



**La Sapienza**

Università degli Studi di Roma

Dipartimento di Informatica e Sistemistica

# RETI DI CALCOLATORI II

## QoS in reti IP

**Emiliano Trevisani**

**[trevisani@dis.uniroma1.it](mailto:trevisani@dis.uniroma1.it)**

**A.A. 2008/2009**

## Overview 1/5

- ❑ Il classico servizio di trasferimento offerto da Internet è
  - punto-punto, connectionless, best effort
- ❑ Il servizio non è adatto alle applicazioni con requisiti di trasferimento real time a causa di:
  - throughput non garantito
  - ritardo variabile di attraversamento della rete
  - perdite di pacchetti dovute a congestione
- ❑ Inoltre:
  - Non sono definite procedure di admission control e di policing
  - Le sotto-reti sono fortemente eterogenee in termini di capacità

Introdurre QoS nel mondo IP: l'idea base è quella di re-introdurre una sorta di "connessione" a livello rete; tuttavia, la natura connection less / best effort del servizio non consente di controllare le risorse di rete in modo semplice

## Overview 2/5

- ❑ La QoS descrive il livello di prestazione che deve essere assicurato per una particolare applicazione
- ❑ La QoS può essere definita in modo
  - **Assoluto (Performance Assurance)**
    - definendo i valori che devono essere rispettati da un insieme di parametri prestazionali (es. ritardo massimo, probabilità di perdita di pacchetti, ...)
  - **Relativo (Service Differentiation)**
    - definendo le modalità di trattamento di una classe di traffico rispetto alle altre (es. livello di priorità di servizio, livello di priorità di scarto, ...)
- L'insieme di regole, assolute e/o relative, che definiscono un livello di QoS è detto Service Level Agreement (SLA)

## Overview 3/5

- Una rete è in grado di **garantire** un prefissato livello di QoS attraverso un duplice approccio
  - **Overprovisioning**
    - le risorse di rete sono dimensionate in modo che, nelle condizioni peggiori, il carico sia inferiore alla soglia minima necessaria a soddisfare gli SLA concordati con gli utenti
  - **Admission Control**
    - Il traffico è controllato preventivamente, sarà accettato solo se le risorse di rete saranno sufficienti a garantire i livelli di qualità richiesti dagli utenti [oltre che da quello corrente]

## Overview 4/5

- ❑ La QoS può essere garantita
  - per flusso, per connessione, per chiamata
  - per aggregati (o classi) di flussi, di connessioni, di chiamate
- ❑ Un **flusso** è definito dalla 5-pla:
  - Source IP address
  - Source port number
  - Destination IP address
  - Destination port number
  - Protocol
- ❑ Una classe o aggregato di traffico comprende un insieme di flussi aventi requisiti simili di QoS

## Overview 5/5

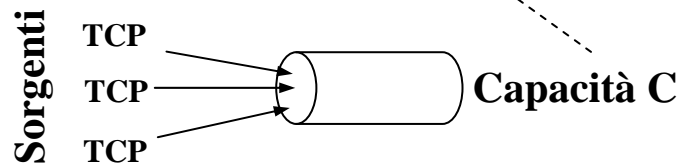
- Una politica di QoS orientata al trattamento dei singoli flussi (**per-flow QoS**)
  - Assicura prestazioni migliori sia per quanto riguarda i singoli flussi che per quanto riguarda l'utilizzazione delle risorse di rete
  - Ha una notevole complessità dovuta ai meccanismi di segnalazione e alla necessità di monitorare i singoli flussi di pacchetti in rete
  - Bassa scalabilità
- Una politica di QoS orientata al trattamento degli aggregati di flussi (**aggregate traffic QoS**)
  - Fornisce prestazioni non ottimali
  - Ha complessità minore
  - È maggiormente scalabile

## Internet traffic 1/2

### □ Elastic Traffic

Traffico con requisiti d'integrità ma non di ritardo / throughput

- traffico tradizionale delle reti TCP/IP
- è indifferente alle variazioni di ritardo di transito dei pacchetti
- Sono normalmente utilizzati i protocolli TCP e UDP
  - TCP: la capacità utilizzata è il massimo tra la banda disponibile e il rate corrente stabilito dai meccanismi di controllo di flusso / congestione
  - UDP: la capacità utilizzata è il massimo tra la banda disponibile e il rate di emissione della sorgente
- la qualità di servizio percepita dall'utente si riferisce al tempo necessario al trasferimento di un intero elemento informativo tipico dell'applicazione



Il comportamento delle sorgenti TCP è di tipo aggressivo nel senso che ciascuna tende ad utilizzare tutta la banda disponibile; poiché tutte le sorgenti si comportano in modo analogo la distribuzione della capacità tende ad essere equa e non differenziata per tipologia di traffico

## Internet traffic 2/2

### □ Real Time Traffic

- la qualità è fortemente dipendente dal ritardo di transito in rete
- Parametri di QoS:
  - Throughput: è richiesto una banda minima garantita dalla rete
  - Delay: è richiesto un ritardo limite garantito dalla rete
  - Jitter: occorre stabilire il valore massimo della variazione del ritardo
  - Perdita di pacchetti: è richiesta la garanzia di un valore massimo di perdita di pacchetti
- Il trattamento del traffico real time necessita di
  - trattamento preferenziale rispetto al traffico elastico
  - procedure di riservazione di risorse

