

STEROWANIE RUCHEM W SIECIACH TELEKOMUNIKACYJNYCH

*Sterowanie przepływem w sieci
telekomunikacyjnej*

Istota sterowania przepływem

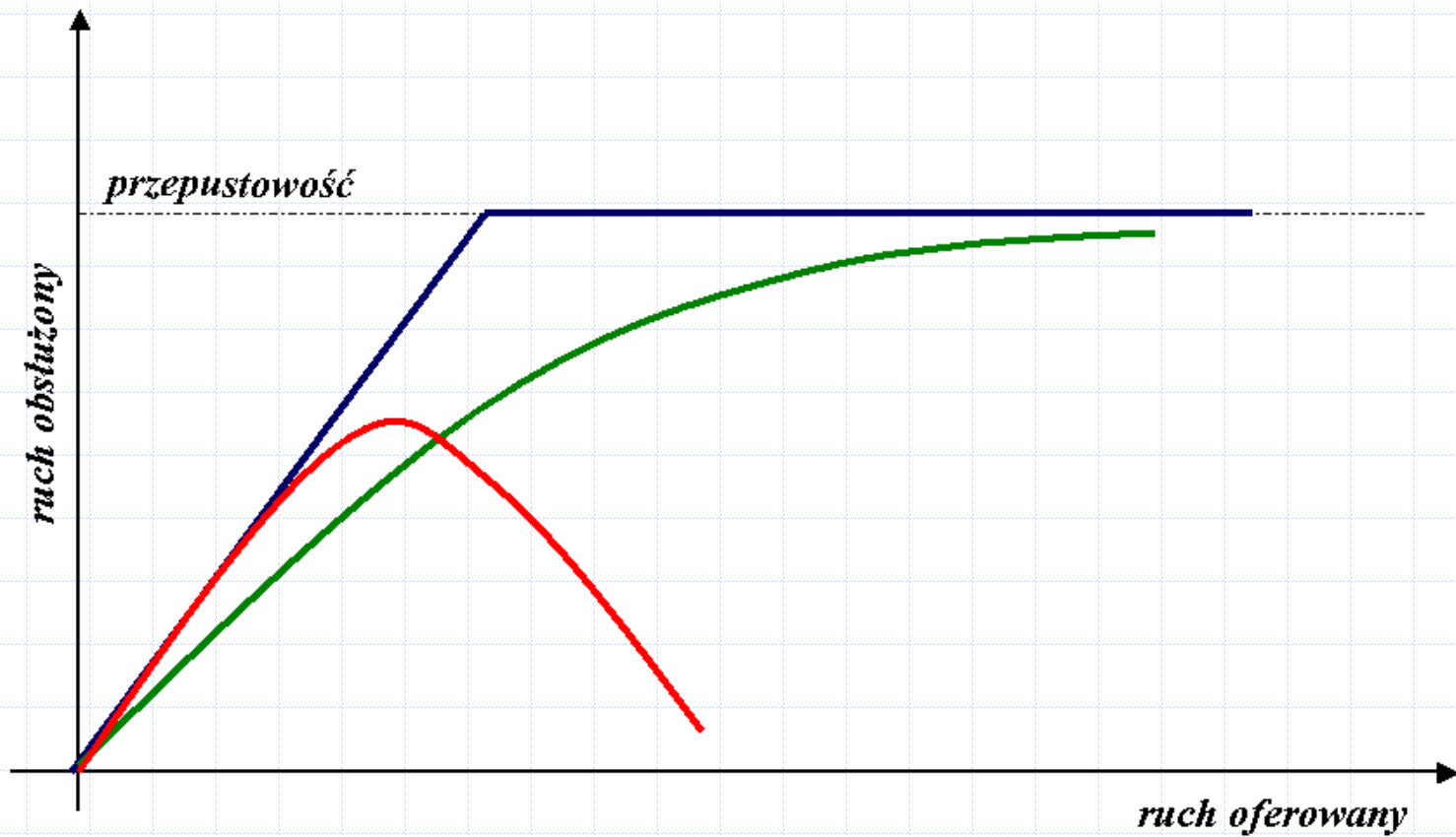
Cel sterowania

Zabezpieczenie systemu przed konsekwencjami przeciążenia zasobów systemu lub braku dopasowania szybkości transferu (szybkości wprowadzani i szybkości wyprowadzania) w wybranych elementach systemu (podsystemach)

Warunki efektywnego sterowania

- ☛ wymiana informacji pomiędzy elementami systemu (węzłami) o stanie systemu i jego elementów niezbędnej do wyboru sterowania i/lub parametrów sterowania*
- ☛ wymiana komend i odpowiedzi (zawierających wartości parametrów procedur sterowania i uzgodnień) niezbędnych do implementacji ustalonej metody sterowania*

