

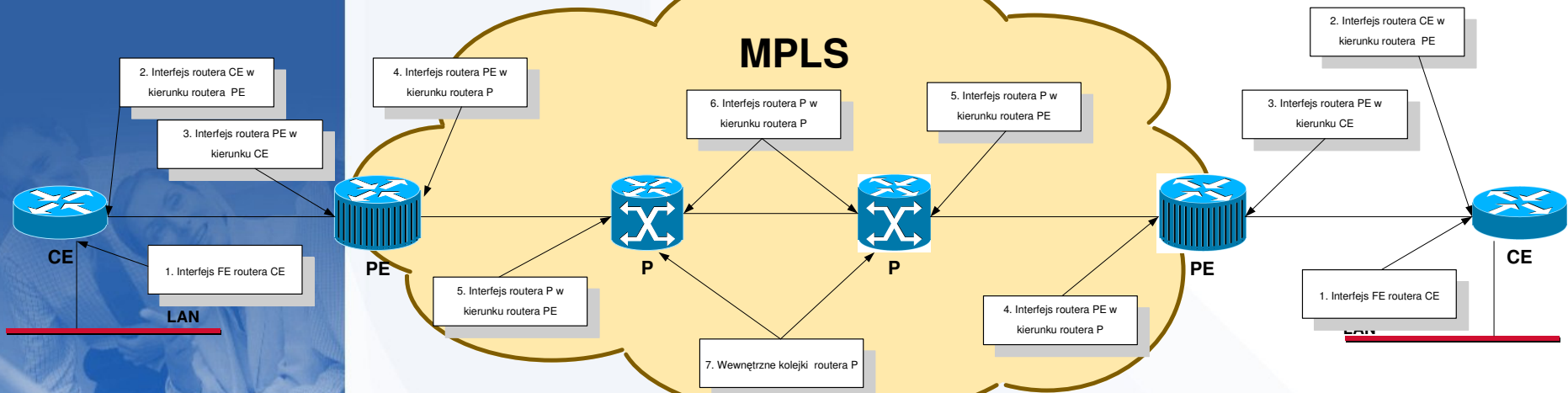
Wdrażanie mechanizmów QoS w sieciach MPLS

Prowadzi: **Przemysław Misiak**

Warszawa 05.02.2008

O czym jest ta sesja?

- Omówienie wdrażania architektury Diffserv w sieci
uwaga: to nie jest **jedyn**e słuszne podejście
- Nie dotyka zagadnień inżynierii ruchowej (MPLS TE)



Agenda

- Co to jest ten Diffserv
- Przykład Modelu Diffserv
- Model QoS klienta (brzegu sieci)
 - Klasy
 - mechanizmy
- Model QoS szkieletu sieci
 - Klasy
 - Mechanizmy
 - Tunelowanie

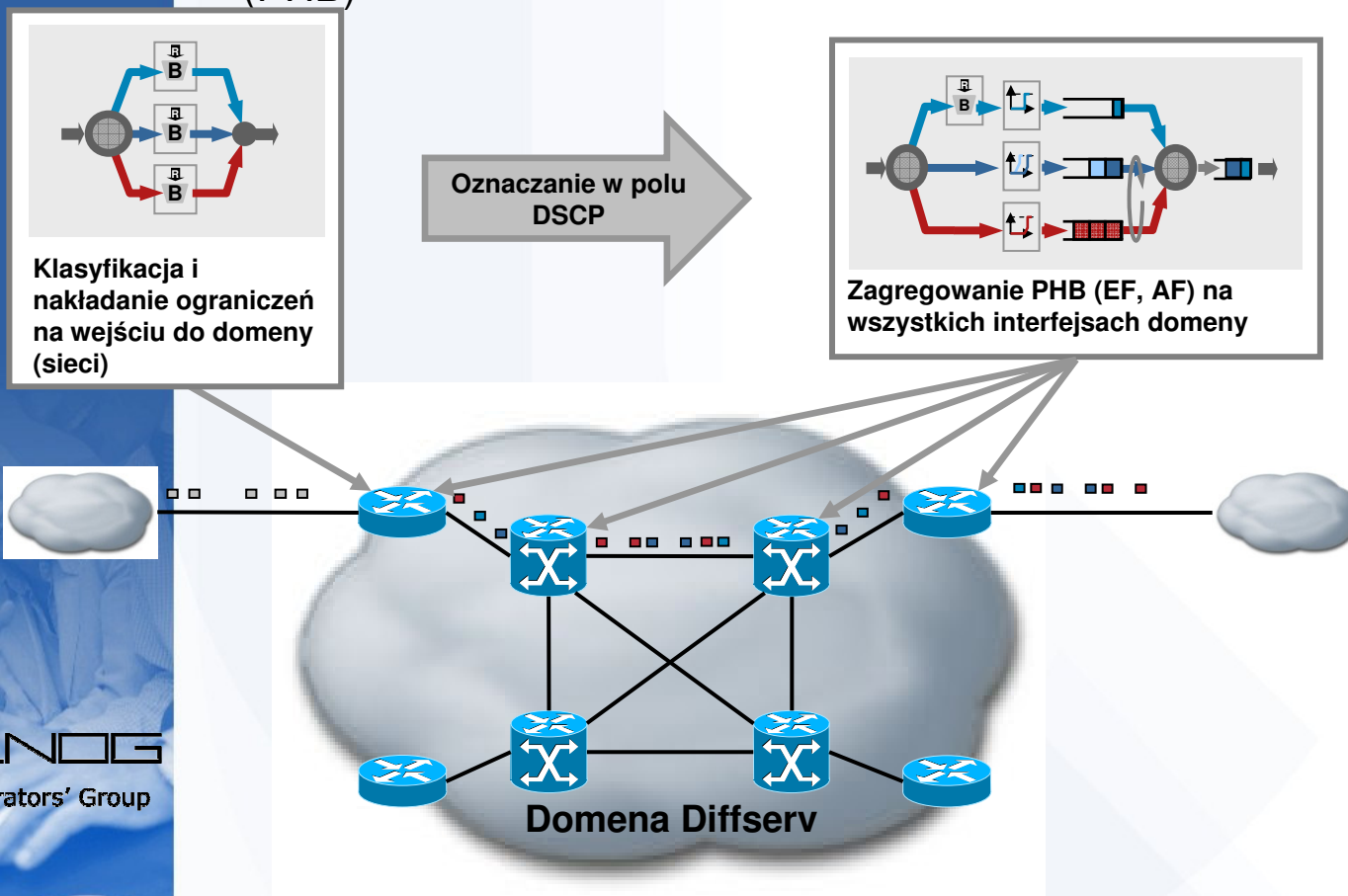
Czym jest Diffserv?

Architektury QoS

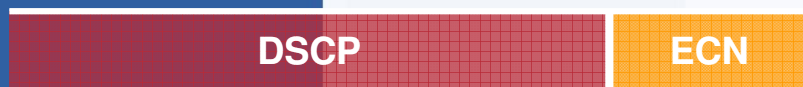
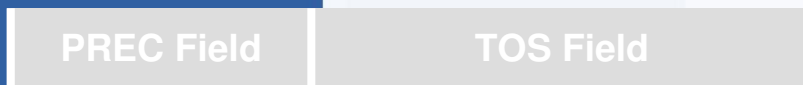
- Architektury QoS opisują sposób w jaki mechanizmy QoS działają w sieciach zapewniając gwarancje end-to-end
- Pierwsze architektury IP QoS nie wspominały wprost o SLA end-to-end
 - IP Precedence (RFC 791), IP ToS (RFC 1349)
- Modele obecnie używane to:
 - Differentiated Services – Diffserv (RFC2475)
 - Integrated Services – Intserv (RFC1633)
 - Modele oparte o MPLS TE
- W sieciach dostawców usług modele mogą współistnieć i uzupełniać się, choć dla „prostych” polityk QoS stosuje się tradycyjnie architekturę Diffserv (mało aplikacji obsługuje RSVP wprost)

Architektura Diffserv – RFC2475

- Mały poziom złożoności – brak utrzymania stanów czy sygnalizacji
- Usługi tworzone są przez kombinację złożonej klasyfikacji, oznaczania, ewentualnie przycinania ruchu na brzegu sieci i konfiguracji szkieletu (PHB)



Pole Diffserv



Wybór klasy (CS)

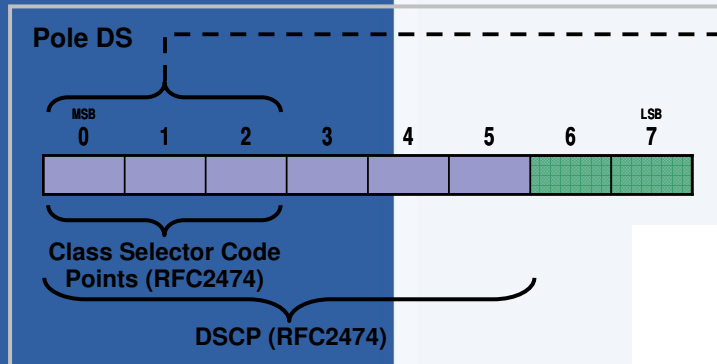
} bajt Type of Service (ToS) (RFC1349)
} pole Diffserv (DS) (RFC 2474)

- Pole DS. zastępuje pola IPv4 ToS/IPv6 Traffic Class
- Pole składa się z 6 bitów
- Pierwsze 3 bity DSCP są selektorem klasy (CS) – funkcjonalnie odpowiednie i wstecznie kompatybilne z IP Precedence
- RFC 3168 definiuje dodatkowo wykorzystanie dwóch bitów z pola ECN (Explicit Congestion Notification)
...nie jest to jednak często wykorzystywane

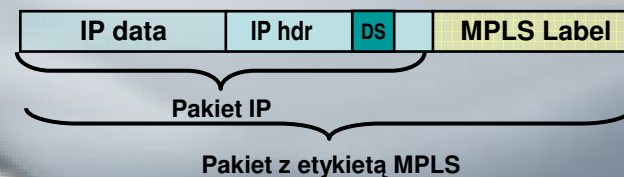
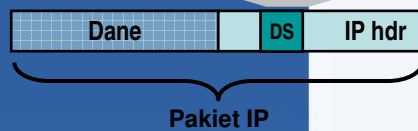
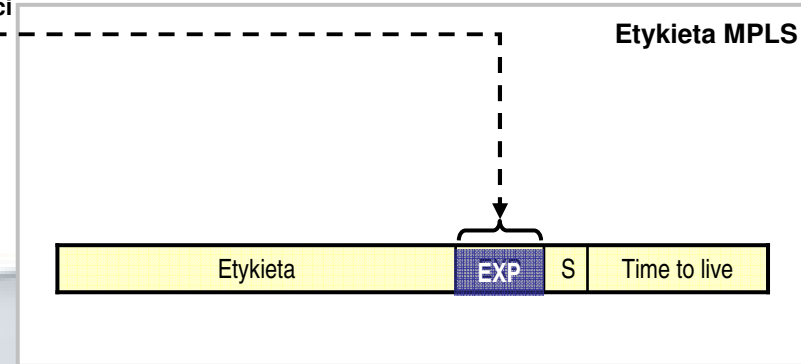
Zarządzanie per-hop (PHB)

