienia wykorzystywane są następujące definicje:

1. Definicja zgodności ruchu z parametrami algorytmu ciekącego wiadra.

Jeżeli $a_i(t_1, t_2)$ jest ilością ruchu generowanego w ramach i -tego przepływu w przedziale czasu $(t_1, t_2]$, to i -ty przepływ jest definiowany jako zgodny z algorytmem ciekącego wiadra, opisanym parametrami (b_{ZNi}, v_{ZNi}) , dla każdego t_1 i t_2 , gdy spełniony jest warunek:

$$a(t_1, t_2) \leq b_{ZNi} + v_{ZNi}(t_2 - t_1).$$

2. Definicja wykładniczo ograniczonej wybuchowości ruchu.

Wybuchowość źródła ruchu generowanego w ramach i -tego przepływu jest ograniczona wykładniczo przez proces o parametrach $(v_{ZNi}, \alpha_i, \gamma_i)$ wtedy, gdy dla każdego t_1 i t_2 ($t_2 \geq t_1$) oraz dowolnego x spełnione jest:

$$P\{a(t_1, t_2) \leq x + v_{ZNi}(t_2 - t_1)\} \leq \alpha_i \exp(-\gamma_i x).$$

3. Definicja gwarantowanej szybkości obsługi pakietów. Gwarantowana szybkość obsługi pakietów *GRC* (ang. *Guaranteed Rate Clock*) w k -tym węźle dla każdego i -tego przepływu jest wartością, charakteryzującą każdą jednostkę danych przepływu, spełniającą warunki:

$$GRC_k(i, 0) = 0,$$

$$GRC_k(i, r) = \max\{t_k(i, r), GRC_k(i, r-1)\} + \frac{l_{ir}}{C_i},$$

gdzie: $t_k(i, r)$ – czas napływu r -tego pakietu i -tego przepływu do k -tego węzła, l_{ir} – długość r -tego pakietu i -tego przepływu a C_i – gwarantowana szybkość dla i -tego przepływu.

4. Definicja przynależności dyscypliny obsługi do klasy gwarantowanej szybkości obsługi pakietów.

Dyscyplina szeregowania w k -tym węźle należy do klasy gwarantowanej szybkości obsługi pakietów i -tego przepływu, gdy dana dyscyplina szeregowania gwarantuje, że r -ty pakiet i -tego przepływu będzie obsługowany w k -tym węźle do chwili upływu czasu:

$$t_{kir} = GRC_k(i, r) + \omega_k,$$

gdzie ω_k jest wartością stałą, charakteryzującą implementację danej dyscypliny szeregowania w k -tym węźle (np. maksymalny czas oczekiwania danego pakietu na zakończenie obsługi pakietów, poprzedzających dany pakiet w k -tym węźle trasy).

Zgodnie z podanymi definicjami, jeżeli dyscyplina szeregowania należy do klasy gwarantowanej szybkości obsługi pakietów, to gwarantuje ona, że całkowite opóźnienie r -tego pakietu i -tego przepływu d_{ir} , przekazywanego przez m kolejnych węzłów sieci, spełnia ograniczenia:

- dla ruchu ograniczonego wartościami parametrów algorytmu ciekącego wiadra:

$$d_{ir} \leq \frac{b_{ZNi} + (m-1) \max_{s \in \{1, 2, \dots, r\}} l_{is}}{v_{ZNi}} + \sum_{k=1}^m \omega_k + \sum_{k=1}^{m-1} \tau_k,$$

- dla przepływu o wykładniczo ograniczonej wybuchowości ruchu:

$$P \left\{ d_{ir} \geq x + (m-1) \frac{\max_{s \in \{1, 2, \dots, r\}} l_{is}}{v_{ZNi}} + \sum_{k=1}^m \omega_k + \sum_{k=1}^{m-1} \tau_k \right\} \leq \alpha_i \exp(-\gamma_i x v_{ZNi})$$

gdzie: $(m-1) l_{is}/v_{ZNi}$ – suma opóźnień paketyzacji wynikających z konieczności skompletowania pakietu przed rozpoczęciem kolejnej fazy jego obsługi w każdym z $m-1$ węzłach, tzn. wszystkich węzłach trasy z wyłączeniem węzła ujścia a τ_k ($k = 1, 2, \dots, m$) – czas propagacji w łączu pomiędzy węzłami k -tym a $k+1$.