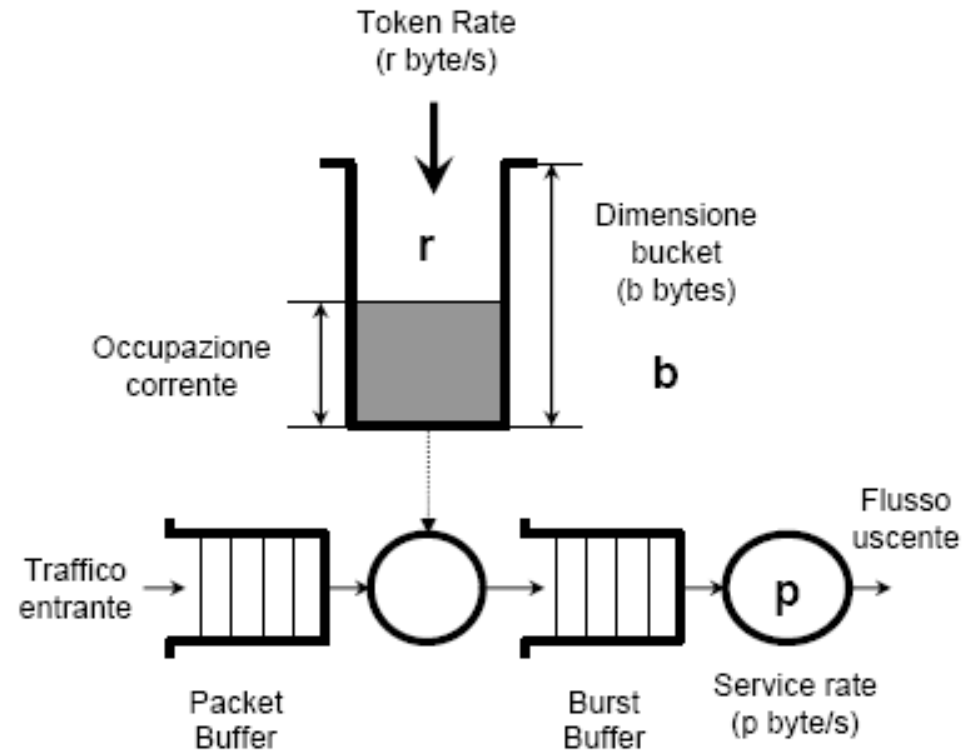
Non ci si può basare solo sull'aggressività delle sorgenti; occorre che la rete riservi risorse preventivamente per garantire QoS

## Caratterizzazione del traffico

- ❑ I parametri caratteristici di un flusso o di un aggregato di flussi sono riassunti nel corrispondente **TSpec (Transmission Specification)**
- ❑ I parametri di servizio richiesti sono definiti nel **RSpec (Request Specification)**
- ❑ I parametri del TSpec sono specificati con riferimento al modello **Token Bucket**
- ❑ Modello **Token Bucket**:
  - Fornisce una descrizione concisa del carico prodotto da un flusso [modello di emissione di una sorgente]
  - Fornisce i parametri per l'algoritmo di policing ed è utilizzato anche come formatore (**shaper**) di traffico
  - È caratterizzato da 5 parametri:
    - **token rate (r)**
    - **maximum service rate (p) ( $p > r$ )**
    - **bucket size (b)**
    - **minimum policed unit (m)**
    - **maximum datagram size (M)**

## Modello Token Bucket 1/5

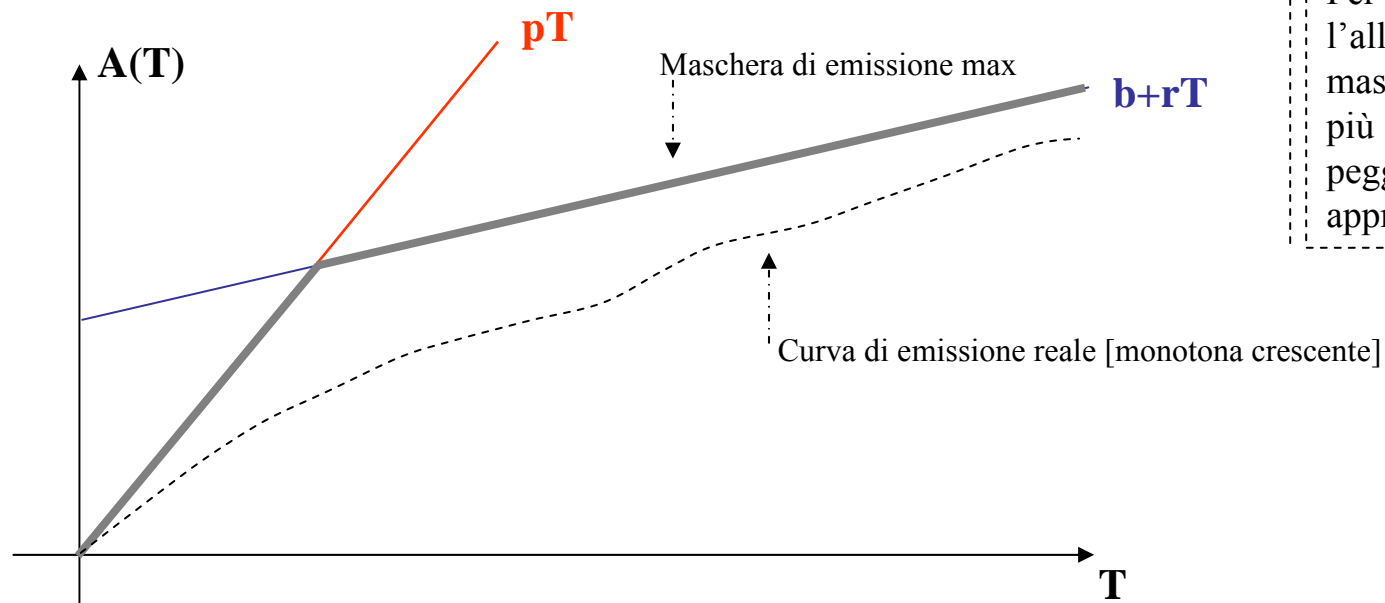
- ❑ Il token rate  $r$  specifica il rate a lungo termine (**average rate**)
- ❑ Il service rate  $p$  specifica il rate massimo (**peak rate**)
- ❑ Il buffer size  $b$  rappresenta la massima lunghezza di burst concessa alla sorgente [grado di **burstiness** concesso alla sorgente]
- ❑ I datagrammi di lunghezza inferiore a  $m$  sono considerati di lunghezza uguale a  $m$  byte
- ❑ Un datagramma di dimensione superiore a  $M$  byte è scartato