- Mechanizmy kształtowania PHB w ramach modelu Diffserv rozróżniają klasy ruchowe w momencie wystąpienia przeciążenia sieci:
 - Expedited Forwarding (EF) PHB – RFC3246
 - zwykle implementowana jako kolejka priorytetowa (zwykle głos lub inny, wysokopriorytetowy ruch)
 - Assured Forwarding (AF) – RFC2597
 - n kolejek dla klas \times m „uprzywilejowania” w odrzucaniu
 - zwykle implementowana jako algorytm kolejek ważonych dla klas danych
 - Domyślna
 - Brak gwarancji
- Każdy węzeł sieciowy może samodzielnie decyduje o losie każdego pakietu zgodnie ze swoją, lokalną polityką QoS – i może rozróżnić wiele klas ruchowych dzięki takiej klasyfikacji

MPLS i Diffserv



Kopiowanie wartości



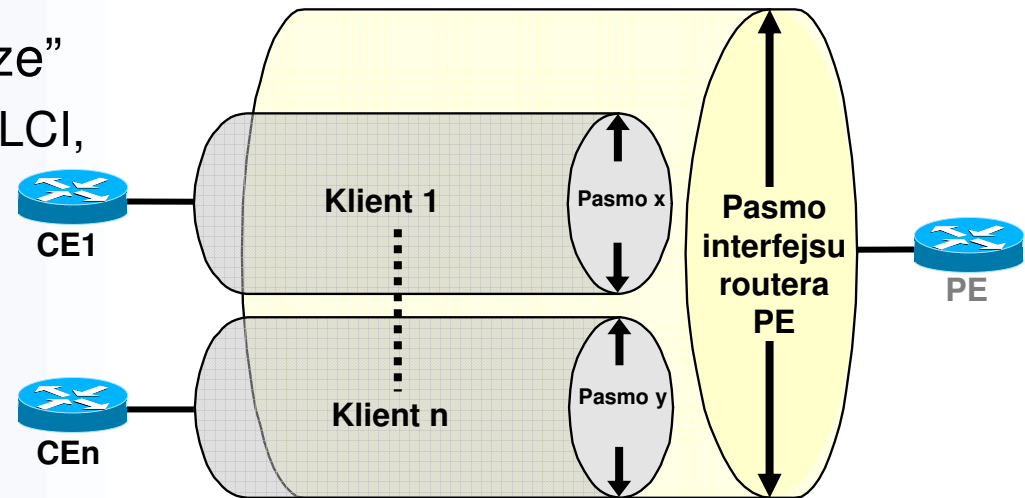
Domena MPLS

- Do „sygnalizacji” QoS używamy 3-bitowego pola EXP w nagłówku pakietu MPLS
Kopia lub mapowanie z wartości pola DSCP oryginalnego pakietu IP
- Decyzja o trafieniu pakietu do konkretnej klasy usługowej może odbywać się na podstawie wartości pola EXP, lub IP Precedence/DSCP oryginalnego pakietu

Przykład zastosowania modelu Diffserv

Projekt wysokopoziomowy – „bez QoS”

- Klient nabywa „wirtualne łącze” reprezentowane przez FR DLCI, ATM VC, Ethernet VLAN na fizycznym interfejsie PE



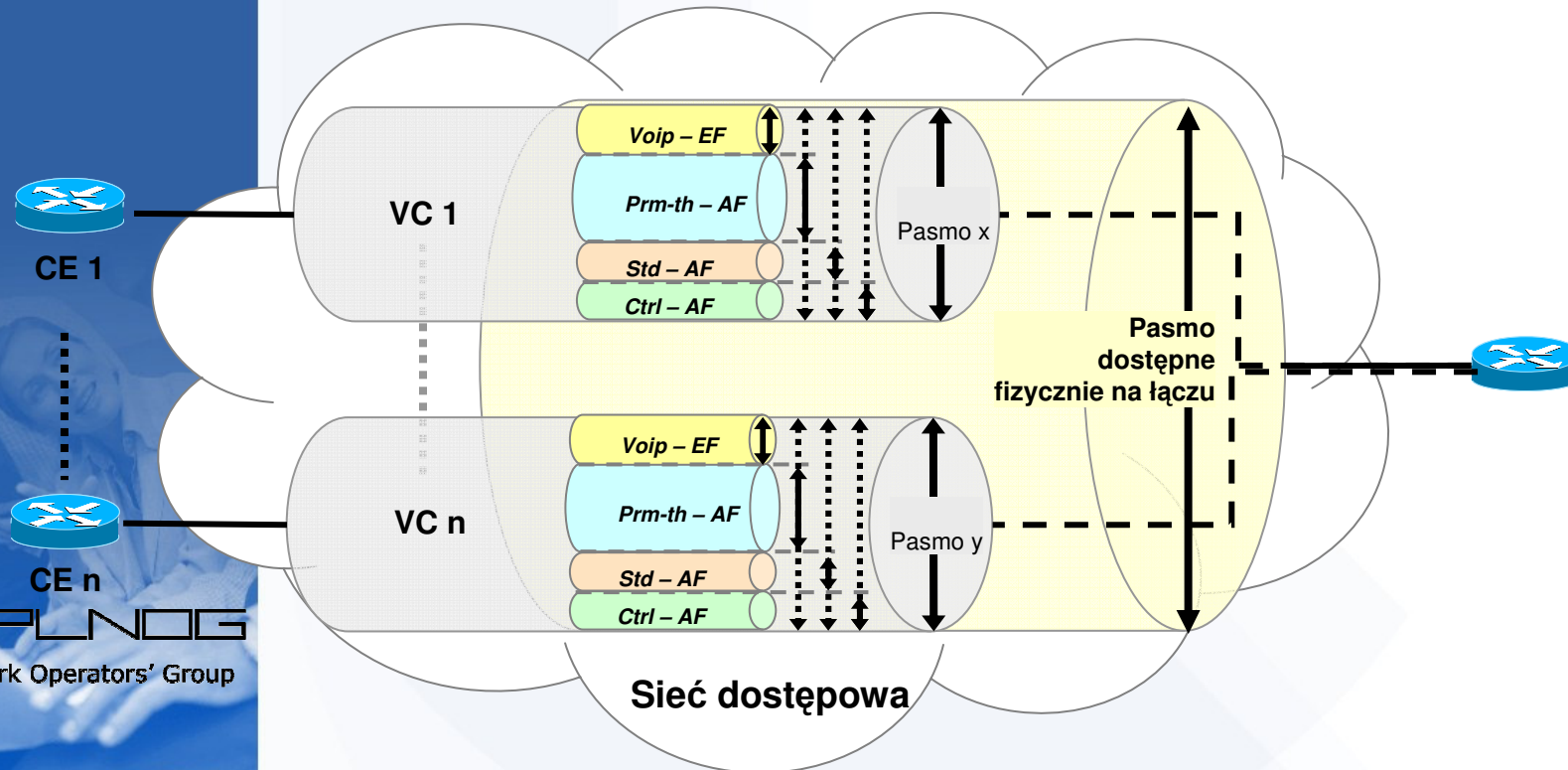
- Klient kupuje konkretne gwarantowane pasmo: $CIR = \min CIR = x$ kbit/s
- SP zapewnia poziom usług przez skonfigurowanie odpowiedniego kształtowania ruchu na poziomie podinterfejsu routera PE (i CE)

Model zakłada brak strat w sieci dostępowej

...zapewnia przy okazji separację klientów (MPLS)

Projekt wysokopoziomowy – z QoS

- Klient nabywa dostęp z IP QoS
- Do (pod)interfejsu routera PE/CE dołączania jest konfiguracja przycinania/gwarancji ruchowych dla tego konkretnie klienta
jeśli klient przekracza CIR, używane jest LLQ (low-latency queuing) do zagwarantowania właściwej obsługi klasy/ruchu priorytetowego



Aplikacje w sieci

| Application | L3 Classification | | |
|-----------------------|-------------------|----------|-------|
| | IPP | PHB | DSCP |
| Routing | 6 | CS6 | 48 |
| Voice | 5 | EF | 46 |
| Interactive-Video | 4 | AF41 | 34 |
| Streaming Video | 4 | CS4 | 32 |
| Mission-Critical Data | 3 | - | 25 |
| Call Signaling | 3 | AF31/CS3 | 26/24 |
| Transactional Data | 2 | AF21 | 18 |
| Network Management | 2 | CS2 | 16 |
| Bulk Data | 1 | AF11 | 10 |
| Scavenger | 1 | CS1 | 8 |
| Best Effort | 0 | 0 | 0 |

Klasy na brzegu – ile ?

| Numer klasy | Nazwa klasy używana w projekcie | Profil ruchu |
|-------------|---------------------------------|---|
| 6,7 | management | Zarządzanie siecią, mechanizmy sieci |
| 5 | voice | VoIP oraz video konferencje |
| 4 | video | Video streaming (np. IPTV) |
| 3 | biznes_critical | Ruch biznesowy Interaktywny krytyczny ze względu na działalność klienta |
| 2 | biznes_interactive | Ruch biznesowy Interaktywny (np. SAP, Telnet, SNA, Citrix) |
| 1 | biznes_transactional | Ruch biznesowy transakcyjny (np. Intranet, FTP, Poczta korporacyjna) |
| 0 | be | Ruch internetowy |

Model z podziałem na ruch w kontrakcie i poza kontraktem

| Klasa | DSCP | | IP Prec (CS) | Binarnie |
|---------|----------------|---------------|--------------|--------------------|
| VoIP | D_{voip} | 40 46 (EF) | 5 5 | 101 000 101 110 |
| Bus_in | D_{bus_in} | 24 | 3 | 011 000 |
| Bus_out | D_{bus_out} | 32 | 4 | 100 000 |
| Std | D_{std} | 0 | 0 | 000 000 |
| RP | D_{rp} | 48 | 6 | 110 000 |
| Mgt | D_{mgt} | 16 | 2 | 010 000 |

Klasy brzegowe

- Zakładamy wyodrębnienie czterech klas usługowych pomiędzy PE-CE:

VoIP

Klasa dla aplikacji interaktywnych – wymagających konkretnego pasma, niskiej stopy utraty pakietów i bardzo ścisłych wymogów co do opóźnienia i jego zmiany (jitteru) – tak jak zwykle VoIP

