## Calidad de Servicio

Redes de Ordenadores para Robots y Máquinas Inteligentes

 $\mathsf{GSyC}$ 

Departamento de Teoría de la Señal y Comunicaciones y Sistemas Telemáticos y Computación

Febrero 2024



©2024 Grupo de Sistemas y Comunicaciones.
Algunos derechos reservados.
Este trabajo se distribuye bajo la licencia
Creative Commons Attribution Share-Alike 4.0
disponible en http://creativecommons.org/licenses/by-sa/4.0/deed.es

- Aplicaciones multimedia en redes
- 2 Evolución Calidad de servicio
- 3 Técnicas para proporcionar múltiples clases de tráfico
- Mecanismos de planificación
- 5 Mecanismos de vigilancia (policing & shaping)
- 6 Referencias

- Aplicaciones multimedia en redes
- 2 Evolución Calidad de servicio
- 3 Técnicas para proporcionar múltiples clases de tráfico
- 4 Mecanismos de planificación
- Mecanismos de vigilancia (policing & shaping)
- 6 Referencias

## Aplicaciones multimedia en redes

#### Clases de aplicaciones multimedia en redes

- Flujos de audio/vídeo almacenados
- Plujos de audio/vídeo en vivo
- 3 Audio/vídeo interactivo en tiempo real

#### Características fundamentales

- Habitualmente son sensibles a los retardos
  - retardo de propagación terminal a terminal (end-to-end)
  - Variación/fluctuación del retardo de los paquetes en un mismo flujo (packet jitter)
- Toleran pérdidas de datos ocasionales: pérdidas poco frecuentes no causan grandes inconvenientes
- Justo lo contrario de las aplicaciones elásticas de transmisión de datos tradicionales, que no toleran las pérdidas (precisan completitud e integridad) pero no les afectan gravemente los retardos de propagación

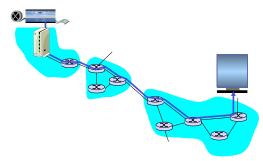
- Aplicaciones multimedia en redes
  - 1. Flujos de audio/vídeo almacenados
  - 2. Flujos de audio/vídeo en vivo
  - 3. Flujos de audio/vídeo interactivo en tiempo real

- - Planificación FIFO
  - Colas con prioridad
  - Colas de turno rotatorio
  - Colas equitativas ponderadas (WFQ)
- - Token bucket.
  - Token bucket +WFQ

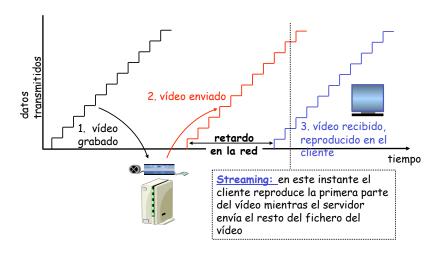
GS<sub>v</sub>C - 2024 Calidad de Servicio

## 1. Flujos de audio/vídeo almacenado

- Audio/vídeo almacenado en un servidor
- Se transmite al cliente
- Transmisión de flujos (streaming): el cliente comienza a reproducir antes de que todos los bits del fichero hayan llegado
  - El único requisito temporal para los datos aún no transmitidos es que lleguen a tiempo para cuando tengan que ser reproducidos

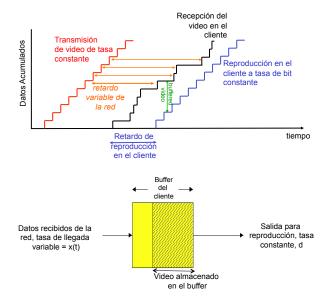


## Flujos de audio/vídeo almacenado



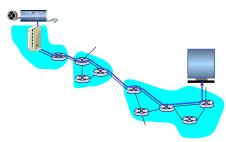
### Buffer en el cliente

 En el cliente, el almacenamiento de los datos en un buffer previo a la reproducción permite controlar los retardos variables de la red



## Interactividad con los flujos de audio/vídeo almacenados

- Funcionalidad tipo vídeo con operaciones interactivas como pausa, rebobinado, avance rápido
  - Es aceptable un retraso de hasta 10s antes de comenzar la reproducción
  - Es aceptable un retraso de 1s o 2s para las operaciones interactivas
- El único requisito temporal para los datos aún no transmitidos es que lleguen a tiempo para cuando tengan que ser reproducidos