Sterowanie przepływem w sieci



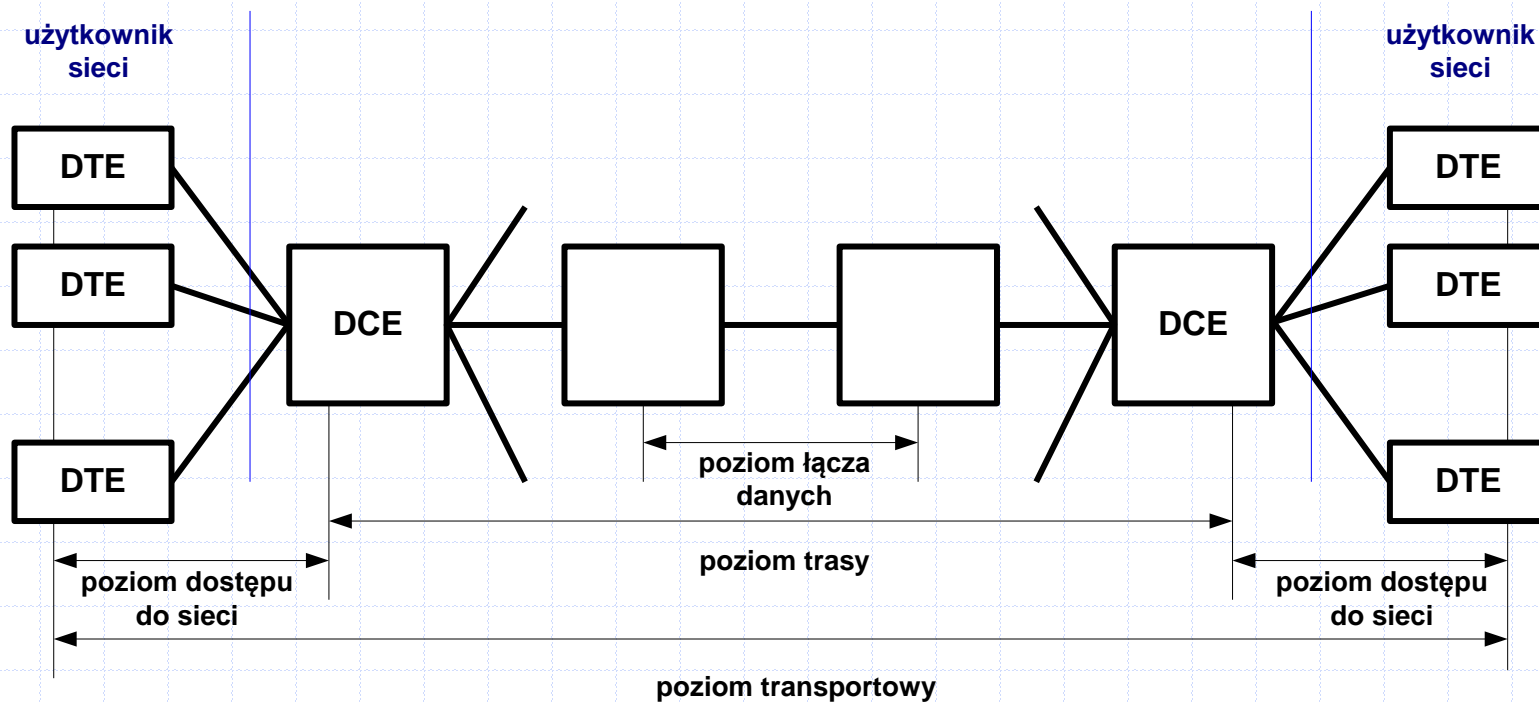
Sterowanie przepływem w sieci

- *Dopasowanie szybkości (speed matching) transferu pomiędzy siecią a procesami korzystającymi z usług sieci*
- *Sprawiedliwa alokacja zasobów (fair resource allocation) współdzielonych pomiędzy współzawodniczące procesy*
- *Zapobieganie degradacji przepustowości (throughput degradation) i utracie efektywności z powodu przeciążenia (overload)*
- *Unikanie przeciążeń/zakleszczeń (deadlock avoidance)*

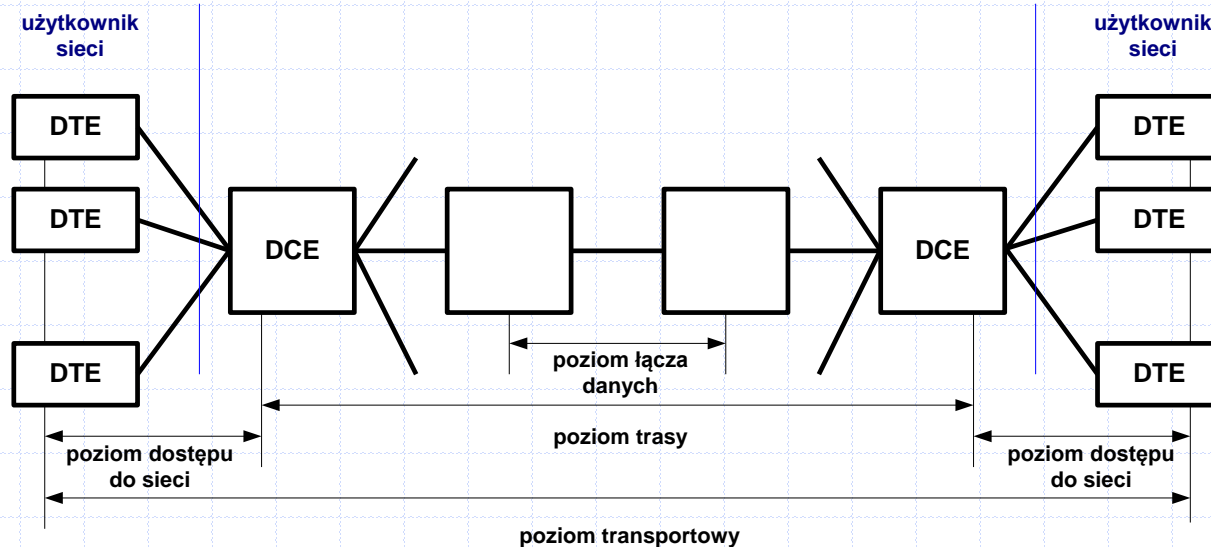
Przeciążenie

- *Stan długotrwałego przeładowania sieci, kiedy zapotrzebowanie na ograniczone jej zasoby jest zbliżone do przepustowości sieci lub ją przekracza*

Architektura systemu sterowania przepływem



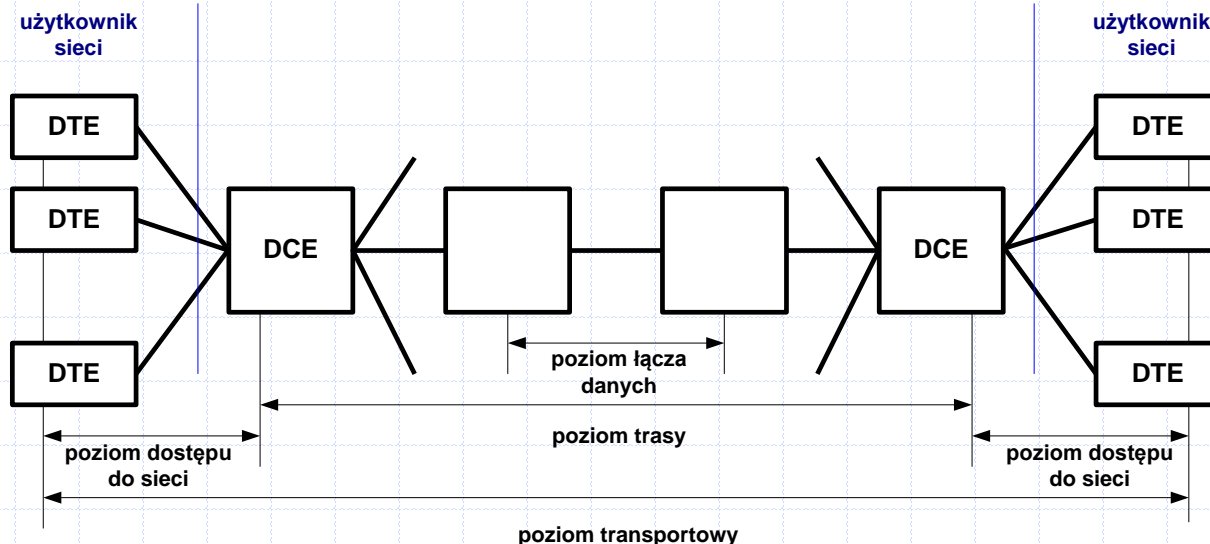
Zakres sterowania



Sterowanie na poziomie transportowym

- ➔ *dotyczy wszystkich elementarnych strumieni ruchu*
- ➔ *pozwala dopasować szybkość nadawania do szybkości odbioru w procesie komunikacji źródło - ujście*
- ➔ *możliwość zapobiegania przeciążeniu globalnemu*