Dane biznesowe (Bus)

Aplikacje przesyłające ruch wymagający priorytetu i konkretnego pasma – ale mniej czułego na opóźnienia/jitter

Dane standardowe (Std)

Cała reszta ruchu klienta

Ruch kontrolny (Ctrl)

Dedykowana do ruchu kontrolnego (administracyjnego)

} Klasy 'widoczne' przez użytkownika

} Klasa wewnętrzna

- Ilość klas zdefiniowana na brzegu niekoniecznie musi od razu narzucać ilość klas w szkieletcie sieci

Wymagania dla różnych typów ruchu

- **Cisco Service Provider Quality of Service Design Guide**
- **ITU-T Y.1540** (Internet protocol data communication service – IP packet transfer and availability performance parameters)
- **ITU-T Y.1541**(Network performance objectives for IP-based services)
- **ITU-T G.114** (General Recommendations on the transmission quality for an entire international telephone connection)

| Nazwa klasy używana w projekcie | Opóźnienie (w jedną stronę) | Warjancja opóźnień (Jitter) | Utrata pakietów |
|---------------------------------|------------------------------|-----------------------------|-----------------|
| voice | 50 ms | 20 ms | 1,00% |
| video | 70 ms | 30 ms | 0,20% |
| biznes_critical | 70 ms | Brak gwarncji | 0,20% |
| biznes_interactive | 70 ms | Brak gwarncji | 0,20% |
| biznes_transactional | 70 ms | Brak gwarncji | 0,20% |
| be | Brak gwarncji | Brak gwarncji | Brak gwarncji |

SLA: klasa VoIP

- SLA definiujemy zarówno dla warstwy brzegowej jak i szkieletu

Ruch jest przycinany na brzegu zgodnie z kontraktem pomiędzy usługodawcą a klientem – do R_{voip} z dopuszczalnym pikiem B_{voip}

pole DSCP dla ruchu pasującego do profilu oznaczane jest wartością D_{voip}

ruch nadmiarowy jest odrzucany

Mała ilość ruchu odrzucanego

klasa zaprojektowana do bardzo niskiej stopy utraty pakietów – zwykle jedynymi błędami są tu problemy w warstwie drugiej lub problemy sieci ($P_{voip} < 1\%$)

Niskie opóźnienie i jitter

dostawca oferuje opóźnienie end-to-end w najgorszym wypadku nie gorsze niż L_{voip}

SLA: klasa danych biznesowych

- Aplikacje wymieniające ruch w sieci, takie jak SNA, SAP R/3, etc.
- SLA definiujemy zarówno dla warstwy brzegowej jak i szkieletu
 - pole DSCP dla ruchu pasującego do profilu oznaczane jest wartością D_{bus}

zapewnienie ograniczonego opóźnienia i jitteru nie występuje zwykle na brzegu sieci, choć może być określone dla szkieletu – jako L_{bus}

typowy RTT w granicach 25-100ms, w zależności od medium (opóźnienia związanego z propagacją i serializacją) i konkretnej relacji ruchowej w punktach dostępu do domeny Diffserv w architekturze MPLS

SLA: klasa danych biznesowych c.d.

- Definicja SLA – ciąg dalszy założeń

Niska stopa utraty pakietów

sterowanie warunkami oferowanego SLA odbywa się przez kontrolowanie ilości jednocześnie obsługiwanych klientów

określony nieprzekraczalny próg utraty pakietów

klasa zaprojektowana do bardzo niskiej stopy utraty pakietów – zwykle jedynymi błędami są tu problemy w warstwie drugiej lub problemy sieci ($P_{bus} < 0.2\%$)

SLA: klasa ruchu standardowego

- Definicja SLA

Brak ograniczeń na brzegu

→ może wykorzystać nieużywane pasmo z innych klas ruchowych do wartości pasma wykupionego w usłudze

ruch oznaczany jest wartością DSCP D_{std}

Brak zapewnienia ograniczonego opóźnienia i jitteru

typowy RTT w granicach 30-100ms, w zależności od medium (opóźnienia związane z propagacją i serializacją)

Brak gwarancji utraty pakietów

SLA: klasa kontrolna

- Ruch protokołów routingu (RP) i zarządzający (Mgt)
 - Zapewniamy pasmo – R_{ctrl} – aby:
 - protokoły routingu mogły bezpiecznie wymienić informacje o osiągalności i podtrzymać zestawione sesje nawet jeśli łącze będzie przeciążone
 - dostawca mógł zarządzać urządzeniem – za pomocą protokołów takich jak TFTP, SNMP, SSH nawet jeśli łącze będzie przeciążone
 - ...klasa może wykorzystywać nieużywany ruch z innych klas (np. zdalny upload/download oprogramowania, etc.)
 - Ruch RP oznaczamy w polu DSCP wartością D_{rp}
 - Ruch Mgt oznaczamy w polu DSCP wartością D_{mgt}

Projekt brzegu sieci (niskopoziomowy)

Klasa VoIP

- Klasa ruchowa VoIP

EF → obsługiwana w PQ (kolejce priorytetowej) dla zapewnienia najniższego opóźnienia i możliwie stałego, niskiego jitteru

ograniczany do $R_{voip} = 256\text{kbit/s}$ (w omawianym scenariuszu)

→ nadmiar odrzucany

rozmiar chwilowego ruchu nadmiarowego zależy od najgorszego przewidywanego opóźnienia

$$B_{voip} / \text{pasmo_łacza} + L_s \leq L_{voip}$$

gdzie L_s to wpływ najgorszego przewidywanego opóźnienia na ruch w klasie EF z uwagi na działanie systemowego schedulera i wielkości/ustawień tx-ring

$$\text{Zatem } B_{voip} \leq (L_{voip} - L_s) * \text{pasmo_łacza}$$

jeśli $L_{voip} = 15\text{ms}$ i $L_s = 10\text{ms}$:

$$\rightarrow B_{voip} \leq 1,250 \text{ bajtów}$$

Klasa VoIP c.d.

- Klasa ruchowa VoIP c.d.

Należy się upewnić, że rozmiar kolejki na ruch nadmiarowy jest wystarczający by obsłużyć ruch ze wszystkich jednoczesnych połączeń

$G.729-20ms = 66 \text{ bajtów na pakiet} * 50p/s = 26.4\text{kbit/s}$ na połączenie czyli 9 połączeń na 256kbit/s

$B_{voip} >$ maksymalny rozmiar kolejki na ruch nadmiarowy (czyli $9 * 66b = 594$ bajtów)

domyślne ustawienie to 200ms przy 256kbit/s = 6400 bajtów

ruch sygnalizacyjny VoIP współdzieli kolejkę priorytetową (PQ) z ruchem VoIP

rozmiar kolejki nadmiarowej może wymagać dopasowania by „pomieścić” sygnalizację

Klasa biznesowa

- Gwarantacja pasma $R_{bus} = 1.4\text{Mbit/s}$ w tym scenariuszu (to ~80% pozostałego pasma po odjęciu klasy VoIP)
ruch oznaczany D_{bus}
- WRED

Klasa biznesowa cd.

Tabela 12 Ustawienia WRED dla klasy Business.

| Link Speed in kbps | B | Link BW minT H | Link BW maxT H | Class BW 10% in kbps | Class BW 10% minTH | Class BW 10% maxTH | Class BW 20% in kbps | Class BW 20% minTH | Class BW 20% maxTH | Class BW 25% in kbps | Class BW 25% minTH | Class BW 25% maxTH | Class BW 30% in kbps | Class BW 30% minTH | Class BW 30% maxTH |
|--------------------|--------|----------------|----------------|----------------------|--------------------|--------------------|----------------------|--------------------|--------------------|----------------------|--------------------|--------------------|----------------------|--------------------|--------------------|
| 64 | 6 | 3 | 9 | | | | | | | | | | | | |
| 128 | 11 | 3 | 9 | | | | | | | | | | | | |
| 256 | 22 | 3 | 9 | | | | | | | | | | | | |
| 512 | 43 | 3 | 9 | | | | | | | | | | | | |
| 1024 | 86 | 3 | 9 | | | | | | | | | | | | |
| 2048 | 171 | 6 | 18 | | | | | | | | | | | | |
| 10000 | 834 | 26 | 84 | 1000 | 3 | 9 | 2000 | 6 | 17 | 2500 | 7 | 21 | 3000 | 8 | 26 |
| 34684 | 2891 | 87 | 290 | 3468 | 9 | 29 | 6937 | 18 | 58 | 8671 | 22 | 73 | 10405 | 27 | 87 |
| 100000 | 8334 | 251 | 834 | 10000 | 26 | 84 | 20000 | 51 | 167 | 25000 | 63 | 209 | 30000 | 76 | 251 |
| 155000 | 12917 | 388 | 1292 | 15500 | 39 | 130 | 31000 | 78 | 259 | 38750 | 97 | 323 | 46500 | 117 | 388 |
| 622000 | 51834 | 1556 | 5184 | 62200 | 156 | 519 | 124400 | 312 | 1037 | 155500 | 389 | 1296 | 186600 | 467 | 1556 |
| 2400000 | 200000 | 6000 | 20000 | 240000 | 600 | 2000 | 480000 | 1200 | 4000 | 600000 | 1500 | 5000 | 720000 | 1800 | 6000 |

Klasa standardowa i kontrolna

- Klasa standardowa

Klasa AF czyli z gwarancją dla pasma

W tym scenariuszu zakładamy $R_{std} =$
~360kbit/s (~20% pozostałego pasma)