## Modello Token Bucket 2/5

- Dato un Token Bucket  $(r, b, p)$ , il numero massimo di bit  $A(T)$  che possono essere emessi da una sorgente in un intervallo  $(t, t+T]$  è dato da:

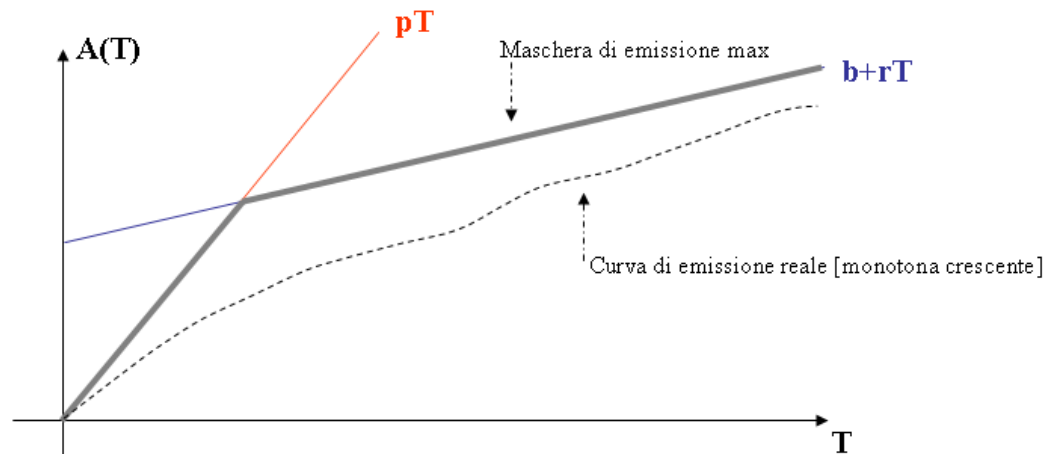
$$A(T) = \min [ p \cdot T, b + r \cdot T ] \leq b + r \cdot T$$



Per semplificare l'allocazione uso una maschera di emissione più semplice (caso peggiore)  $\Rightarrow$  approssimazione

