

QoS setting MQC cisco design

Massimiliano Sbaraglia

INTSERV and DIFFSERV differences

▶ Integrated Services

- ▶ Ogni flusso ha una esplicita reservation end-to-end via RSVP per la richiesta di specifiche QoS
- ▶ Non scala bene su reti di grandi dimensioni in quanto richiede di mantenere le informazioni di ogni flusso/path su ogni singolo routers della catena end-to-end QoS.

▶ Differentiated Services

- ▶ Il traffico è raggruppato in classi di servizio
- ▶ La classificazione è definita a livello network edge inbound
- ▶ La classificazione è codificata all'interno dei pacchetti: Marking o Colorazione
- ▶ **Marking:**
- ▶ Definisce la classificazione dei pacchetti con diverse metodologie:

- ▶ IPv4 e IPv6 (layer 3 TOS)
 - ▶ DSCP (6 bit)
 - ▶ IP Precedence (3 bit)

- ▶ Layer 2
 - ▶ Frame Relay DE (1 bit)
 - ▶ MPLS EXP (3 bit)
 - ▶ 802.1q/ISL CoS (3 bit)



DSCP and CS value

- ▶ Differentiated Services Code Point classifica i pacchetti su determinati valori, quali:
 - ▶ BE (Best Effort) ; DSCP = 0 ; value in bit = 000000 ; definisce traffico senza priorità.
 - ▶ EF (Expedited Forwarding) ; DSCP = 46 ; value in bit = 101110 ; definisce priorità alta per flussi sensibili
 - ▶ AF (Assured Forwarding) ; DSCP = xxxyy0 value in bit ; stabilisce differenti classi di traffico per probabilità di scarto per flussi a banda garantita
 - ▶ AF_{xy} dove:
 - ▶ X = da 1 a 4 (il valore più alto è preferito) ; valore di classe
 - ▶ Y = da 1 a 3 (il valore più alto è preferito) ; valore di scarto (drop)
- ▶ Class Selector (CS)
 - ▶ Compatibile con il tag IP Precedence
 - ▶ Definisce 7 differenti classi
 - ▶ Il valore più alto è preferito



Tabella PHB value

| PHB | DSCP | DSCP | Drop Preced. | IP Precedence | Description | MPLS EXP |
|------|------|--------|--------------|---------------|---------------------------------------|----------|
| EF | 46 | 101110 | | 101 | CRITICAL | |
| AF43 | 38 | 100110 | HIGH | 100 | FLASH OVERRIDE | |
| AF42 | 36 | 100100 | MEDIUM | 100 | FLASH OVERRIDE | |
| AF41 | 34 | 100010 | LOW | 100 | FLASH OVERRIDE | |
| AF33 | 30 | 011110 | HIGH | 011 | FLASH | |
| AF32 | 28 | 011100 | MEDIUM | 011 | FLASH | |
| AF31 | 26 | 011010 | LOW | 011 | FLASH | |
| AF23 | 22 | 010110 | HIGH | 010 | IMMEDIATE | |
| AF22 | 20 | 010100 | MEDIUM | 010 | IMMEDIATE | |
| AF21 | 18 | 010010 | LOW | 010 | IMMEDIATE | |
| AF13 | 14 | 001110 | HIGH | 001 | PRIORITY | |
| AF12 | 12 | 001100 | MEDIUM | 001 | PRIORITY | |
| AF11 | 10 | 001010 | LOW | 001 | PRIORITY | |
| CS7 | 56 | 111000 | | 7 | NETWORK | 7 |
| CS6 | 48 | 110000 | | 6 | INTERNET | 6 |
| CS5 | 40 | 101000 | | 5 | CRITICAL (for Voice RTP) | 5 |
| CS4 | 32 | 100000 | | 4 | FLASH OVERRIDE | 4 |
| CS3 | 24 | 011000 | | 3 | FLASH (for Voice signaling and Video) | 3 |
| CS2 | 16 | 010000 | | 2 | IMMEDIATE | 2 |
| CS1 | 8 | 001000 | | 1 | PRIORITY | 1 |
| BE | 0 | 000000 | | 000 | ROUTINE (BEST EFFORT or DEFAULT) | 0 |