Wred

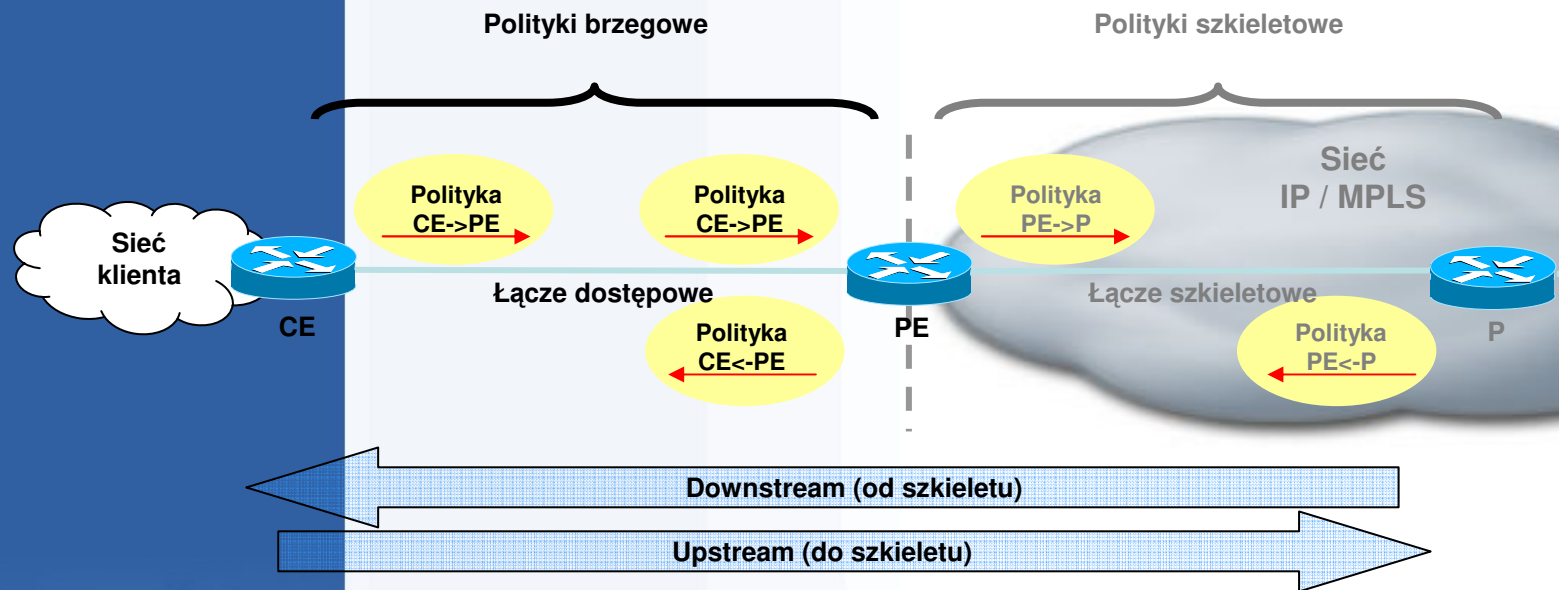
- Klasa kontrolna

Klasa AF czyli z gwarancją dla pasma

W tym scenariuszu zakładamy $R_{ctrl} =$
~20kbit/s (~1% pozostałego pasma)

Konfiguracja routerów CE i PE

Konfiguracja CE



- Konfiguracja dotyczy ruchu opuszczającego CE w stronę szkieletu

Komponenty MQC

1. Klasyfikacja ruchu

“class-map”

Identyfikuje ruch i
przydziela go do klasy

2. Polityka Diffserv

“policy-map”

Grupuje klasy w politykę
Opisuje sposób w jaki
Diffserv potraktuje klasy

3. Dołączenie polityki do konkretnego interfejsu

“service-policy”

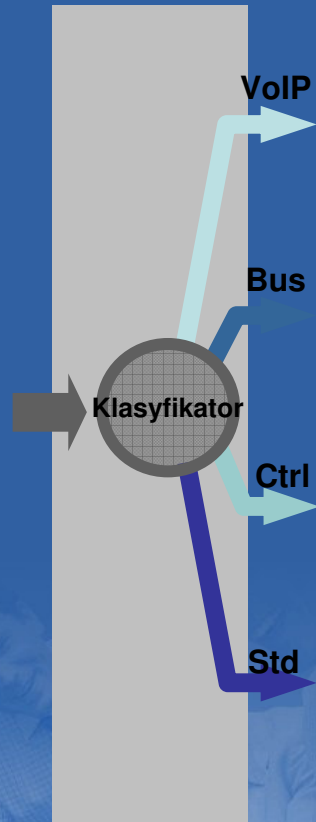
```
class-map match-any voip
  match ip dscp 40
  match access-group 100
class-map match-any bus
  match access-group 101
class-map match-all ctrl
  match access-group 103
  match access-group 104
```

!

```
policy-map DIFFSERV_POLICY
  class voip
    priority
    police 64000
  class bus
    bandwidth remaining percent 90
```

!

```
interface Serial0
  ip address 192.168.2.2 255.255.255.0
  service-policy output DIFFSERV_POLICY
```

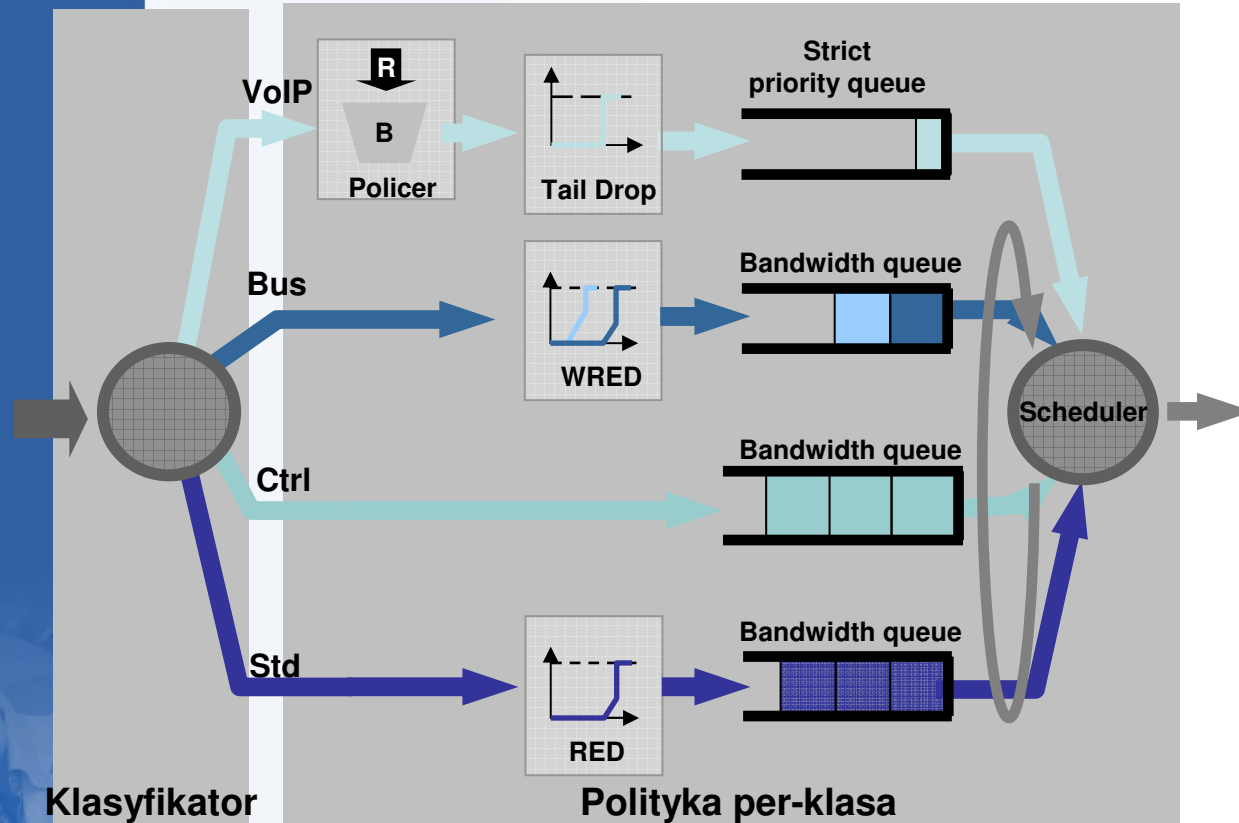


Klasyfikacja ruchu na CE

```
ip telnet tos 64
snmp-server ip dscp 16
!
access-list 100 permit udp any any eq 16383
access-list 100 permit udp any any eq 16384
access-list 101 permit <ruch biznesowy>
!
class-map match-any voip
  match ip dscp 40 46
  match access-group 100
class-map match-any bus
  match access-group 101
class-map match-any ctrl
  match ip dscp 16 48
```

- „Proste” (bazujące na DSCP) lub „skomplikowane” (bazujące na ACL) przypisanie ruchu do klas

Polityka dla każdej z klas



- Mechanizm LLQ zapewnia zróżnicowanie klas EF i AF jeśli klient przekroczy określone wartości ruchowe

Konfiguracja klasy VoIP

```
policy-map child
  class voip
    priority {256000}
    police 256000
    set ip dscp 46
  ...
```

- Polecenie *priority* określa że ruch należący do klasy ma być traktowany priorytetowo (klasa EF)
- Polecenie *police* określa bezwarunkowe przycinanie ruchu jeśli wykroczy poza wskazaną wartość – uznaje się, że deterministyczne zachowanie jest lepsze niż niedeterministyczne gdy kolejka mogłaby używać pasma z innych klas

Konfiguracja klasy biznesowej

```
policy-map child
...
class bus
  bandwidth 1400

  random-detect dscp-based
  random-detect exponential-weighting-constant 7
  random-detect dscp 32    23    69    1
  random-detect dscp 24    550   600   1
...
```

- Polecenie *bandwidth* wskazuje na klasę typu AF (w momencie przeciążenia gwarantujemy konkretne pasmo)
- WRED