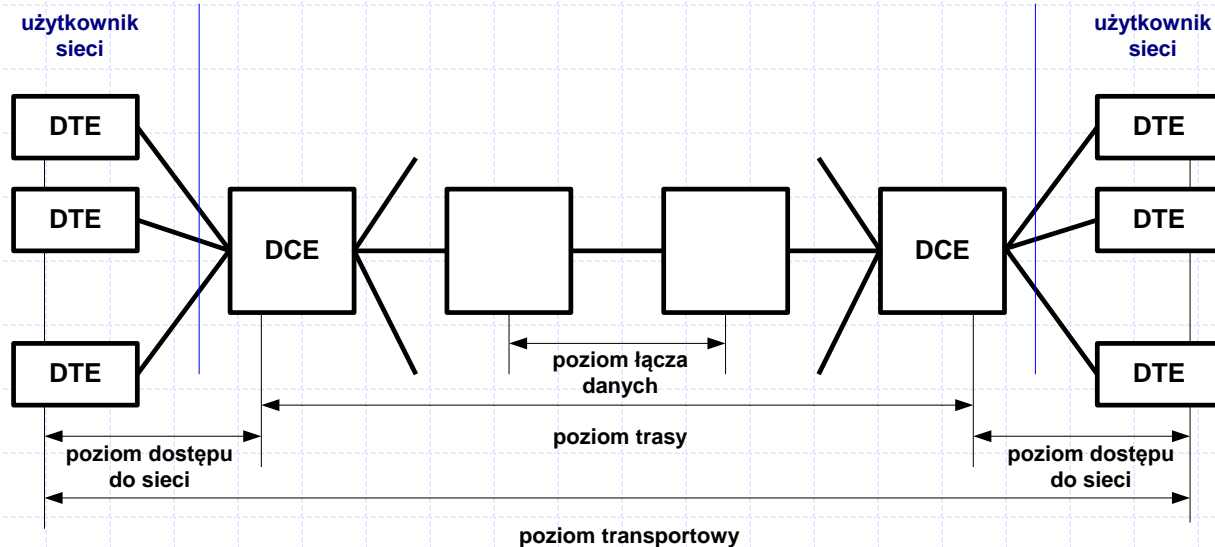
Zakres sterowania



Sterowanie na poziomie trasy

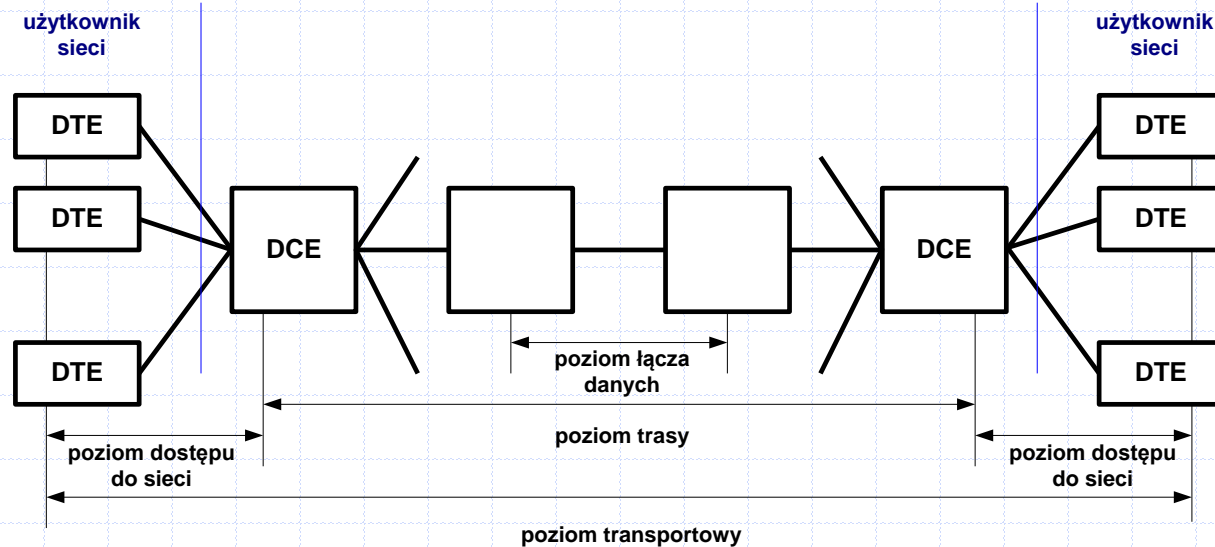
- ➔ *dotyczy przepływu między dwoma wyróżnionymi węzłami końcowymi sieci*
- ➔ *pozwała ograniczyć napływ pakietów do pamięci buforowej węzła docelowego (ujścia)*
- ➔ *możliwość zapobiegania przeciążeniom lokalnym w węzłach ujściowych i przeciążeniom tras*

Zakres sterowania



- ✓ **Sterowanie na poziomie łącza danych**
 - ➔ dotyczy przepływu między dwoma sąsiednimi węzłami sieci
 - ➔ pozwala ograniczyć napływ pakietów do pamięci buforowej węzła odbiorczego
 - ➔ możliwość zapobiegania przeciążeniom lokalnym