MPLS QoS LSP (DiffServ)

- ▶ MPLS INTSERV utilizza il protocolloRSVP per la reservation bandwidth end-to-end
- ▶ MPLS DIFFSERV introduce due nuovi tipi di LSP per caratteristiche ed operazioni:
 - ▶ **E-LSP (EXP-inferred-class):**
 - ▶ Supporta la trasmissione simultanea di più classi di traffico all'interno di un tunnel/path MPLS
 - ▶ All'interno di un singolo tunnel/path possiamo avere traffici classificati come EF + AF1 + AF2 + etc...
 - ▶ **L-LSP (Label-inferred-class):**
 - ▶ Trasporta una sola classe di traffico all'interno di un tunnel/path MPLS
 - ▶ Abbiamo differenti tunnel/path dove per ciascuno di essi sono trasportati classi EF, AFxy, etc...



QoS setting: Class-Map

▶ Classificazione del Traffico

- ▶ Vengono usate class-map command e si possono usare una o più match conditions
- ▶ Pacchetti che non rientrano nei criteri di nessuna classe sono parte della classe implicita di default che è referenziate come class-default

| | |
|---|--|
| class-map <name> | <i># identifica il nome di una class-map</i> |
| ! | |
| match access-group { value name ipv4 ipv6 value } | <i># numero o nome di una access-list</i> |
| match precedence <list> | <i># list of precedence in ipv4 and/or ipv6</i> |
| match dscp <list> | <i># list of DSCP in ipv4 and/or ipv6</i> |
| match mpls experimental topmost list | <i># list of EXP value in MPLS</i> |
| match packet length { min max value } | <i># IP packet size (including IP header)</i> |
| match protocol <protocol> | <i># list of protocols (ipv4 ipv6 TCP RTP etc...)</i> |
| match vlan <range> | <i># list of ethernet vlan-id</i> |
| match vlan inner <range> | <i># list of inner vlan-id for packet with double vlan encapsul.</i> |
| match cos <list> | <i># list of ethernet 802.1q user priority value</i> |
| Etc.... Etc.... | |



QoS setting: Policy-Map

▶ Colorazione del traffico

- ▶ Si basa sul valore di DSCP, IP Precedence o EXP impostato della trama IP packet e definita all'interno di una policy-map che riferenzia le class-map
- ▶ Il comando set supporta un'ampia gamma di layer 2 e layer 3 criteria

```
policy-map <name>           # identifica il nome di una policy-map (colorazione del traffico)
class <name>                 # riferenzia una specifica class-map (classificazione del traffico)
!
```



QoS setting: Policy-Map

▶ Marking Criteria for IP and MPLS Packets

| Syntax | Marking Criteria |
|--|--|
| set precedence <i>value</i> | IPv4 and IPv6 precedence |
| set precedence tunnel <i>value</i> | Precedence to be used by IP tunnel header |
| set dscp <i>value</i> | IPv4 and IPv4 DSCP |
| set dscp tunnel <i>value</i> | DSCP to be used by IP tunnel header |
| set mpls experimental imposition <i>value</i> | EXP bit to be used by push operation |
| set mpls experimental topmost <i>value</i> | EXP bit in MPLS header on top of the label stack |



QoS setting: Policy-Map

▶ Default MLS EXP Marking actions

| MPLS Forwarding Operation | Default Marking Action |
|---------------------------|--|
| push | Set MPLS EXP on all imposed label using marking in encapsulation header (MPLS EXP - IP Precedence - Ethernet 802.1q user priority) |
| swap | Maintain MPLS value |
| pop | Do not modify marking in exposed header |

▶ Default IP tunnel Marking action

| IP Tunnel Operation | Default Marking Action |
|----------------------|---|
| tunnel encapsulation | Set tunnel header DSCP using encapsulated DSCP for IP over GRE or encapsulated EXP for MPLS over GRE For L2TP, set DSCP to default value = 0 |
| tunnel decapsulation | Do not modify DSCP in exposed header |