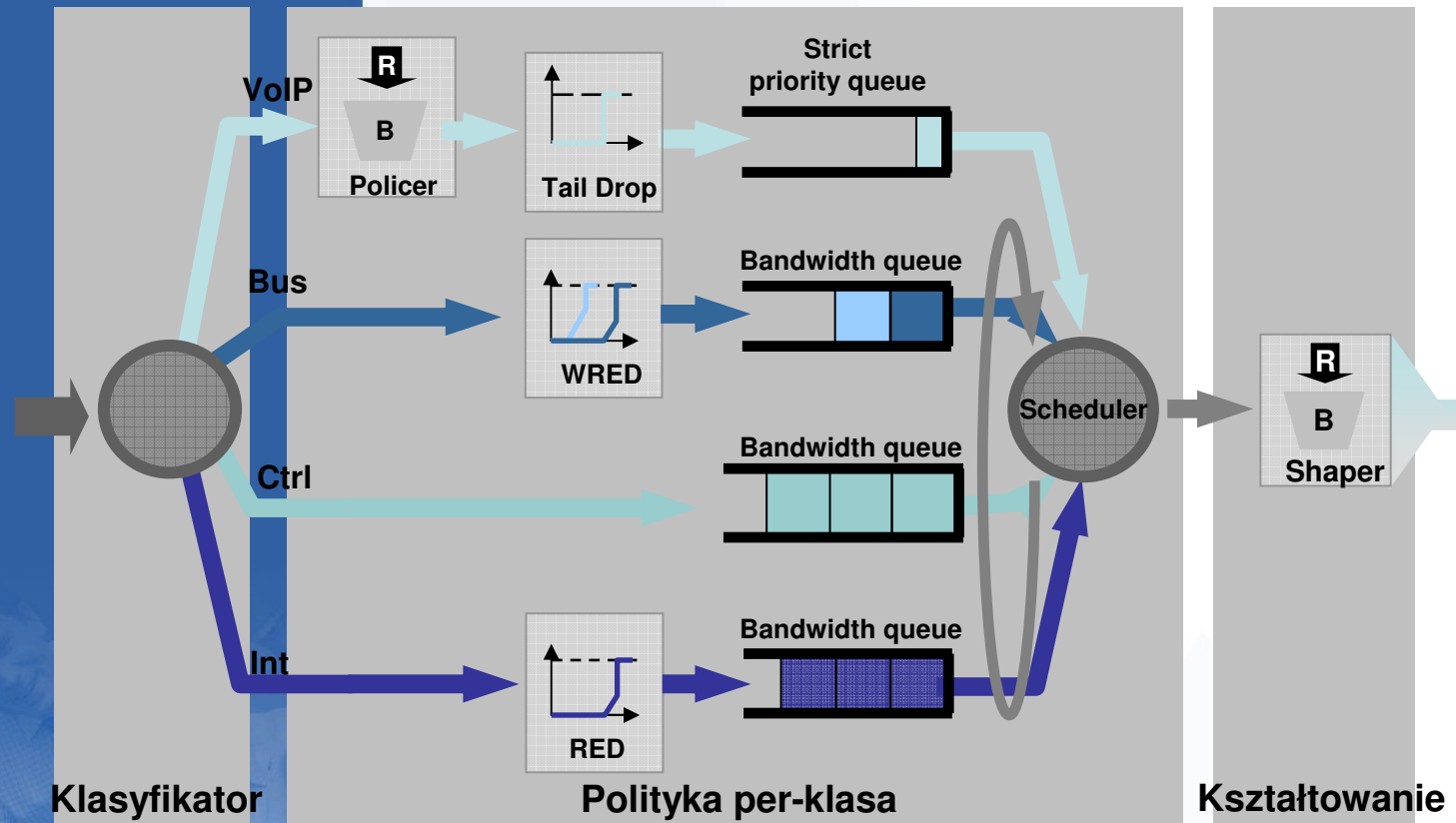
Konfiguracja klasy standardowej i kontrolnej

```
policy-map child
...
  class ctrl
    bandwidth 22

class class-default
  bandwidth 370
  random-detect dscp-based
  random-detect exponential-weighting-constant 5
  random-detect dscp 0 33 99 1
```

- Klasa standardowa wykorzystuje wbudowaną klasę “class-default”
 - dotyczy całego ruchu do tej pory nie ‘obsłużonego’
- pracuje WRED – aby zmaksymalizować przepustowość sesji TCP

Zbiorcza polityka kształtowania ruchu

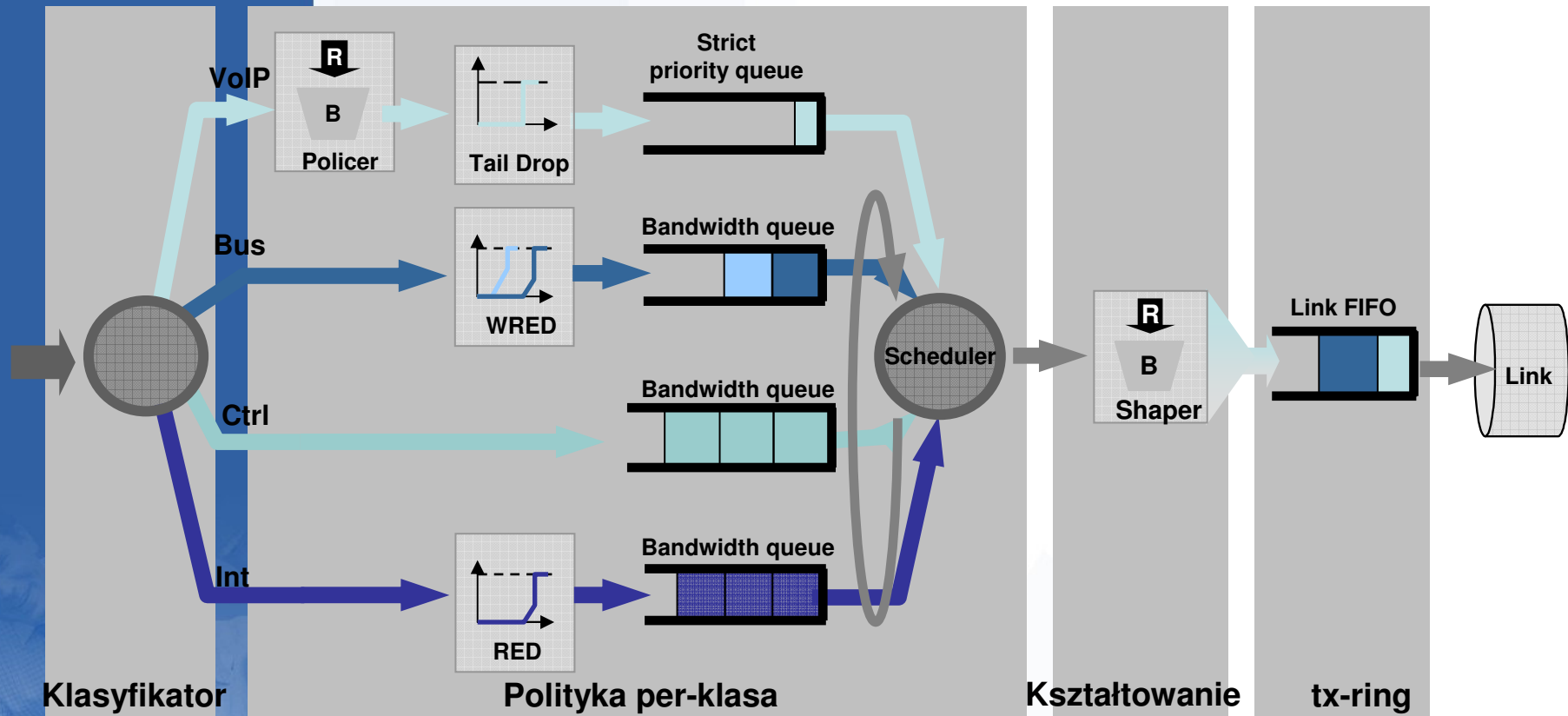


- Polityka kształtowania ruchu obejmuje definicje QoS klas „dzieci” – na (pod)interfejsie routera PE skierowanego w stronę CE

Zbiorcza polityka kształtowania ruchu

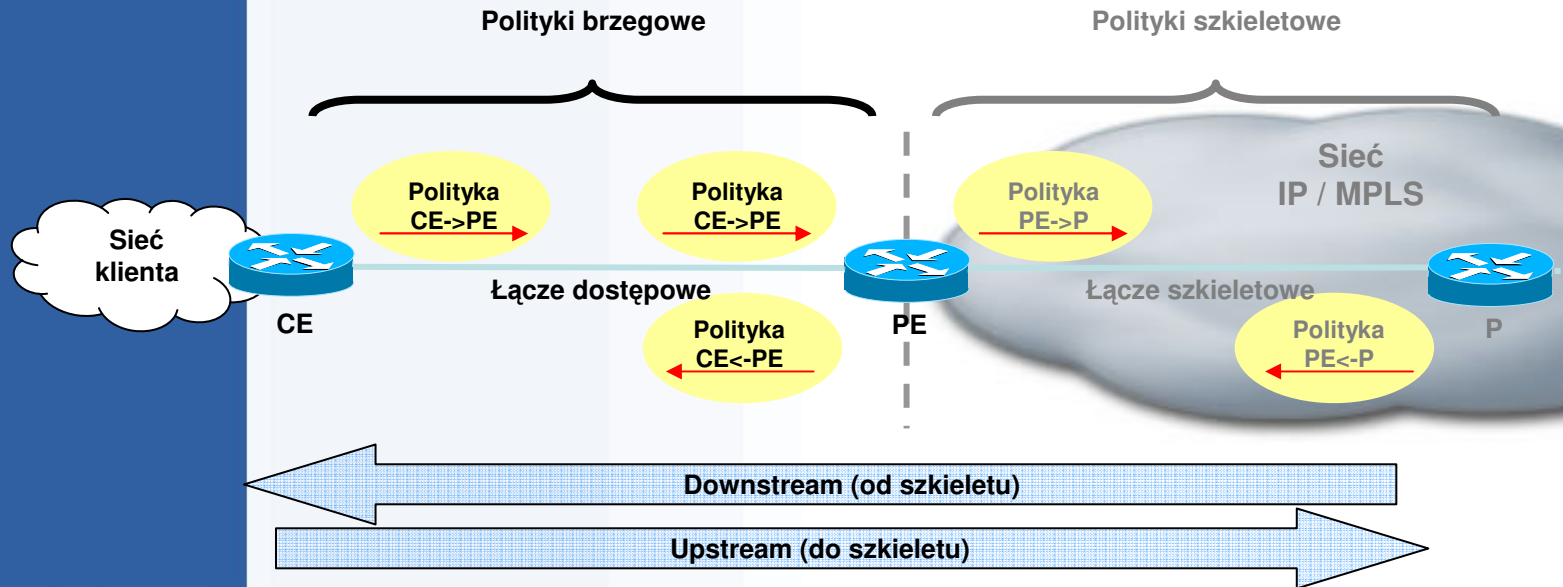
```
policy-map parent
  class class-default
    shape average 2048000
    service-policy child
!
interface GigabitEthernet1/0/0.102
  encapsulation dot1Q 102
  ip address 192.168.48.4 255.255.255.0
  service-policy output parent
```

tx-ring



- tx-ring to kolejka FIFO przypisana do interfejsu przed sterownikiem sprzętu
- IOS w normalnych warunkach sam optymalnie dobiera długość kolejki tx-ring

Konfiguracja PE



- Konfiguracja QoS dotyczy ruchu od strony szkieletu (routera PE) w stronę CE
- W modelu „niezarządzanego” CE powinna dotyczyć również ruchu wejściowego do PE od strony CE choć dodatkowa, choćby ‘ogólna’ polityka QoS może nie będzie zbyt problematyczna w utrzymaniu a zabezpieczy sieć w razie nieprzewidzianych zdarzeń

Konfiguracja PE

- Konfiguracja podobna do tej na CE:
cały ruch od klienta powinien zostać
właściwie oznaczony w polu DSCP na CE