Zakres sterowania

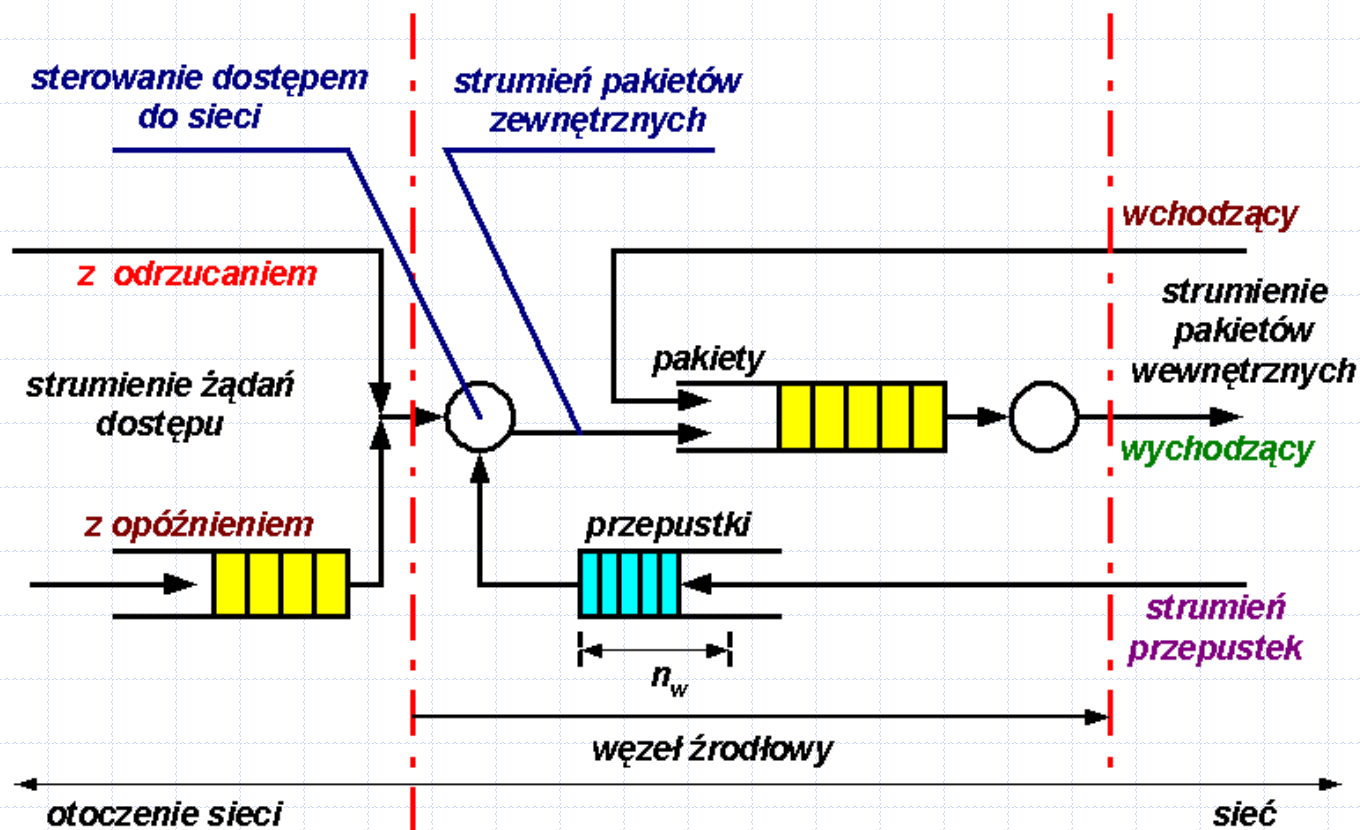


- ✓ **Sterowanie na poziomie dostępu do sieci**
 - ➔ dotyczy wszystkich przepływów w sieci
 - ➔ ingeruje w ilość i szybkość wprowadzania nowego ruchu
 - ➔ możliwość zapobiegania przeciążeniu globalnemu

Metody sterowania na poziomie dostępu do sieci

- ✓ **Metoda izarytmiczna (przepustek)**
 - ograniczanie liczby pakietów przez stosowanie ustalonej liczby przepustek (żetonów)
- ✓ **Metoda bufora wejściowego**
 - ograniczanie pojemności pamięci buforowych dostępnych dla ruchu wejściowego
 - ruch tranzytowy nie ma ograniczeń w dostępie do pamięci buforowych
- ✓ **Metoda pakietów dławiących**
 - generowanie w obszarach sieci zagrożonych przeciążeniem specjalnych pakietów powodujących ograniczenie intensywności napływu nowych pakietów
- ✓ **Metoda stopniowych ograniczeń**
 - wprowadzanie progowych wartości intensywności napływu nowych pakietów w zależności od długości kolejek
- ✓ **Metoda okna**
 - ograniczenie liczby połączeń wirtualnych i liczby pakietów w każdym połączeniu

Metoda izarytmiczna



Ograniczenia metody przepustek

- *zależność jakości obsługi pakietów od rozkładu generowanego ruchu*
- *zależność efektywności wykorzystania zasobów od przyjętego sposobu dystrybucji przepustek zwalnianych w węzłach ujściowych:*
 - ✓ *równomierny rozdział zwalnianych przepustek pomiędzy wszystkie węzły sieci*
 - ✓ *przekazywanie przepustek do węzła źródłowego*
- *konieczność implementacji systemu zarządzania przepustkami*

Dobór parametrów metody przepustek

➤ Parametry metody przepustek

- ✓ maksymalna całkowita liczba przepustek w sieci: n_p .
- ✓ maksymalna liczba przepustek oczekujących oczekujących w każdym z węzłów sieci: n_w
- ✓ schemat dystrybucji w sieci zwalnianych przepustek