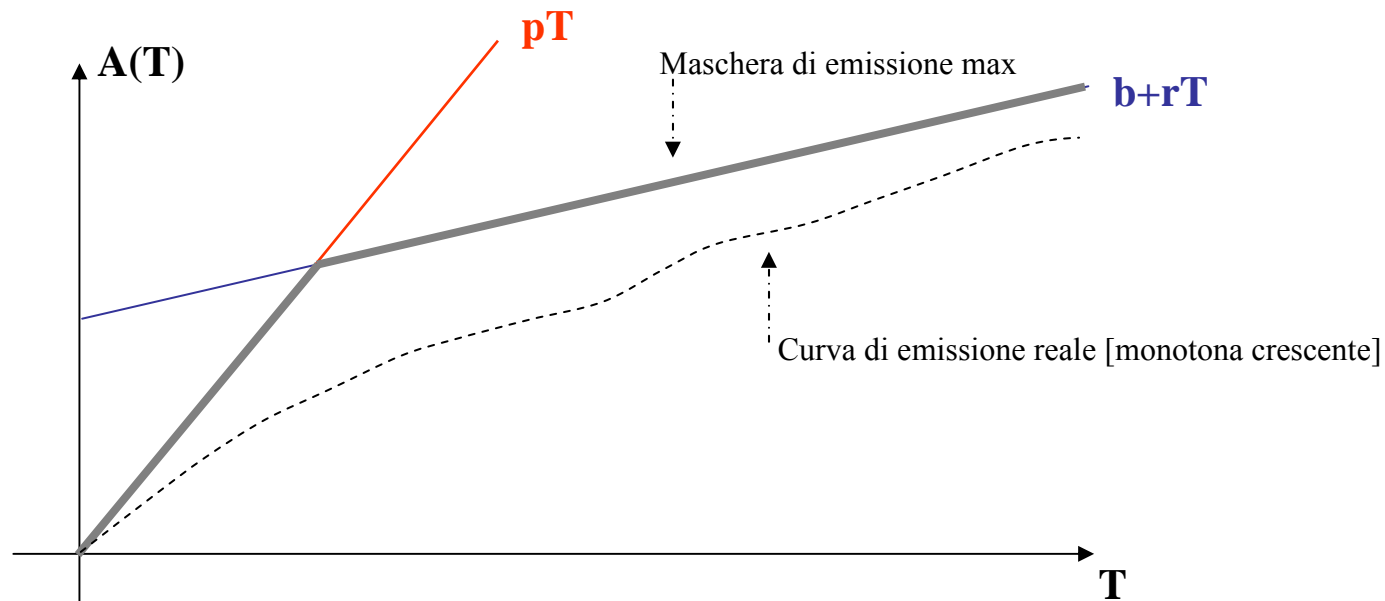
## Modello Token Bucket 3/5

- Al tempo 0 il bucket è pieno [ $\Leftrightarrow$  b autorizzazioni a trasmettere]
  - il caso più favorevole è quello per il quale la sorgente emette un burst di dati che il server è autorizzato a trasmettere al rate di picco p
  - successivamente il server può emettere al più al ritmo di generazione dei token
  
- Notare le funzioni di sagomatura e controllo del traffico:
  - La sagomatura di bordo indica il limite di emissione della sorgente; naturalmente questa può operare al di sotto [è autorizzata a trasmettere ma non ci sono dati da trasmettere]



## Modello Token Bucket 4/5

- Il modello Token Bucket introduce 2 livelli d'inefficienza:
  - I meccanismi di allocazione delle risorse tengono conto della maschera di emissione max e non della curva di emissione reale
  - La maschera di emissione max “semplificata” fa riferimento ad una condizione ancora peggiore

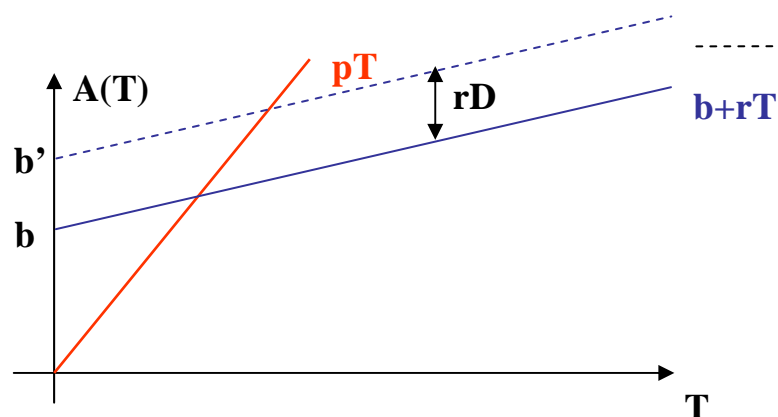
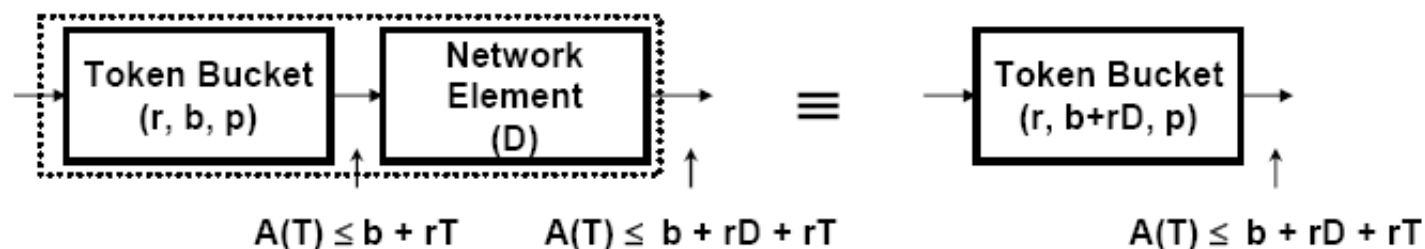


## Modello Token Bucket 5/5

### □ Teorema:

Vincolando le sorgenti ad emettere secondo il modello Token Bucket siamo in grado di capire come degrada il traffico nel transito nei nodi di rete; in particolare, attraversare un nodo con ritardo max  $D$ , peggiora la maschera di emissione con la quale dovrò fare allocazione

- Un flusso compreso nel profilo uscente di un token bucket  $(r, b, p)$  che transita attraverso un elemento di rete che garantisce un ritardo massimo uguale a  $D$ , risulta essere compreso nel profilo uscente da un token bucket di parametri  $(r, b+rD, p)$



La rete garantirà la banda necessaria ed una probabilità di perdita nulla per saturazione di buffer; l'allocazione è eseguita sulla base del caso peggiore [maschera max]. Notare come ogni nodo di rete attraversato sovradimensioni l'allocazione

## Integrated Service Architecture (Intserv) 1/2

- ❑ Fornisce supporto esplicito ai servizi a qualità garantita in una rete IP
- ❑ E' un'estensione del modello classico IP
  - Le nuove componenti sono aggiuntive rispetto a quelle comprese nel modello IP classico e non le sostituiscono
- ❑ Il modello **Intserv** definisce le tre classi di servizio:
  - **Best effort service**
  - **Guaranteed service** [Servizio a qualità garantita]
  - **Controlled load** [Servizio a carico controllato]