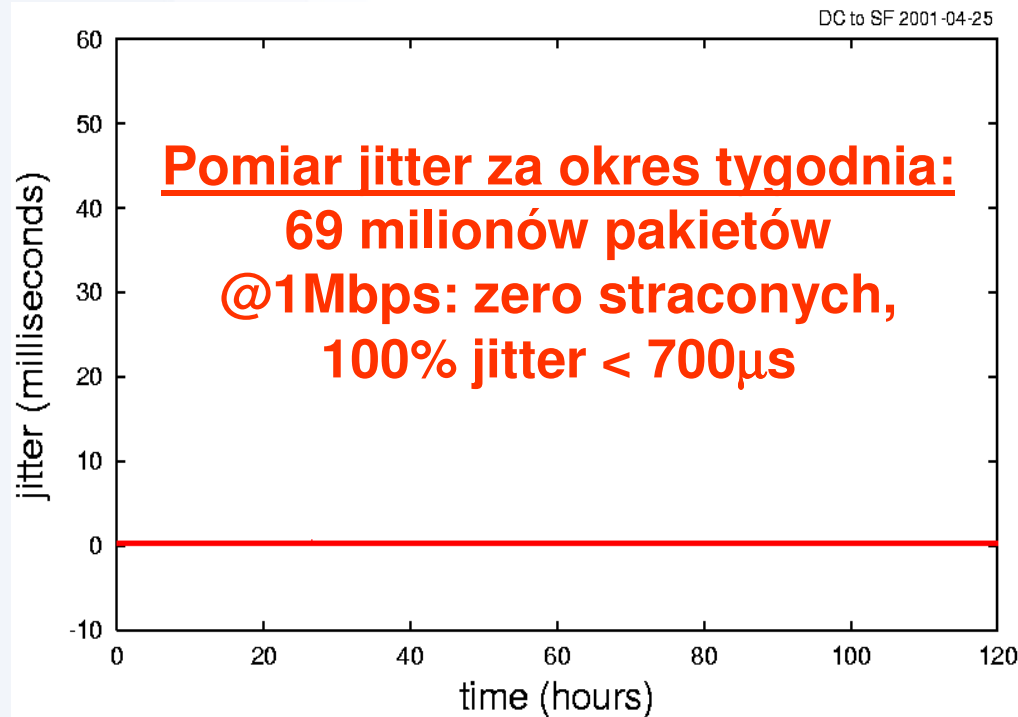
```
class-map match-any voip
  match ip dscp 40 46
class-map match-any bus
  match ip dscp 24 32
class-map match-any ctrl
  match ip dscp 16 48
```

Projektowanie szkieletu: Czy potrzebujemy tu Diffserv?

Jak myśleć o ruchu w szkielecie?

- W szkielecie ruch jest agregowany, wymagania QoS do poszczególnych klas przetłumaczyć można na wymogi co do pasma
 - Problem odpowiedniej konfiguracji QoS można zatem efektywnie zredukować do odpowiedniego zarządzania pasmem
- Aby upewnić się, że gwarancje związane z opóźnieniem i jitter nie zostaną naruszone, szkielet może wymagać dodatkowego pasma aby być w stanie „pomieścić” dodatkowy ruch z poszczególnych klas

Więcej pasma w szkielecie (źródło: Stephen Casner, NANOG 22)



- Nadmiar pasma w szkielecie może być efektywnym narzędziem dla wsparcia polityki QoS
- Dostępne pasmo jest większe niż łączne szczytowe zapotrzebowanie na pasmo
- Sieć jest prosta do wdrożenia i utrzymania – skomplikowane polityki QoS znajdują się tylko na jej brzegu

Więcej pasma – czy to praktyczne?

- Nie zawsze jest to możliwe (uzasadnione ekonomicznie) a przeciążenie sieci może i tak się wydarzyć
 - Problemy z planowaniem
 - Awarie sieci**
 - Gwałtowny wzrost zapotrzebowania na ruch
 - Ataki DDoS**
 - TCP zajmuje całe dostępne pasmo jeśli może
- Jednakowa obsługa – nie ma rozróżnienia pomiędzy ruchem ważniejszym i mniej ważnym
 - np. w przypadku problemów jakość połączeń głosowych ucierpi
- Złożoność tego podejścia polega na tym, że trudno zapewnić nadmiarowość pasma w szkielecie w każdym możliwym scenariuszu

Zatem Diffserv w szkielecie?

- Praktyczniejsze podejście zakłada, że nie zawsze można efektywnie kosztowo lub praktycznie wyposażyć szkielet w większe przepustowości niż to wymagane
- W szczycie w szkielecie może być mniej pasma niż suma ruchu oferowanego
- Wdrożenie modelu Diffserv odsuwa ryzyka związane z awariami – zapewnia rozróżnienie usług
- ...będzie jednak trudniejsze we wdrożeniu i utrzymaniu niż nadmiarowe pasmo

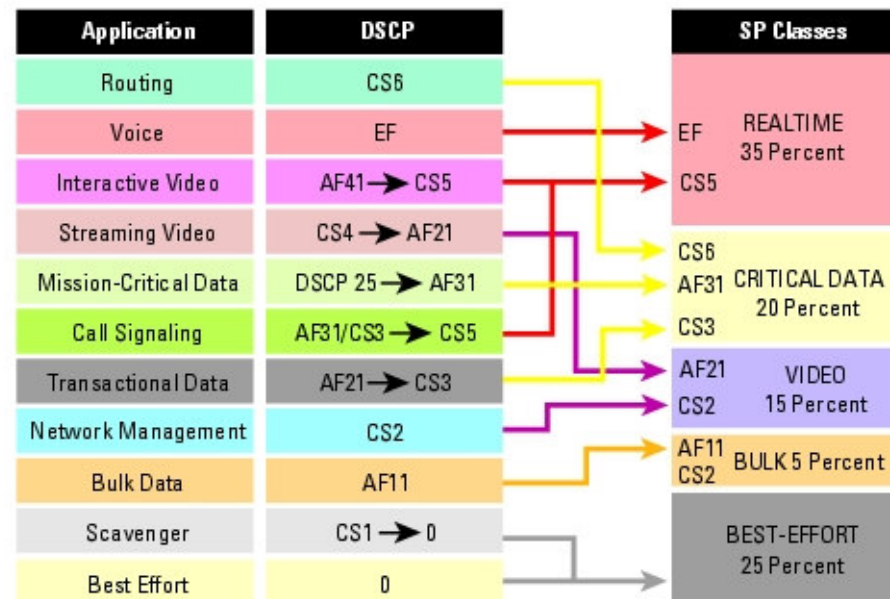
Projektowanie szkieletu: Implementacja

Klasy w szkielecie – ile?

- Szkielet nie musi mieć tylu klas ruchowych ile brzeg sieci

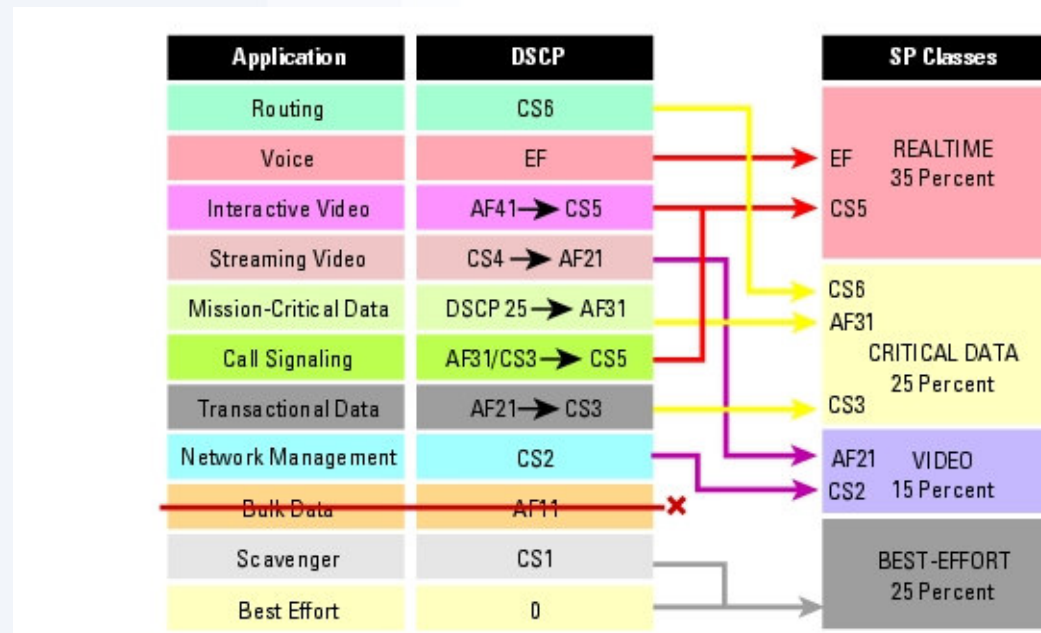
Klasy w szkielecie – ile?

Five-Class Service Provider Model Remarking Diagram



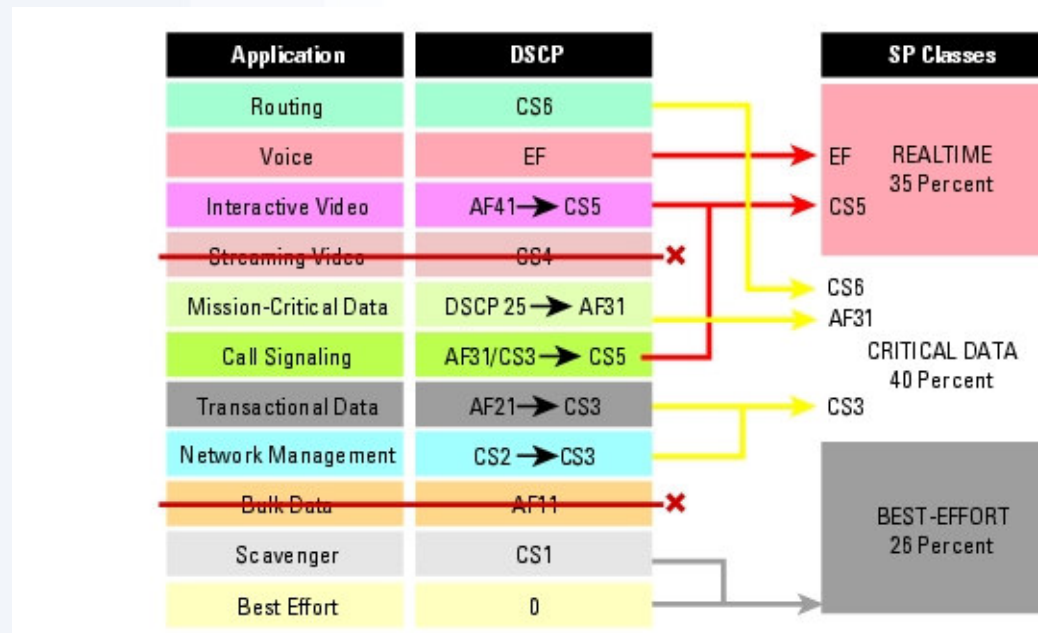
Klasy w szkielecie – ile?