➤ Krytyczne wartości liczby przepustek

$$n_p = 3N$$

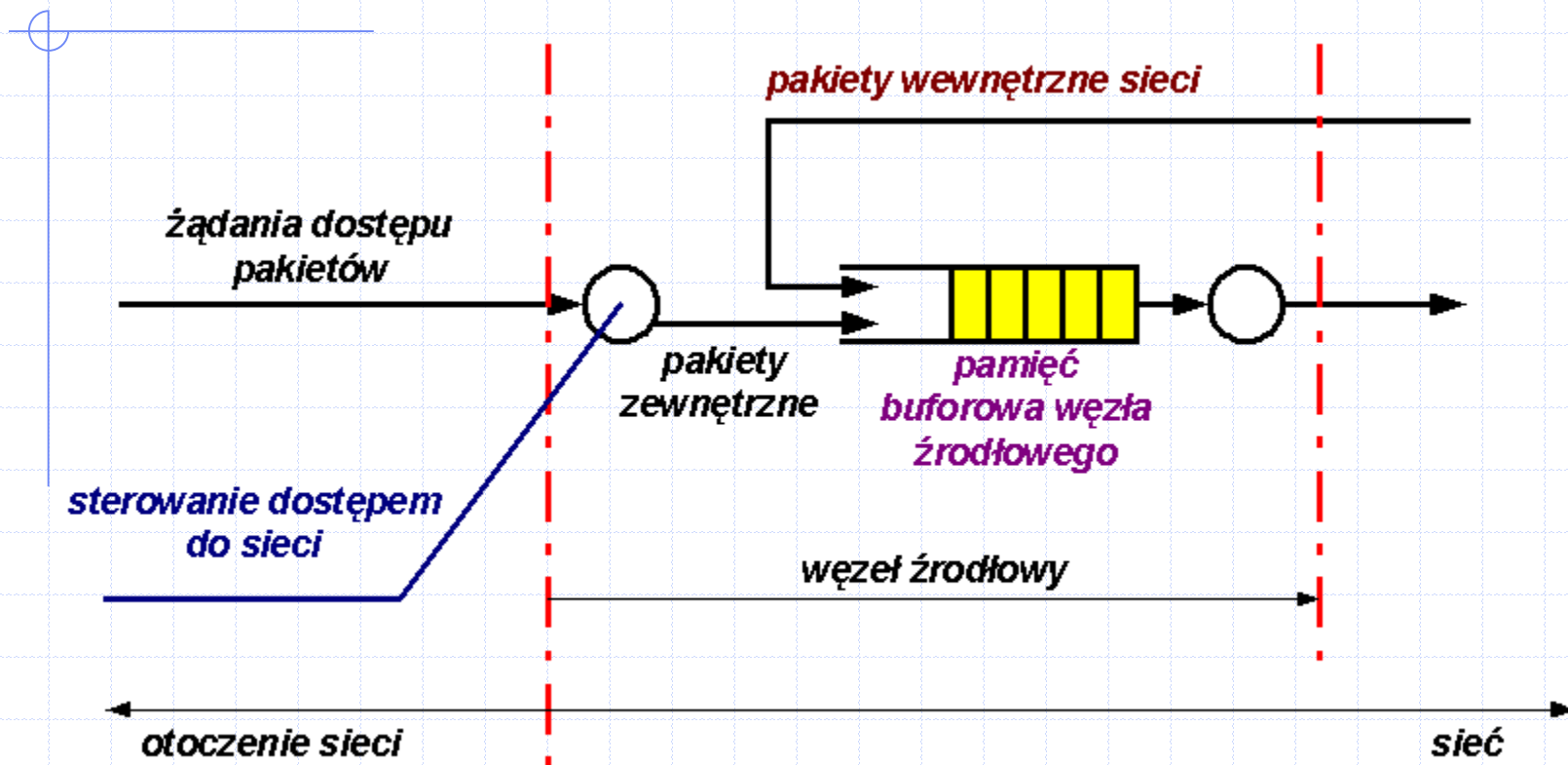
$$n_w = 3$$

N - liczba wszystkich węzłów sieci

Metoda bufora wejściowego

- *procedura lokalnego sterowania dostępem do sieci*
- *założenia metody*
 - ✓ *istnieje jednoznaczna zależność pomiędzy przeciążeniem węzłów źródłowych sieci i przeciążeniem globalnym sieci*
 - ✓ *zapobieganie przeciążeniom węzłów źródłowych sieci (na brzegu sieci) jest sposobem zapobiegania przeciążeniu globalnemu sieci*
 - ✓ *priorytet obsługi pakietów, które już wykorzystywały zasoby sieci powinien być wyższy od priorytetu obsługi nowych pakietów*

Istota metody bufora wejściowego



Implementacja metody bufora wejściowego

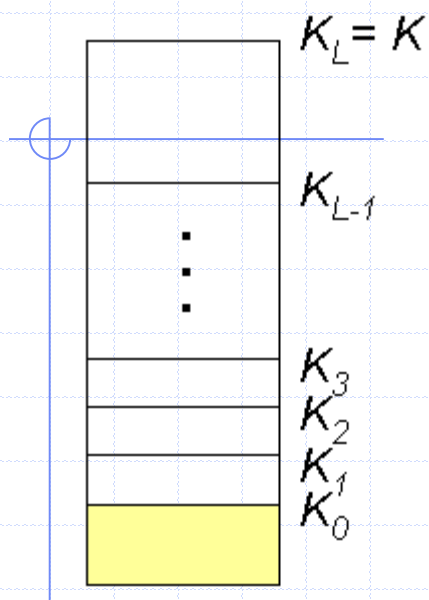
➤ Założenia

- ➔ napływające do bufora pakiety wewnętrzne należą do jednej z klas Cl_i , gdzie: $i = 1, 2, \dots, L$ – liczba węzłów, w których był obsługiwany pakiet wewnętrzny
- ➔ pakiety zewnętrzne napływające do danego węzła źródłowego zalicza się do klasy Cl_0 i zajmują bufory należące do zbioru K_0
- ➔ zasady akceptacji żądań dostępu do sieci w węźle źródłowym są determinowane zasadami dostępu do pamięci buforowej pakietów należących do różnych klas

➤ Procedury akceptacji żądań dostępu

- ✓ procedura GMDNET
- ✓ procedura Lama
- ✓ procedura Price'a
- ✓ procedura Kamouna (DTFC)

Procedura GMDNET



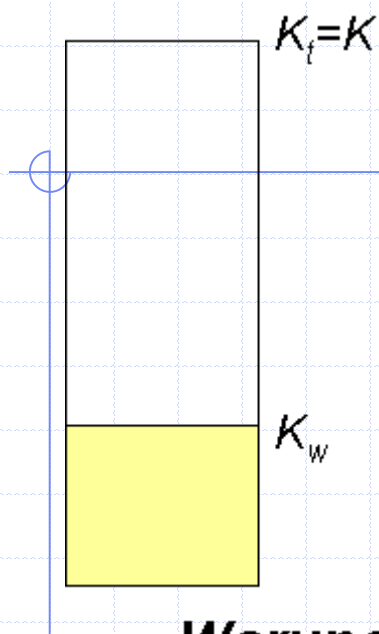
Warunki akceptacji żądań dostępu

- ✓ bufory należące do zbioru K_0 nie są zajęte przez pakiety należące do klas wyższych
- ✓ w zbiorze buforów K_0 jest co najmniej jeden wolny bufor, tzn., że jest spełniony warunek

$$k_0 < \min \left(K_0, K - \sum_{i=1}^L k_i \right) \quad 0 \leq k_i \leq K_i \quad K_0 \leq K_1 \leq \dots \leq K_L = K$$

gdzie: k_i jest liczbą pakietów i -tej klasy w pamięci buforowej węzła źródłowego

Procedura Lama



Założenia

- *dwie klasy pakietów*
 - ✓ wejściowe
 - ✓ tranzytowe
- *pakiety tranzytowe mogą zajmować wszystkie bufory pamięci buforowej węzła*

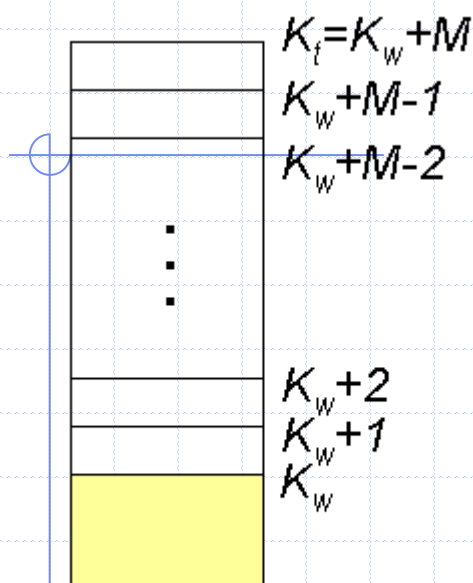
Warunek akceptacji żądań dostępu

$$k_w < \min(K_w, K - k_t) \quad 0 \leq k_t \leq K_t = K$$

gdzie: k_w – liczba buforów pamięci zajętych przez pakiety wejściowe

k_t – liczba buforów pamięci zajętych przez pakiety tranzytowe

Procedura Price'a



Założenia

- dwie klasy pakietów: wejściowe i tranzytowe
- liczba buforów dostępnych w pamięci buforowej węzła źródłowego dla pakietów wejściowych spełnia warunek:

$$K_w = K - M$$

gdzie: K – całkowita liczba buforów

M – liczba kanałów wyjściowych (kolejek)