## Integrated Service Architecture (Intserv) 2/2

### ❑ Utilizza procedure di **reservation**

- procedure di set-up esplicite
- resource reservation
- admission control
- policing

Come faccio a differenziare il traffico in un ambiente connectionless per applicare i criteri di QoS? Introduco il concetto di flusso di datagrammi generati dalla stessa sorgente [processo] e con lo stesso livello di QoS associato

### ❑ Utilizza un approccio per flusso (**per flow - QoS**)

### ❑ Definisce dei flow state nei router e non solo negli host

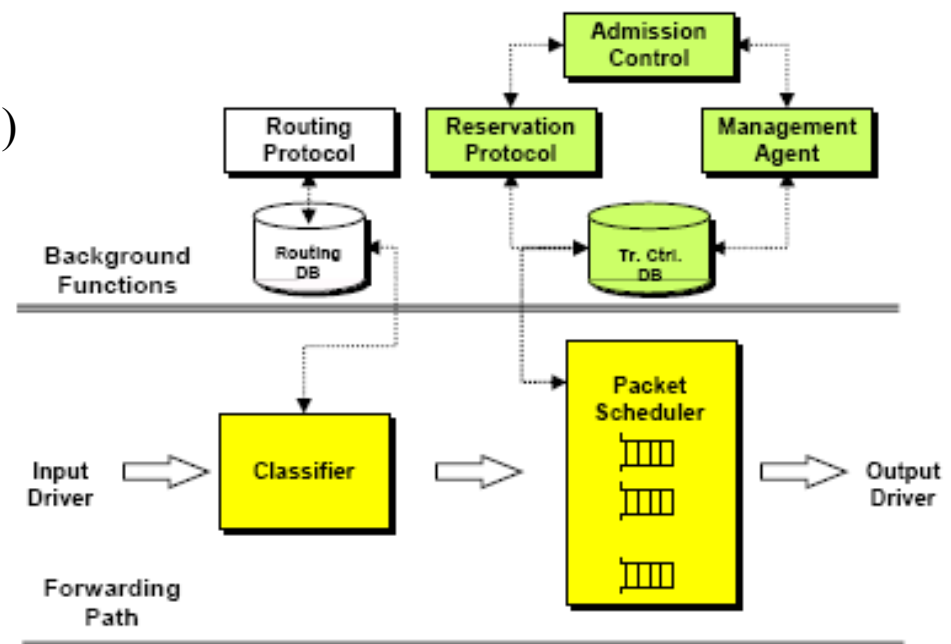
### ❑ Utilizza stati di tipo soft nei router

- robustezza e semplicità

## Modello di router Intserv 1/3

□ Il controllo del traffico in un router Intserv è realizzato mediante quattro componenti logiche:

- Packet Scheduler
- Classifier
- Admission Control
- Reservation Protocol (RSVP)



## Modello di router Intserv 2/3

### □ Packet Scheduler

L'idea è quella di passare da un buffer fisico gestito con politica FIFO [router tradizionali] ad un insieme di buffer logici ritagliati nel buffer fisico in modo che gli algoritmi di scheduling possano garantire livelli di QoS distinti per ciascun buffer logico

- stabilisce l'ordine di rilancio e l'eventuale scarto dei datagrammi dalla coda di uscita
- utilizza un insieme di code e altri meccanismi (es. timer)
- esegue funzionalità di policing

### □ Classifier

Blocco logico che individua il flusso al quale appartiene il datagramma; opera packet by packet

- individua il flusso a cui appartiene ogni datagramma entrante e determina la classe di servizio con cui verrà trattato dal packet scheduler
- la classe di servizio può essere definita localmente e essere diversa in diversi router

### □ Admission Control

All'uscita del classifier il piano di controllo indica quale buffer logico utilizzare in funzione della QoS da raggiungere

- implementa l'algoritmo di decisione per l'accettazione di un nuovo flusso

## Modello di router Intserv 3/3

### ❑ ReSerVation Protocol (RSVP)

Sul piano di controllo devo aggiungere segnalazione per la riservazione delle risorse

- crea e mantiene lo stato riguardante un flusso negli host e nei router
  - trasporta le specifiche del flusso (TSpec) e le richieste di servizio (RSpec)
  - Un TSpec è una lista di parametri che definiscono le caratteristiche di un flusso
  - Un RSpec è una lista di parametri che specifica la qualità di servizio richiesta
- ❑ Tra sorgente e rete viene stabilito un **contratto di traffico**
- la sorgente comunica alla rete il TSpec e l'RSpec
  - la rete esegue l'algoritmo di admission control e comunica il risultato alla sorgente
  - la rete provvede garantire durante l'emissione della sorgente la QoS concordata
  - la rete verifica che i parametri del TSpec siano rispettati dalla sorgente
  - eventuali pacchetti che non rispettano il TSpec sono trattati come traffico best effort