Four-Class Service Provider Model Remarking Diagram



Klasy w szkielecie – ile?

Three-Class Service Provider Model Remarking Diagram



Klasy w szkielecie – ile?

Przykład rzeczywisty

| QOS Klienta | | | | QOS Core | | | | | |
|-----------------|---------|---|---|---------------------------------|-------------|-----|---|---|---------------------------------|
| Wartość IP Prec | IP PREC | | | Nazwa klasy używana w projekcie | Wartość EXP | EXP | | | Nazwa klasy używana w projekcie |
| 7 | 1 | 1 | 1 | management | 7 | 1 | 1 | 1 | management |
| 6 | 1 | 1 | 0 | | 6 | 1 | 1 | 0 | |
| 5 | 1 | 0 | 1 | voice | 5 | 1 | 0 | 1 | voice |
| 4 | 1 | 0 | 0 | video | 4 | 1 | 0 | 0 | video |
| 3 | 0 | 1 | 1 | biznes_critical | 3 | 0 | 1 | 1 | biznes |
| 2 | 0 | 1 | 0 | biznes_interactive | | | | | |
| 1 | 0 | 0 | 1 | biznes_transactional | | | | | |
| 0 | 0 | 0 | 0 | be | 0 | 0 | 0 | 0 | be |

Podział pasma

| Wartość EXP | Nazwa klasy używana w projekcie | Rodzaj łącza | | | |
|-------------|---------------------------------|------------------------|-------------------|--------------------|-------------------|
| | | STM-16 | | STM-1 | |
| | | Pasmo | Udział procentowy | Pasmo | Udział procentowy |
| 7 | management | 124,00 Mbps | 5,00% | 7,75 Mbps | 5,00% |
| 6 | | | | | |
| 5 | voice | 620,00 Mbps | 25,00% | 38,75 Mbps | 25,00% |
| 4 | video | 297,60 Mbps | 12,00% | 0,00 Mbps | 0,00% |
| 3 | biznes | 421,60 Mbps | 17,00% | 32,55 Mbps | 21,00% |
| 2 | | | | | |
| 1 | | | | | |
| 0 | be | 1 016,80 Mbps | 41,00% | 75,95 Mbps | 49,00% |
| | | Σ 2 480,00 Mbps | 100,00% | 155,00 Mbps | 100,00% |

Klasa ruchu czasu rzeczywistego

- W konserwatywnym ujęciu - < 25% fizycznego pasma łączy
- Dokładniejsze określenie gwarancji można uzyskać stosując scenariusze „a co jeśli...”:
 - stworzyć matrycę relacji ruchowych i wymogów co do pasma
 - narzędzia symulacyjne mogą stworzyć model takiej sieci i określić ryzyka związane z utratą poszczególnych węzłów i łączy
 - warto w tym momencie pomyśleć o rozbudowie sieci w przyszłości

Klasa ruchu czasu rzeczywistego

- Kolejka priorytetowa nie powinna „zagłodzić” innego ruchu w szkielecie, jest zatem idealna z punktu widzenia klas ruchowych wymagających gwarancji nieprzekraczalnego opóźnienia i jitter
- Należy odgórnie ograniczać ze względu na możliwe sytuacje gdy:
 - boimy się ataku DDoS z użyciem ruchu oznaczonego EF
 - wolimy odrzucać pakiety niż zaburzać jitter lub zwiększać opóźnienie w przypadku przeciążenia łącz/urządzeń agregacyjnych

Klasa ruchu biznesowego

- Klasa obsługuje 17% pasma łącza

Warto zastosować WRED dla zapewnienia odpowiedniego porządku odrzucania ruchu z klas z rozróżnieniem poszczególnych rodzajów ruchu biznesowego oraz optymalizacji sesji TCP

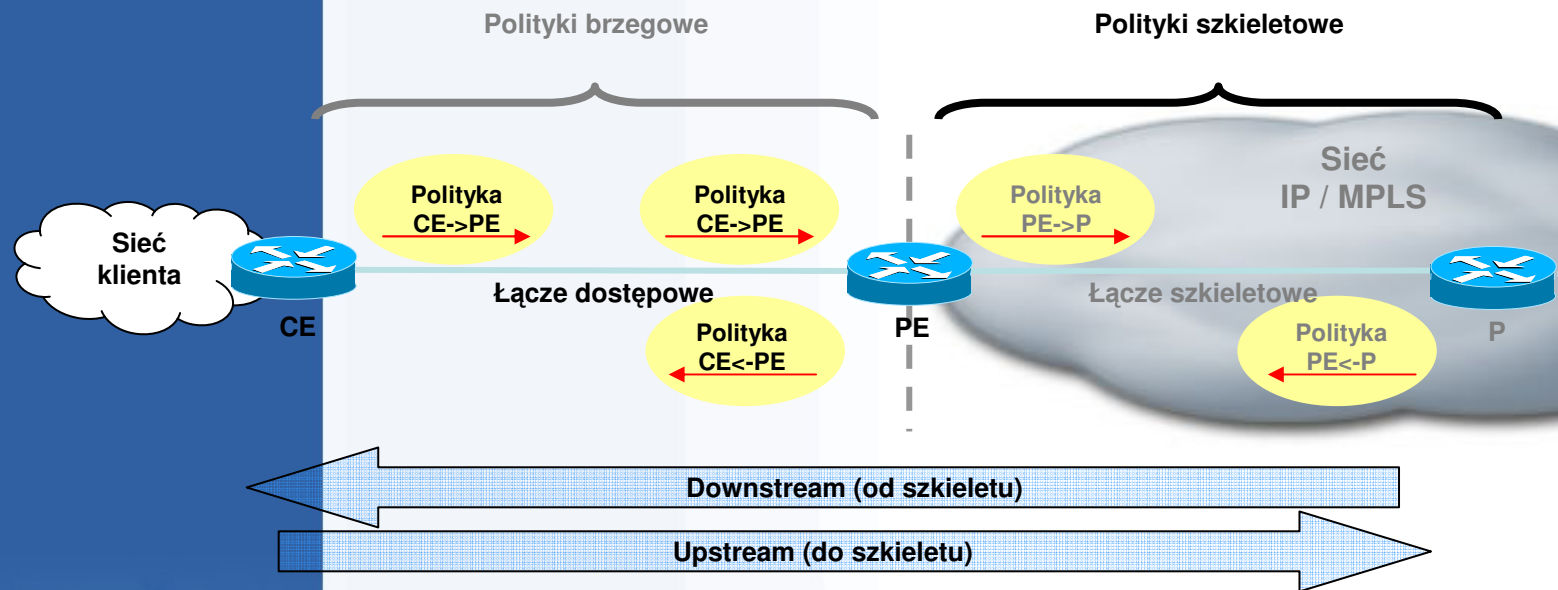
Klasa standardowa i kontrolna

- Dla klasy kontrolnej gwarantujemy 41% pasma łącza
- Dla klasy standardowej gwarantujemy 5% pasma łącza

może korzystać z nieużywanego obecnie pasma
stosujemy WRED

Projektowanie szkieletu: Konfiguracja routerów P i PE

Konfiguracja P i PE



- Polityka dotyczy ruchu wychodzącego wszystkimi interfejsami
- Obsługa ruchu wchodzącego

Konfiguracja – Cisco 12000

Stara notacja MDRR

```
cos-queue-group STM-16
precedence 0 random-detect-label 0
precedence 3 queue 1
precedence 3 random-detect-label 3
precedence 4 queue 3
precedence 5 queue low-latency
precedence 5 random-detect-label 5
precedence 6 queue 2
precedence 7 queue 2
random-detect-label 0 4557 15190 1
random-detect-label 3 1302 4340 1
random-detect-label 5 3662 3663 1
exponential-weighting-constant 10
queue 0 41
queue 1 17
queue 2 5
queue 3 12
queue low-latency alternate-priority 1
```

```
int pos 5/0
tx-cos STM-16
```

- WRED wykorzystuje wartość pola MPLS-EXP

Konfiguracja – Cisco 12000 MDRR !

```
class-map match-any rt
  match ip dscp 40 46
  match mpls experimental 5
class-map match-any bus
  match ip dscp 24 32
  match mpls experimental 3 4
class-map match-any ctrl
  match ip dscp 16 48
  match mpls experimental 2 6
!
policy-map oc48_policy
  class rt
    priority
    police 1200000000 150000000
      conform transmit exceed drop
  class bus
    bandwidth remaining percent 90
    random-detect
    random-detect pre 4 500 1012 1
    random-detect pre 3 1500 9692 1
  class ctrl
    bandwidth remaining percent 1
    random-detect
    random-detect pre 2 1500 9692 1
    random-detect pre 6 1500 9692 1
  class class-default
    bandwidth remaining percent 9
    random-detect
    random-detect pre 0 1500 9692 1
```

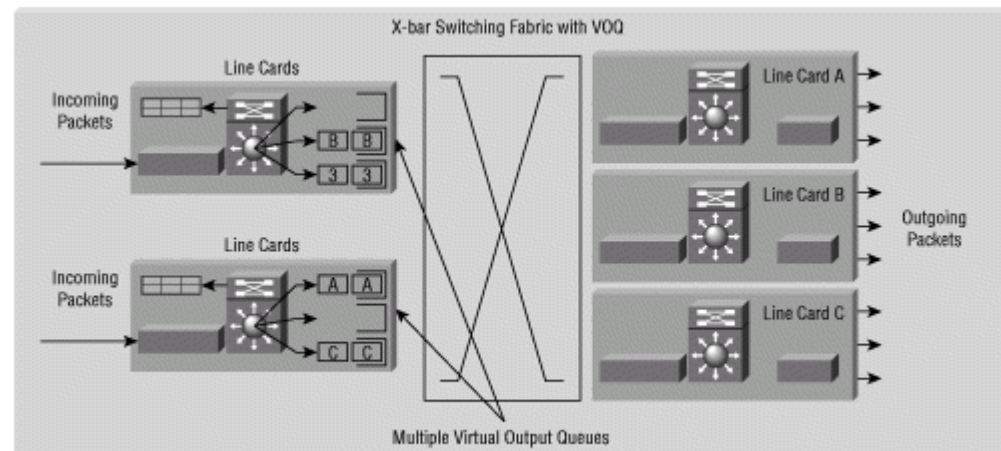
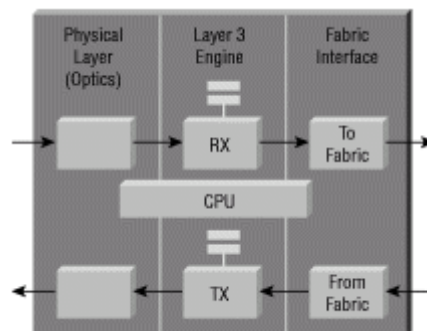
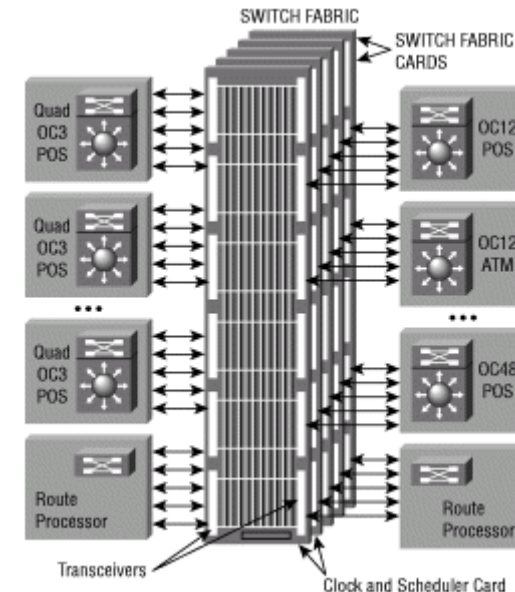
- WRED wykorzystuje wartość pola MPLS-EXP jeśli skonfigurowany zostaje do bazowania na IP Precedence