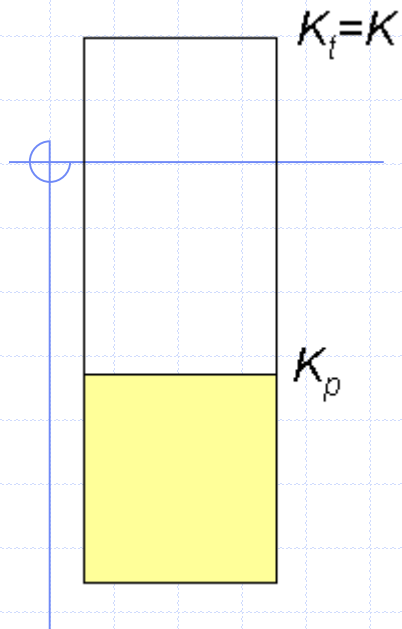
Warunek akceptacji żądań dostępu

$$k_w < \min \left(K_w, K - \sum_{m=1}^M k_{mt} \right) \quad 0 \leq k_{mt} \leq K_{mt} = K - M + 1$$

gdzie: k_w, k_t – liczby pakietów wejściowych i tranzytowych w pamięci buforowej

k_{mt} – liczba pakietów tranzytowych w m -tej kolejce wyjściowej węzła źródłowego

Procedura Kamouna



Założenia

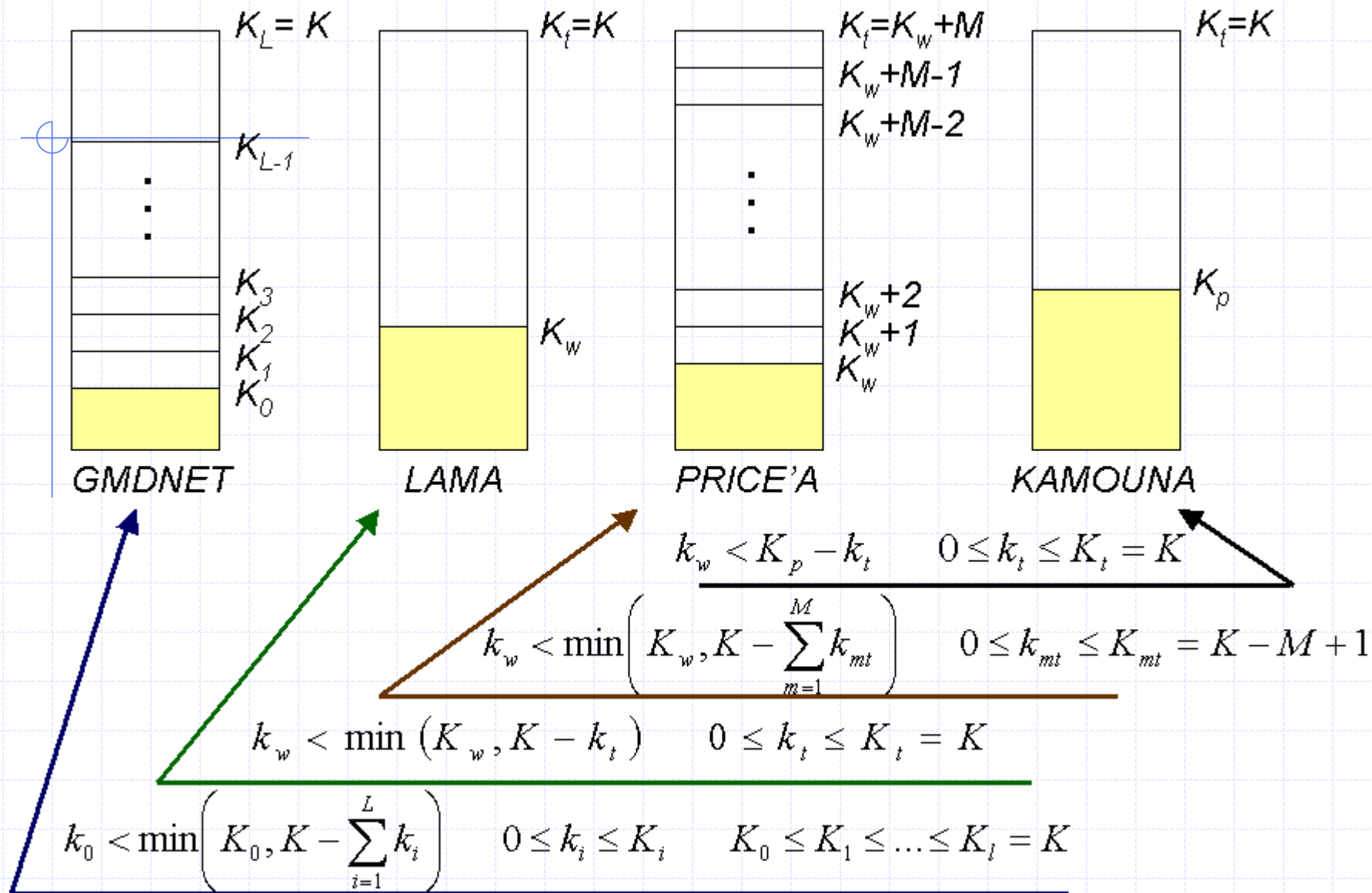
- *dwie klasy pakietów: wejściowe i tranzytowe*
- *zadana wartość progowa liczby pakietów w pamięci buforowej węzła źródłowego K_p*

Warunek akceptacji żądań dostępu

$$k_w < K_p - k_t \quad 0 \leq k_t \leq K_t = K$$

gdzie: k_w, k_t – liczby pakietów wejściowych i tranzytowych w pamięci buforowej

Metoda bufora wejściowego



Metoda pakietu dławiącego

Kryterium przeciążenia trasy

„trasa jest przeciążona, jeżeli co najmniej jedno łącze wchodzące w skład trasy jest przeciążone”