QoS setting: Policy-Map

- ▶ Criteria for Marking of internal device field

| Syntax | Marking Criteria |
|---------------------------------------|--|
| set qos-group <i>value</i> | Internal field for packet class |
| set discard-class <i>value</i> | Internal field for packet drop profile |

- ▶ Criteria for Marking Ethernet, ATM, Frame Relay

| Syntax | Marking Criteria |
|-----------------------------|-------------------------------|
| set cos <i>value</i> | Ethernet 802.1q user priority |
| set atm-clp | ATM CLP bit |
| set fr-de | Frame Relay DE bit |



QoS setting: Service-Policy

- ▶ **Applicazione della policy-map su base interfaccia**
 - ▶ Una volta eseguita la classificazione (class-map) e la colorazione/marketing del traffico (policy-map); quest'ultima viene applicata alla interfaccia di ingresso di raccolta del traffico in ingresso oppure in uscita.

Interface ethernet 3/21

service-policy input / output <name policy-map>



QoS setting: Traffic Policing

▶ Politiche applicate alla QoS (Traffic Policing)

- ▶ **police** command ha un grande numero di opzioni e flessibilità
- ▶ Include sempre un profilo di traffico su base rate e burst (token buckets) ed un gruppo di azioni che possono essere implicitamente o esplicitamente specificate. di azioni
- ▶ CIR = la quantità/percentuale/velocità (RATE) di traffico trasmesso attraverso un link o virtual circuit (VC) misurato in bps (Committed Information Rate su base di un contratto SLA)
- ▶ Bc = numero di bit che può essere trasmesso (su base CIR) per ogni intervallo di tempo Tc; indica di quanto il CIR può essere superato o su base bit-rate o su base tempo
- ▶ Be = numero di bit in eccesso trasmessi (cioè oltre il numero di bit permessi del Bc) per un determinato periodo di inattività (Burst excess); indica di quanto il PIR può essere superato o su base bit-rate o su base di tempo
- ▶ Tc = è un intervallo di tempo settato per “modellare” la quantità di traffico richiesto su base SLA da configurare (Committed Rate Measurement Interval)
- ▶ PIR = Peak Information Rate ed è specificato come bit-rate oppure come percentuale di un link-rate



QoS setting: Traffic Policing

- ▶ SINGLE-RATE Policier (Traffic Profile) TCM

police *rate-value* [*bc-value* [*be-value*]]

police cir *value* [**bc-value** [**be-value**]]

police rate *value* [**burst** *value* [**peak-burst** *value*]]

police cir percent *value* [**bc** *value ms* [**be** *value ms*]]

police rate percent *value* [**burst** *value ms* [**peak-burst** *value ms*]]

Nota:

police cir e **policy rate** sono sintassi equivalenti;

rate e **burst** sono equivalenti a **cir** e **bc**



QoS setting: Traffic Policing

- ▶ DUAL RATE Policier (Traffic Profile) TCM

police cir *value [bc-value]* **pir** *value [be-value]*

police rate *value [burst value]* **peak-rate** [**peak-burst** *value*]

police cir percent *value [bc value ms]* **pir percent** *value [be value ms]*

police rate percent *value [burst value ms]* **peak-rate percent** *value [peak-burst value ms]*

Nota:

Come nel single-rate policier, il **cir** e **bc** definiscono il primo token bucket, nel dual-rate il **pir** ed il **be** definiscono il secondo token bucket.

L'equivalenza del **police cir** e **police rate** syntax è applicata anche nel dual-rate policier **peak-rate** e **peak-burst** sono equivalenti a **pir** e **be**



QoS setting: Traffic Policing

▶ Policier Action Type

| Syntax | Trigger Single-Rate | Trigger Dual-Rate |
|------------------------|---|--|
| conform -action | Quanto basta di tokens (frequenza costante in kbit/s tra un byte q e l'altro q) nel primo bucket (contenitore Q) abilitati a trasmettere in modo conforme ($q < Q$) | Quanto basta di tokens in entrambi i bucket ($q < Q$) |
| exceed -action | Quanto basta di tokens nel solo secondo bucket ($q < Q$) | Quanto basta di tokens nel solo secondo bucket ($q < Q$) |
| violate -action | Non quanto basta di tokens in entrambi i buckets ($q > Q$) | Non quanto basta di tokens in entrambi i buckets ($q > Q$) |