To fabQoS – Cisco 12000

```

rx-cos-slot 1 TO-FABRIC
rx-cos-slot 2 TO-FABRIC
rx-cos-slot 3 TO-FABRIC
rx-cos-slot 5 TO-FABRIC
!
slot-table-cos TO-FABRIC
  destination-slot all ToFab
!
cos-queue-group ToFab
  precedence 3 queue 1
  precedence 4 queue 1
  precedence 5 queue low-latency
  precedence 6 queue 1
  precedence 7 queue 1
  queue 0 41
  queue 1 34
  queue low-latency alternate-priority 1
!

```



Projektowanie szkieletu: Przenoszenie oznaczeń QoS i tunelowanie Diffserv w MPLS

Jakiego używać schematu?

- Dwa podejścia:

Model #1: jednolita architektura end-to-end

polityki wszystkich urządzeń identyczne z punktu widzenia traktowania ruchu

można wykorzystać domyślne zachowanie – kopiowanie wartości pola IP Precedence do pola MPLS EXP i propagowania go w stosie etykiet

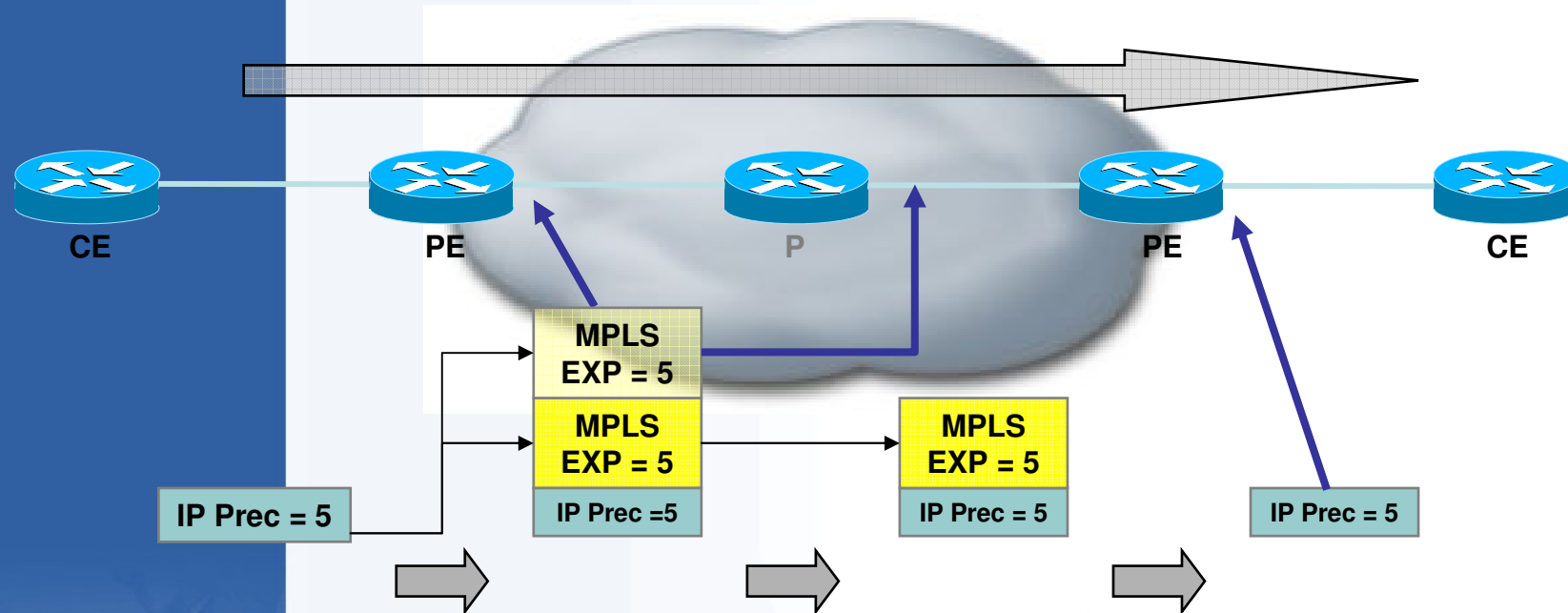
Model #2: różne modele QoS na brzegu i w szkielecie

różne polityki współistnieją w ramach globalnego modelu

wymaga od urządzeń i konfiguracji upewnienia się, że oznaczenia QoS są przenoszone w sposób transparentny

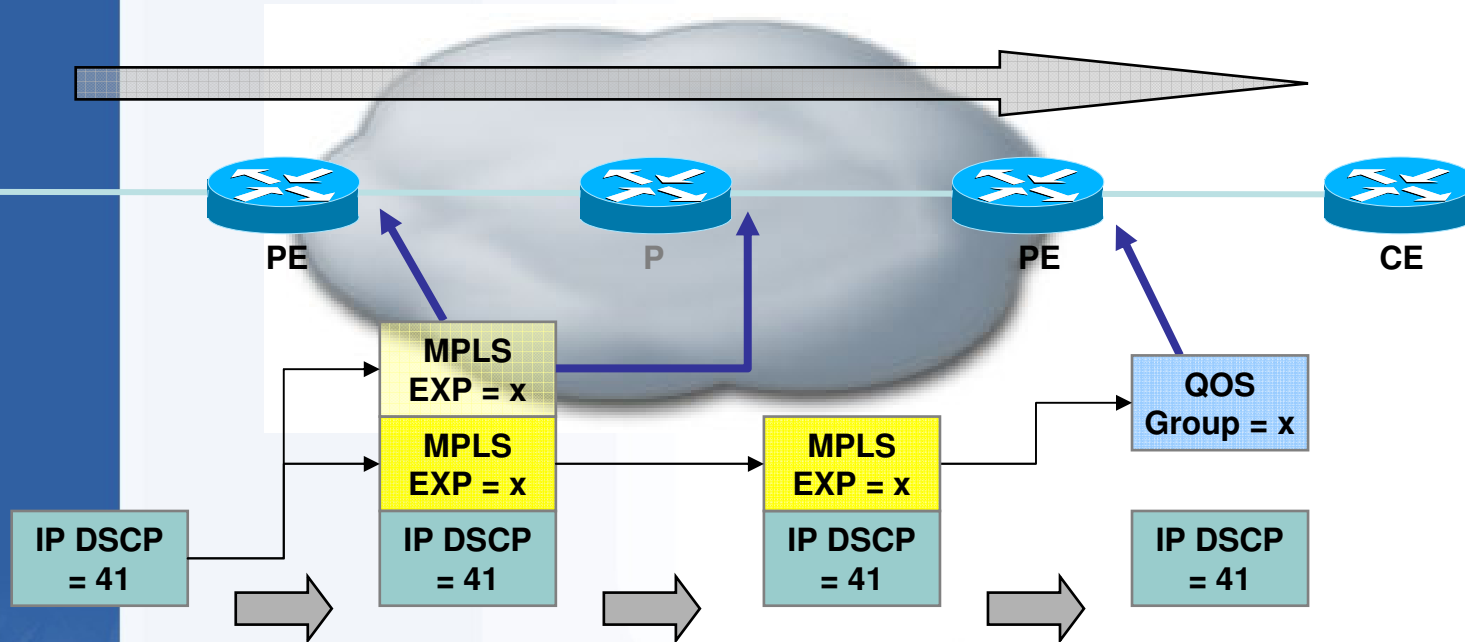
...zgodnie z modelami określonymi w RFC3270 (pipe/short pipe)

Model #1: zachowanie domyślne



- Wartość pola IP Precedence (pierwsze trzy bity DSCP) kopiowane są do pola MPLS EXP
- Pole EXP jest kopiowane do nowych etykiet w trakcie ich wymiany podczas podróży pakietu przez szkielet
- Przy zdjęciu ostatniej etykiety wartość pola **EXP nie jest kopiowana do pakietu IP** przed jego wystąpieniem

Model #2: MPLS Diffserv Tunnelling



- Wartość pola EXP jest ustalana niezależnie od wartości oryginalnego pola IP Precedence / DSCP
- Oryginalne oznaczenie nie zostaje zmienione
- Klasyfikacja na routerze PE którym pakiet opuszcza sieć wykonywana jest na podstawie pola EXP za pomocą konfiguracji qos-group i discard-class
- Dokładniej proces opisuje dokument na CCO: ID 47815

Parę uwag na koniec...

- ...cokolwiek nie wyda się bardziej atrakcyjne, warto pamiętać o prostych uwagach które mogą zapewnić prostsze życie:
 - nie używaj DSCP **56** – czyli MPLS EXP **7** –używają ich keepalive
 - DSCP 40 – czyli MPLS EXP 5 jest najczęściej wykorzystywany do obsługi VoIP (mimo, że wartością zalecaną dla ruchu EF jest DSCP 46)
 - stosuj wprost oznaczenie DSCP 0 (MPLS EXP 0) dla ruchu „domyślnego” – mniej konfiguracji i zmiany oznaczeń dla większości ruchu nie podlegającego wprost klasyfikacji
- Model konfiguracji będzie zależał od konkretnego sprzętu do dyspozycji – należy wziąć pod uwagę wszelkiego rodzaju notki dotyczące platformy i oprogramowania

Dziękuję za uwagę



Lumena SA
ul. Reja 6, 02-053 Warszawa
tel. +48 22 5765500
fax +48 22 5765501
www.lumena.com

Bezpiecznej drogi do domu !