Kryterium przeciążenia łącza

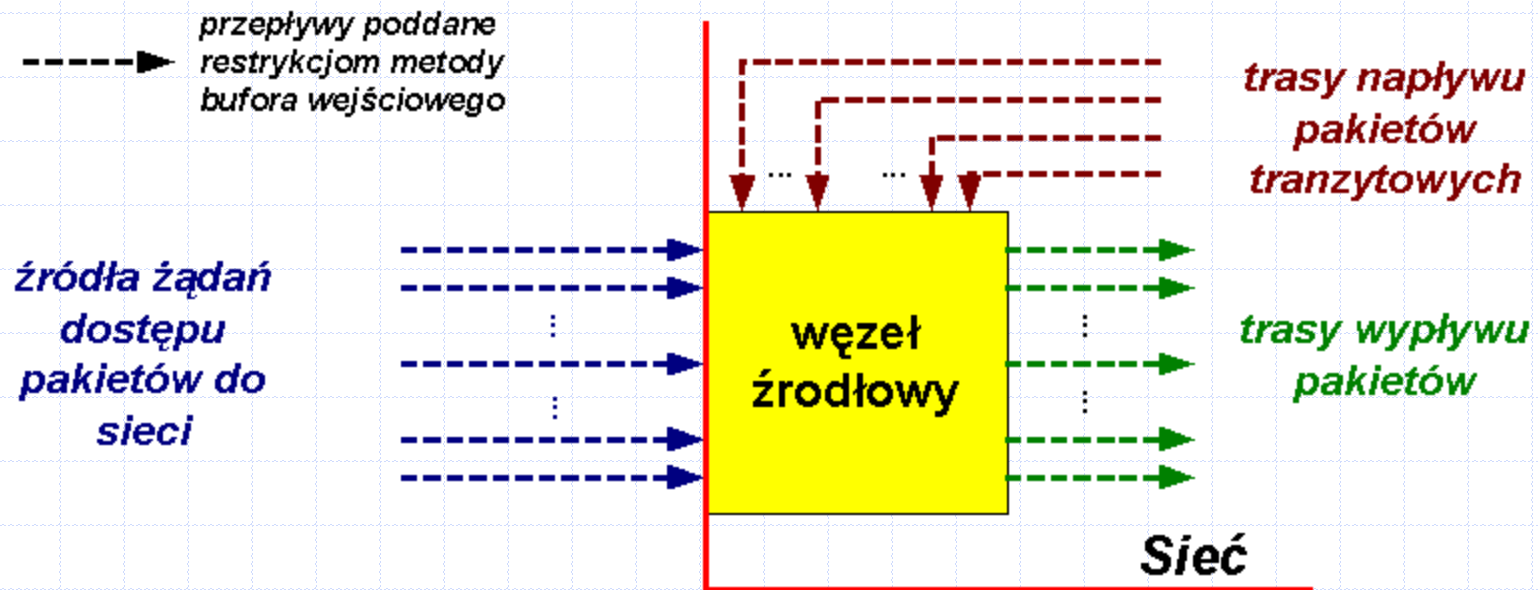
„łącze jest przeciążone, gdy stopień jego wykorzystania jest wyższy od ustalonej wartości progowej”

Warunki akceptacji żądania dostępu

1. Jeżeli do węzła napływa żądanie dostępu pakietu do sieci kierowane do węzła, do którego prowadzi trasa spełniająca kryterium przeciążenia, to nie jest ono akceptowane
2. Jeżeli napływający do węzła pakiet jest pakietem tranzytowym kierowanym do ujścia, do którego prowadzi trasa spełniająca kryterium przeciążenia, to:
 - ✓ pakiet jest wysyłany przeciążoną trasą,
 - ✓ w danym źródle jest generowany pakiet dławiący (choke packet), którego adresatem jest węzeł źródłowy odebranego wcześniej pakietu tranzytowego. Przenosi on informację o przeciążeniu i prowadzi do zablokowania przeciążonej trasy

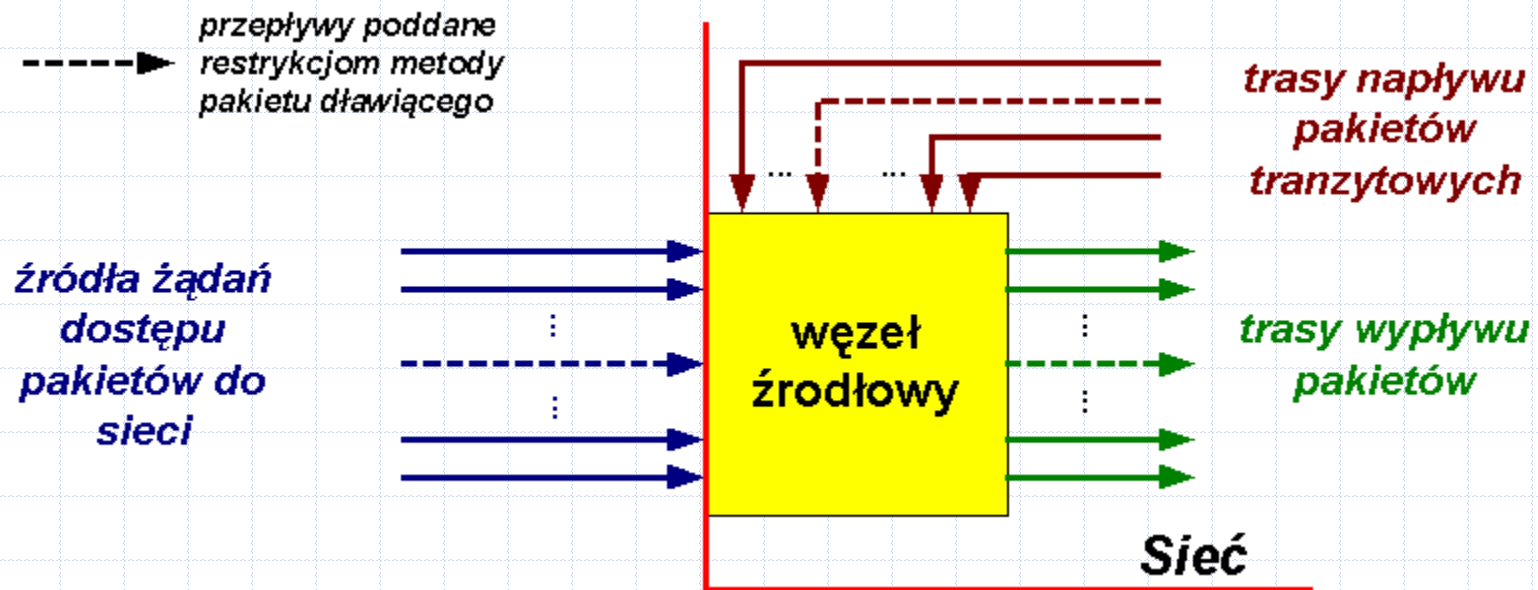
Całościowe i selektywne sterowanie przepływem

Metoda bufora wejściowego – sterowanie całościowe



Całościowe i selektywne sterowanie przepływem

Metoda pakietu dławiącego – sterowanie selektywne



STEROWANIE RUCHEM W SIECIACH TELEKOMUNIKACYJNYCH

Alokacja zasobów

Mechanizmy sterowania w sieci telekomunikacyjnej

Mechanizmy inżynierii ruchowej

Alokacja zasobów
Wymiarowanie sieci
Wyznaczanie przebiegu dróg w sieci



Strategie taryfikacji

Zachęta do racjonalnego korzystania z sieci



Zarządzanie siecią

Nadzór operatora nad działaniem sieci

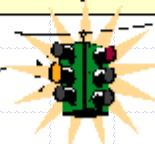


Sygnalizacja

Ustanawia połączenie i powinna pozwolić terminalowi przesłać do sieci deskryptor ruchu i wymagania QoS