6. PODSUMOWANIE

Sieć komputerowa z komutacją pakietów jest zbiorem zasobów współdzielonych przez strumienie ruchu generowane przez użytkowników sieci. Ograniczona ilość zasobów powoduje konflikty mogące prowadzić do degradacji jakości dostarczanych usług sieciowych. Konieczność dostarczania jakości usług oznacza, że nie jest możliwe przyjmowanie ruchu do obsługi w sieci bez sterowania, którego zadaniem jest regulowanie ilości ruchu przyjmowanego do obsługi i koordynacja przepływów w sieci. Zbiór mechanizmów realizujących te funkcje to procedury sterowania przepływem.

Metody sterowania implementowane w sieciach pakietowych można dzielić na lokalne i globalne. Sterowanie lokalne jest stosowane w elementach sieci i wykorzystuje dane o ruchu i wykorzystaniu zasobów w tych elementach sieci i ich najbliższych sąsiadach. Ze względu na ograniczenia w dostępie do danych o stanie innych elementów sieci, procedury sterowania lokalnego nie przeciwdziałają powstawaniu przeciążeń; niezbędne jest stosowanie metod sterowania globalnego.

W przypadku realizacji koncepcji usług zintegrowanych opartej na udostępnianiu wcześniej zarezerwowanych zasobów sterowanie globalne nie jest wymagane; potrzeba sterowania globalnego jest eliminowana przez rezerwację zasobów. W przypadku takich sieci zadania sterowania są realizowane lokalnie i obejmują: sterowanie dostępem, sterowanie ruchem (kształtowanie ruchu) i szeregowanie pakietów..

Od mechanizmów sterowania implementowanych w węzłach sieci pakietowych wymagane jest prostota implementacji i realizacji, odporność na zmiany ilościowe ruchu oraz sprawiedliwość dostępu ruchu gene-

rowanego przez użytkowników sieci do ograniczonych i współdzielonych zasobów sieci.

W pracy przedstawiono wybrane mechanizmy sterowania stosowane w zadaniach kształtowania ruchu i szeregowania pakietów w sieciach z gwarancją jakości usług, tj. w sieciach, w których implementowana jest koncepcja usług zintegrowanych.

CONTROL OF TRAFFIC IN PACKET-SWITCHED NETWORKS TO ASSURE QUALITY OF SERVICE

Abstract: Providing different levels of service in packet-switched network introduces a number of mechanisms that are mainly implemented in two layers: control path and data path. Quality of service in general means that particular traffic flows or classes of traffic are differentiated in sense of access to limited resources and mechanisms through which different services are enforced. Delivery of quality of service in packet-switched networks requires new specification, classification, scheduling and congestion control mechanisms. The mechanism consists of buffer management schemes (decide which packets can be stored as they wait for transmission) and scheduling algorithms (control the actual transmission of packets). The gain of the paper is to present some methods implemented in network nodes to offer different concepts of quality of service.

Literatura

- [1] Grzech A. (2002) Sterowanie ruchem w sieciach teleinformatycznych, Oficyna PWR, Wrocław.
- [2] Guerin R.A., Orda A. (1999) QoS routing in networks with inaccurate information: theory and algorithms, IEEE/ACM Transactions on Networking, 7, 350–364.
- [3] Guerin R.A., Peris V. (1999) Quality-of-Service in packet networks basic mechanisms and directions, Computer Networks, 31, 169–179.
- [4] Hallot C.V., Misra V., Towsley D., Gong Wei-Bo (2001), On designing improved controllers for AQM routers supporting TCP flows, IEEE INFOCOM.
- [5] Papier Z. (2001) Ruch telekomunikacyjny i przeciążenia sieci pakietowych, WKiŁ, Warszawa.
- [6] Perekh A.K., Gallager R.G. (1994), A GPS approach to flow control in integrated services networks: the multiple node case, IEEE/ACM Transactions on Networking, 2, 137–150.
- [7] Wang Z. (2001) Internet QoS: Architectures and mechanisms for Quality of Service, Academic Press, London.



Instytut Badań Systemowych
Polskiej Akademii Nauk

ISBN 83-89475-01-4