## Guaranteed Service 1/3

- E' orientato al traffico real time
  - Garantisce un livello minimo di throughput
  - Garantisce l'assenza di perdita di pacchetti per overflow dei buffer dei router
  - Garantisce un ritardo massimo di transito in rete
- Il ritardo massimo ( $D_{Max}$ ) di transito in rete è dato da due componenti:
  - $D_{fix}$  è la componente fissa del ritardo (trasmissione, propagazione)
  - $D_{Max\_que}$  è il ritardo massimo complessivo per attese in coda

$$D_{Max} = D_{fix} + D_{Max\_que}$$

- È garantito un valore massimo non superato di  $D_{Max\_que}$

## Guaranteed Service 2/3

- Nell'invocazione del servizio, la sorgente deve specificare il TSpec e l'RSpec
- I parametri del TSpec sono:
  - $r$             byte/s     token rate
  - $b$             byte        buffer size
  - $p$             byte/s     peak rate ( $p \geq r$ )
  - $M$             byte        Maximum datagram size
  - $m$             byte        minimum policed unit
- I parametri dell'RSpec sono:
  - $R$             byte/s     requested rate ( $R \geq r$ )
  - $S$              $\mu\text{s}$         termine di correzione (slack term)

## Guaranteed Service 3/3

- ❑ Il termine di correzione  $S$  indica la differenza tra il ritardo desiderato dalla sorgente e ritardo ottenuto tramite l'assegnazione della banda  $R$
- ❑ È utilizzato dagli elementi di rete per ridurre l'assegnazione di banda al flusso
- ❑ In un elemento di rete vengono assegnate ad un flusso
  - una banda  $R$  (byte/s) (richiesta dalla sorgente)
  - un buffer di dimensione  $B$  (byte)
- ❑ Ogni elemento di rete, in base ai parametri di  $TSpec$  e  $RSpec$ , calcola il valore di  $B$  in modo che sia evitata la perdita di datagrammi
- ❑ Il livello di servizio è deterministicamente determinato ma ha inefficienza a causa del sovradimensionamento di risorse

## Controlled Load 1/2

- ❑ E' orientato al trattamento di traffico real-time di tipo adattativo (sono tollerati picchi di ritardo)
- ❑ Idea: recuperare efficienza rispetto a Guaranteed Service rinunciando alle garanzie
- ❑ Questo servizio non prevede esplicitamente parametri di qualità e regole di allocazione; l'implementazione è lasciata al provider
  - La sorgente specifica TSpec e non RSpec [il provider deciderà quante risorse assegnare (banda e buffer) ]
- ❑ La QoS approssima quella percepita dall'utente che utilizza un servizio di tipo best effort nel caso di rete scarica
- ❑ Il ritardo di coda non è limitato superiormente
  - Un'elevata percentuale di pacchetti non subirà un ritardo molto superiore a quello minimo di attraversamento della rete

## Controlled Load 2/2

- ❑ Una percentuale molto elevata di pacchetti sarà consegnata a destinazione
  - la percentuale di pacchetti persi deve essere dell'ordine del packet error rate dovuto ai mezzi trasmissivi
- ❑ Guaranteed Service vs. Controlled Load:

<b>Caratteristica</b>	<b>Guaranteed Service</b>	<b>Controlled Load</b>
<b>Servizio end-to-end</b>	Ritardo Max garantito Prob. di perdita nulla	Equivalente ad un servizio Best effort su rete scarica
<b>Applicazioni</b>	Real Time	Sensibili alla congestione in rete

## Differentiated Service Architecture (Diffserv) 1/2

- L'approccio Intserv è difficilmente applicabile a reti in cui il numero di flussi per link è molto elevato
  - reti di grandi dimensioni
  - link ad elevata velocità (Gbit/s)
- La limitata scalabilità deriva dalla scelta di gestire il singolo flusso di pacchetti (per flow approach)
  - User plane
    - necessità delle funzioni di classifier, policer e scheduler in ogni router
  - Control plane
    - trattamento dei messaggi di segnalazione (RSVP) in ogni router
    - memorizzazione e gestione degli stati di riservazione per singolo flusso (path reservation state) in ogni router

## Differentiated Service Architecture (Diffserv) 2/2

- L'approccio Diffserv offre una maggiore scalabilità utilizzando i seguenti accorgimenti:
  - Le funzioni più complesse (policing e classificazione) sono eseguite solo nei router di bordo della rete
  - Ad ogni datagramma entrante nella rete è assegnata una classe di servizio (aggregato di flussi)
  - Ad ogni datagramma entrante è assegnato un codice (DS codepoint) che indica esplicitamente la sua classe di servizio
  - Nei router interni della rete sono definite regole di trattamento dei datagrammi (Per Hop Behaviour - PHB) valide per ogni singola classe di servizio
  - Sono eliminati i reservation state nei router interni alla rete
  - Almeno in una prima fase, i servizi di trasferimento sono offerti su base contrattuale o semi-permanente (provisioning) e quindi non è necessario l'uso di un protocollo di segnalazione

## IPv4: Service Type field

### Bit 0-2: Precedence

111 - Network Control	011 - Flash
110 - Internetwork Control	010 - Immediate
101 - CRITIC/ECP	001 - Priority
100 - Flash Override	000 - Routine

### Bit 3-5: TOS

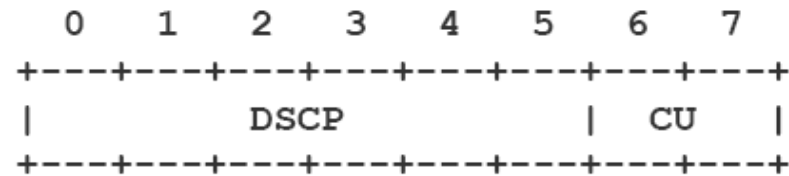
Bit 3: 0 = Normal Delay,	1 = Low Delay.
Bits 4: 0 = Normal Throughput	1 = High Throughput
Bits 5: 0 = Normal Reliability	1 = High Reliability