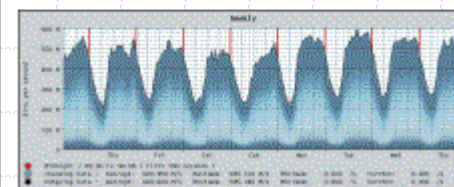
Sterowanie przyjmowaniem nowych wywołań

Dopuszcza nowe zgłoszenie jeśli sieć będzie w stanie przesłać ruch z założonym poziomem QoS, nie degradując ruchu już obsługiwanego



Pomiary i monitorowanie sieci

Wsparcie niektórych mechanizmów sieciowych (np. sterowania przyjmowaniem wywołań i inżynierii ruchowej) poprzez dostarczenie aktualnej wiedzy o stanie sieci

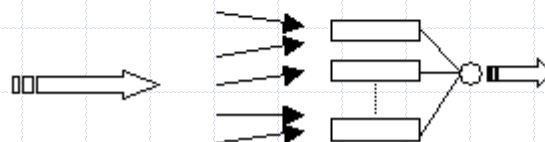
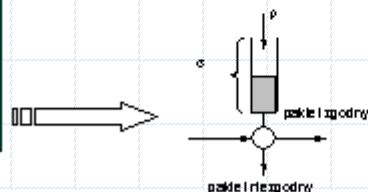


Mechanizm monitorowania zgodności

Monitoruje ruch w ramach danego połączenia i sprawdza jego zgodność z profilem opisanym w kontrakcie ruchowym

Mechanizmy szeregowania pakietów

Regulują dostęp różnych strumieni ruchu do wspólnych zasobów transmisyjnych łączy



Sterowanie



Przekaz danych

Skala czasowa

wg. W. Burakowski

Alokacja zasobów

- Sieć telekomunikacyjna może być traktowana jako rozproszony zbiór zasobów współdzielonych pomiędzy współzawodniczące o nie procesy komunikacyjne
 - dopuszczalność sytuacji, w której chwilowe zapotrzebowanie jest większe od przepustowości sieci
 - dynamiczny przydział zasobów dla komunikujących się procesów (efektywne wykorzystanie zasobów)

Przydział zasobów

- *Równomierny przydział zasobów*
- *Sprawiedliwy przydział zasobów*
 - ✓ *technika „max-min fair share”*
 - ✓ *technika „max-min share allocation”*

Technika „max-min fair share”

- ✓ przydział zasobów wg żądanej przepływności (od najmniejszej do największej)
- ✓ żadne źródło nie uzyskuje więcej przepływności niż żądana
- ✓ źródła, których żądania nie zostały spełnione, dzielą wolną przepływność między siebie po równo

Technika „max-min fair share”

- liczba strumieni: n
- żądania zasobów: $x_1 \leq x_2 \leq x_3 \dots \leq x_n$
- przepustowość łącza: C
- zasób przydzielany strumieniowi o najmniejszym żądaniu: C/n
- jeżeli przydzielony zasób jest większy niż żądanie, wówczas zasób $(C/n - x_1)$ jest dostępny dla pozostałych strumieni
- niewykorzystana przepustowość jest dzielona równo pomiędzy pozostałe strumienie: $C/n + (C/n - x_1)/(n - 1)$
- jeżeli przydzielony zasób jest większy niż żądanie strumienia 2, wówczas niewykorzystany zasób jest dostępny dla pozostałych strumieni
- proces przydziału zasobów kończy się, kiedy każdy ze strumieni nie otrzymuje więcej niż żądał (jeżeli żądanie danego strumienia nie jest w pełni zaspokojone, to mimo wszystko nie otrzymuje on mniej, niż strumień o wyższym żądaniu)

Maksymalizacja minimum przydziału dla strumieni, które nie otrzymały żądanych zasobów

Technika „max-min fair share”

Przykład:

Oblicz przydział max-min dla zbioru 4 strumieni z żądaniami odpowiednio: 2; 4; 6; 8 gdy przepustowość łącza wynosi $C = 16$

Rozwiązanie:

Krok 1: $16/4 = 4 \rightarrow x_1 = 2 \rightarrow$ wolna przepustowość dla pozostałych strumieni: $4 + 2/3 = 4,6666$

Krok 2: $x_2 = 4 \rightarrow$ wolna przepustowość dla pozostałych strumieni: $4,6666 + 0,6666/2 = 5$

Krok 3: $x_3 = 5$

Krok 4: $x_4 = 5$

Technika „max-min share allocation”

- ✓ *stosowana w przypadku, gdy pewnym strumieniom chce się przydzielić więcej zasobów niż pozostałym*
- ✓ *wymaga nadania strumieniom odpowiednich wag*
- ✓ *przydzielone wagi mają wpływ na względne wykorzystanie zasobów przez poszczególne strumienie*

Technika „max-min share allocation”

- ✓ *zasoby są przydzielane według kolejności
wzrastających wag*
- ✓ *żaden strumień nie otrzymuje więcej zasobu od
strumienia o wyższym żądaniu*
- ✓ *strumienie z niezaspokojonymi żądaniami mają
dostęp do wolnych zasobów zgodnie z
przydzielonymi wagami*

Technika „max-min share allocation”

Przykład:

- Oblicz przydział max-min dla zbioru 4 strumieni z żądaniami odpowiednio: 2, 4, 6, 8 z wagami: 2,5; 4; 0,5; 1 gdy przepustowość łącza wynosi $C = 16$

Rozwiązanie:

Krok 1: normalizacja wag: 5; 8; 1; 2

Krok 2: liczba strumieni cząstkowych (N) = suma wag $N = 16$

Krok 3: $C/N \rightarrow 16/16 = 1$

strumień 1: $5 \times 1 = 5$

strumień 2: $8 \times 1 = 8$

strumień 3: $1 \times 1 = 1$

strumień 4: $2 \times 1 = 2$

żądał 2 \rightarrow nadmiar 3 jednostki

żądał 4 \rightarrow nadmiar 4 jednostki

żądał 6 \rightarrow niedobór

Żądał 8 \rightarrow niedobór

Technika „max-min share allocation”

Krok 4: podział wolnej przepustowości pomiędzy strumienie 3 i 4 wg. ich wag

strumień 3: $7 \times (1/3) = 2,3333$

strumień 4: $7 \times (2/3) = 4,6666$

Krok 5: podział ostateczny

strumień 1: 2

strumień 2: 4

strumień 3: $1 + 2,3333 = 3,3333$

strumień 4: $2 + 4,6666 = 6,6666$

→ żądanie zaspokojone

→ żądanie zaspokojone

→ żądał 6

→ żądał 8



Dziękuję za uwagę!