▶ Color Definition for color-aware policier

| Syntax | Description |
|---------------------------|---|
| conform-color name | Classe di traffico associata con il colore conforme |
| exceed-color name | Classe di traffico associata con il colore in eccesso |



QoS setting: Traffic Policing

► Policier Actions

| Syntax | Description |
|--|---|
| drop | drops packets |
| transmit | transmits packet without modification |
| set-prec-transmit <i>value</i> | IPv4 and IPv6 precedence |
| set precedence <i>value</i> | IPv4 and IPv6 precedence |
| set-prec-tunnel-transmit <i>value</i> | Precedence to be used by IP tunneling operation |
| set precedence tunnel <i>value</i> | Precedence to be used by IP tunneling operation |
| set-dscp-transmit <i>value</i> | IPv4 and IPv6 DSCP |
| set dscp <i>value</i> | IPv4 and IPv6 DSCP |
| set-dscp-tunnel-transmit <i>value</i> | DSCP to be use by IP tunneling operation |
| set dscp tunnel <i>value</i> | DSCP to be use by IP tunneling operation |
| set-mpls-exp-imposition-transmit <i>value</i> | EXP bits to be used by push operation |
| set mpls experimental imposition <i>value</i> | EXP bits to be used by push operation |
| set-mpls-exp-topmost <i>value</i> | EXP bits in MPLS header on top of the stack label |
| set mpls experimental topmost <i>value</i> | EXP bits in MPLS header on top of the stack label |



QoS setting: Traffic Policing

► Policer Actions (segue ./.)

| Syntax | Description |
|--|--|
| set-qos-transmit <i>value</i> | internal field for packet class |
| set qos-group <i>value</i> | internal field for packet class |
| set-discard-class-transmit <i>value</i> | internal field for packet drop profile |
| set discard-class <i>value</i> | internal field for packet drop profile |
| set-cos-transmit <i>value</i> | Ethernet 802.1q user priority |
| set cos <i>value</i> | Ethernet 802.1q user priority |
| set-clp-transmit | ATM CLP bit |
| set atm-clp | ATM CLP bit |
| set-frde-transmit | Frame Relay DE bit |
| set fr-de | Frame Relay DE bit |



QoS setting: Traffic Shaping

- ▶ Con il comando **shape** è possibile configurare un valore massimo di bandwidth per una determinata classe di traffico; in caso di eccesso (oltre il profilo accordato) il traffico verrà accodato per essere trasmesso.

| Syntax | Description |
|--|---|
| shape average <i>rate-value</i> [burst] | Definisce un valore di media shaper bandwidth da rispettare all'interno di un intervallo di tempo fisso |
| shape average <i>rate-value</i> [bc-value [be-value]] | Come sopra ma con un intervallo di tempo configurabile |
| shape peak <i>rate-value</i> [bc-value [be-value]] | Definisce un valore di shaper picco da rispettare all'interno di un intervallo di tempo configurabile |
| shape average percent <i>rate value</i> [burst] ms | Definisce un valore di media shaper percentuale in bandwidth ed un intervallo di tempo fisso |
| shape average percent <i>rate-value</i> [bc-value ms [be-value ms]] | Definisce un valore di media shaper BW ed un intervallo di tempo configurabile |
| shape peak percent <i>rate-value</i> [bc-value ms [be-value ms]] | Definisce un valore di media shaper percentuale in bandwidth ed un intervallo di tempo configurabile |
| shape max-buffer <i>value</i> | Innalza il numero massimo di buffers dal valore di default (1000) per evitare perdita di pacchetti. |



QoS setting: Traffic Shaping Formula

- ▶ Se T_c è lo shaper interval, abbiamo le seguenti relazioni:

$$T_c = B_c / CIR$$

$$PIR = (B_c + B_e) / T_c$$

$$PIR = CIR + CIR (B_e / B_c)$$

$$PIR = CIR + EIR$$

$$EIR = CIR (B_e / B_c)$$

EIR significa Excess Information Rate

PIR significa Peak Information Rate



QoS design example