### Bit 6-7: Reserved for Future Use

0	1	2	3	4	5	6	7						
+-----+-----+-----+-----+-----+-----+-----+-----+													
	PRECEDENCE			D		T		R		0		0	
+-----+-----+-----+-----+-----+-----+-----+-----+													

## Diff Serv [DS] field

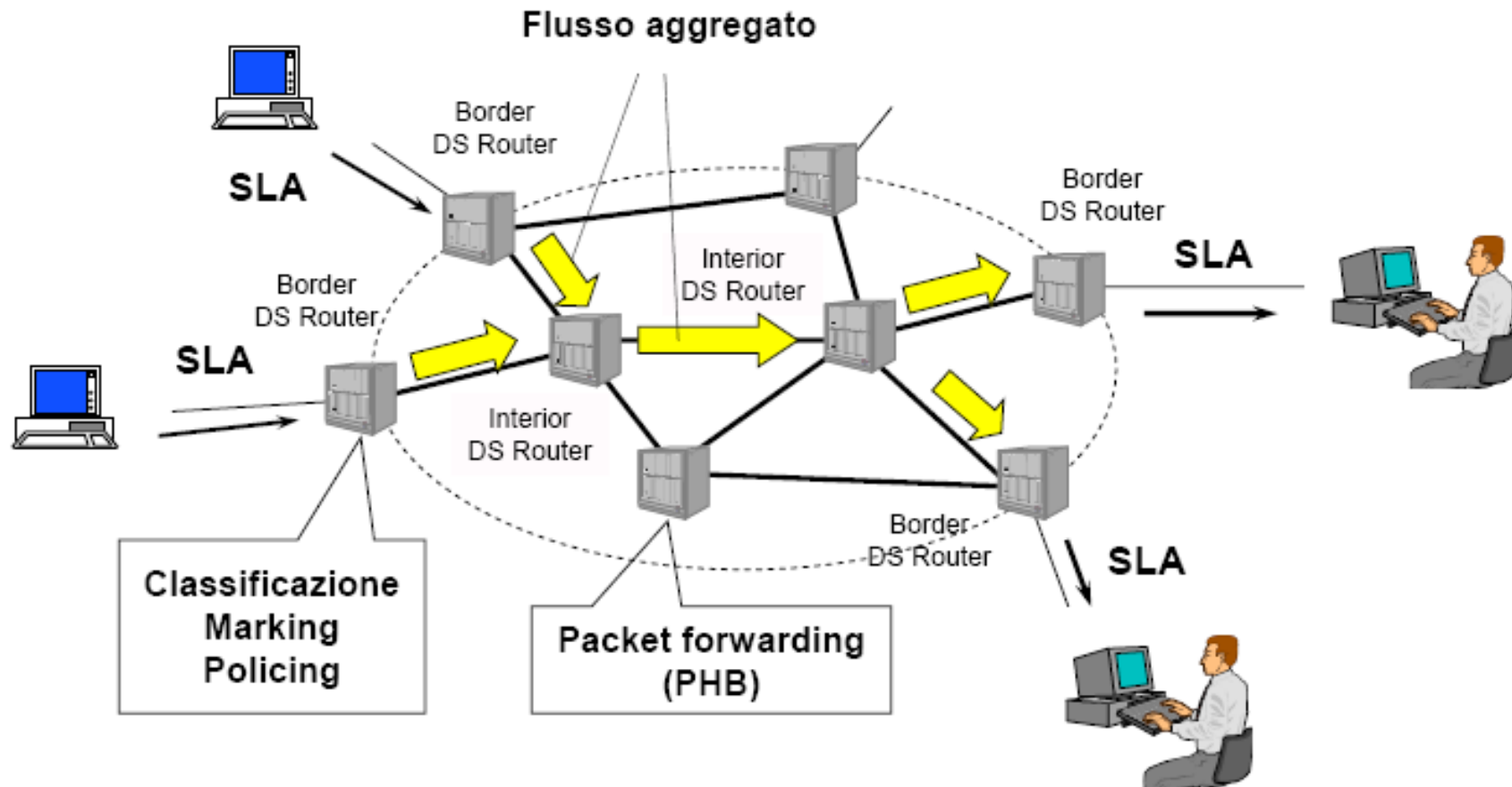
- ❑ Il campo DS sostituisce il campo Service Type nel protocollo IPv4
- ❑ La struttura del campo DS è la seguente



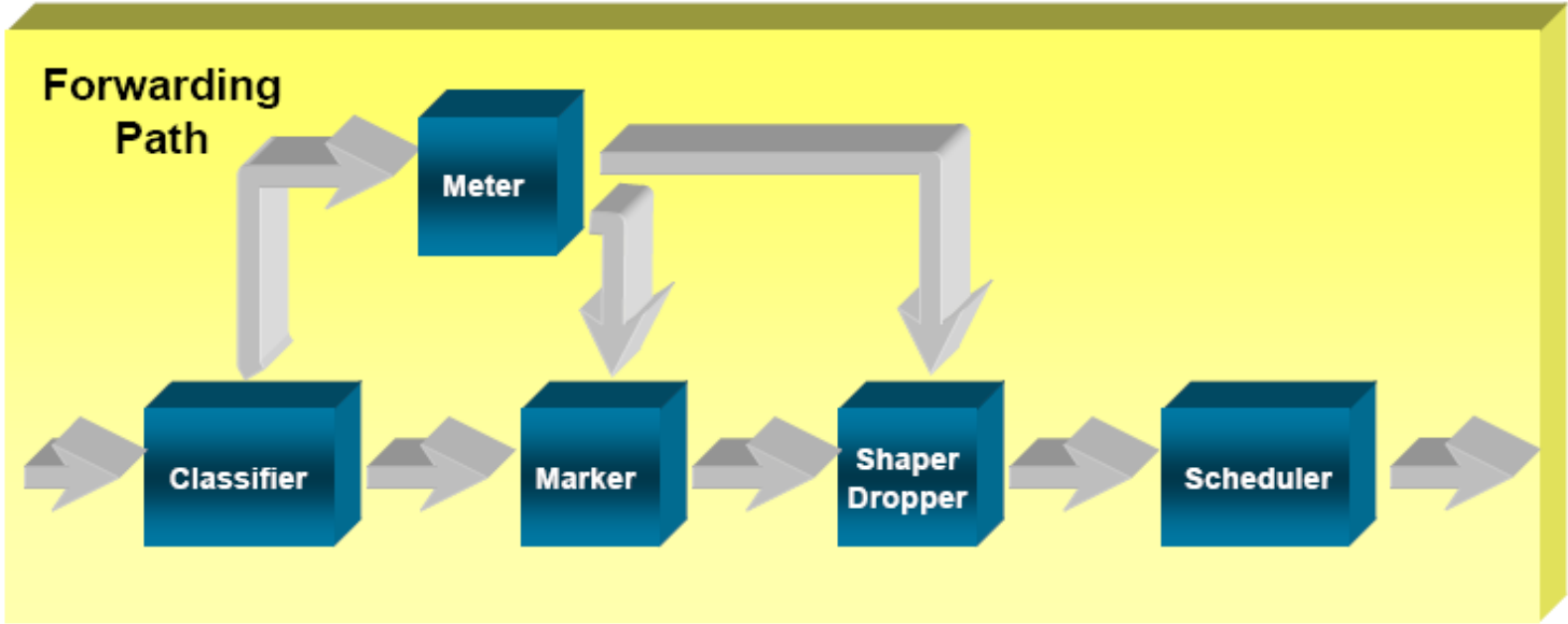
DSCP: differentiated services codepoint  
CU: currently unused

- ❑ Il campo DSCP è usato come indice per determinare il trattamento del pacchetto in un router

# Dominio Diffserv



# Router Diffserv 1/2



## Router Diffserv 2/2

### □ Classifier

- identifica il flusso a cui appartiene il datagramma entrante
- identifica il tipo di trattamento del datagramma

### □ Meter

- misura i parametri temporali del flusso
- effettua il confronto con il profilo di traffico stabilito dallo SLA

### □ Marker

- assegna il valore del campo DSCP al datagramma
- assegna il datagramma ad un particolare flusso aggregato

### ▪ Shaper/Dropper

- sagoma il flusso di datagrammi in accordo al profilo di traffico concordato per il flusso aggregato
- scarta i datagrammi che sono fuori dal profilo di traffico concordato per il flusso aggregato

### □ Scheduler

- determina l'ordine di trasmissione dei datagrammi in accordo alle regole di trattamento del particolare flusso aggregato a cui appartiene il datagramma

## Per Hop Behaviour (PHB) 1/2

- ❑ Un PHB stabilisce le regole con cui devono essere trattati i datagrammi nei router di un dominio DS
- ❑ È definito un insieme di PHB che corrispondono ai diversi livelli di QoS che un dominio DS può offrire
- ❑ Un livello di QoS offerto da un dominio DS è dato dalla composizione dei PHB applicati nei router attraversati dai datagrammi
- ❑ Oggi sono definiti tre PHB
  - **Best effort**
  - **Expedited Forwarding (EF)**
  - **Assured Forwarding (AF)**

## Per Hop Behaviour (PHB) 2/2

- ❑ **Expedited Forwarding (EF)**: utilizzato per costruire un servizio caratterizzato da bassa probabilità di perdita, basso ritardo, bassa variabilità del ritardo e banda garantita
  - Un PHB EF deve assicurare che i datagrammi non subiscano accodamenti eccessivi nei buffer dei router
  
- ❑ **Assured Forwarding (AF)**: offre la possibilità ad un operatore di rete di definire un insieme di livelli di trasferimento (Forwarding Assurances) per differenziare il trattamento dei flussi in un dominio IP
  - Sono definite quattro classi AF ( $AF_{xy}$ ,  $x=1,2,3,4$ ); per ognuna di esse è allocato in ogni router un diverso insieme di risorse (banda e buffer)
  - Per ciascuna delle quattro classi sono definiti tre livelli di priorità di scarto dei datagrammi (drop precedence) ( $AF_{xy}$ ,  $y=1,2,3$ )
    - in caso di congestione i datagrammi sono scartati nell'ordine stabilito dal livello di drop precedence