Calidad de Servicio GS<sub>v</sub>C - 2024

# Interactividad con los flujos de audio/vídeo almacenados (II)



## Aplicaciones multimedia en redes

- 1. Flujos de audio/vídeo almacenados
- 2. Flujos de audio/vídeo en vivo
- 3. Flujos de audio/vídeo interactivo en tiempo real
- 2 Evolución Calidad de servicio
- Técnicas para proporcionar múltiples clases de tráfico
- 4 Mecanismos de planificación
  - Planificación FIFO
  - Colas con prioridad
  - Colas de turno rotatorio
  - Colas equitativas ponderadas (WFQ)
- Mecanismos de vigilancia (policing & shaping)
  - Token bucket
  - Token bucket +WFQ
  - Referencias

## 2. Flujos de audio/vídeo en vivo

#### Ejemplos:

- Emisoras de tv/radio por Internet
- Transmisión de eventos en tiempo real por Internet

#### Es similar a los flujos de audio/vídeo almacenado

- Buffer de reproducción
- La reproducción puede comenzar hasta varias decenas de segundos después de su transmisión, lo que para eventos en directo puede llegar a ser un problema para el usuario.
- Tienen los mismos requisitos temporales.

#### Interactividad con el usuario

- Imposible el avance rápido
- Permiten el rebobinado y la pausa

- Aplicaciones multimedia en redes
  - 1. Flujos de audio/vídeo almacenados
  - 2. Flujos de audio/vídeo en vivo
  - 3. Flujos de audio/vídeo interactivo en tiempo real

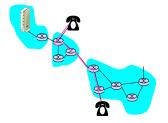
- - Planificación FIFO
  - Colas con prioridad
  - Colas de turno rotatorio
  - Colas equitativas ponderadas (WFQ)
- - Token bucket.
  - Token bucket +WFQ

GS<sub>v</sub>C - 2024 Calidad de Servicio

## 2. Audio/Vídeo interactivo en tiempo real

#### Aplicaciones:

Telefonía IP (audioconferencia), vídeo interactivo en tiempo real (videoconferencia), mundos interactivos



#### Requisitos de retardo terminal a terminal:

- Audio: ideal si es menor de 150ms, aceptable si es menor de 400 ms
  - El retardo terminal a terminal contempla el retardo en la red y en los terminales (paquetización)
  - Los retardos mayores impiden el uso de estas aplicaciones

GS<sub>v</sub>C - 2024 Calidad de Servicio

- Aplicaciones multimedia en redes
- 2 Evolución Calidad de servicio
- 3 Técnicas para proporcionar múltiples clases de tráfico
- 4 Mecanismos de planificación
- 5 Mecanismos de vigilancia (policing & shaping)
- 6 Referencias

## Calidad de Servicio

- Calidad de servicio (Quality of Service, QoS): Medida cuantitativa de la calidad de la red, en términos de:
  - bitrate o throughput
  - retardos de paquetes
  - jitter
  - porcentaje de pérdidas de paquetes
  - ...
- Se dice que una aplicación tiene "requisitos de calidad de servicio" o necesita "calidad de servicio" cuando necesita unas garantías en todos o algunos de esos parámetros: Throughput mínimo, retardo máximo, jitter máximo, porcentaje máximo de pérdidas...
- Los protocolos TCP/IP en Internet proporcionan un servicio Ilamado best-effort ("mejor que se puede"), pues no se garantizan valores para esos parámetros.
- Living with lag https://www.youtube.com/watch?v=\_fNp37zFn9Q

## Multimedia sobre la red Internet actual

- Las aplicaciones multimedia frecuentemente presentan requisitos de calidad de servicio que Internet no da.
- Las aplicaciones multimedia en Internet utilizan técnicas en el nivel de aplicación para paliar en la medida de lo posible los efectos de los retardos y las pérdidas de mensajes:
  - Números de secuencia
  - Buffering en el receptor para compensar el jitter
  - Marcas de tiempo en los paquetes para reproducir de acuerdo a cómo se grabó el audio/vídeo
  - Corrección de errores de transmisión
  - Redes de distribución de contenidos
  - ...

# Evolución de Internet para soportar mejor las aplicaciones multimedia (I)

- Históricamente se estandarizaron dos modelos para proporcionar QoS en Internet:
  - IntServ (Integrated Services) (protocolo RSVP) 1994. Un usuario (emisor o receptor) solicita previamente los recursos que necesitará. Cada router del trayecto ha de tomar nota y efectuar la reserva de recursos solicitada.
  - ② DiffServ (Differentiated Services), 1998. Un usuario (emisor) marca los paquetes con una determinada etiqueta que indica la prioridad y el trato que dichos paquetes recibirán en los routers.
- Hoy no se utiliza ninguno de estos estándares a escala de toda Internet. Predomina la "filosofía" de Diffserv, pero normalmente sólo se aplica a escala de un ISP o de una organización.

# Evolución de Internet para soportar mejor las aplicaciones multimedia (II)

#### Filosofía de servicios integrados (Integrated services):

- Se requieren cambios drásticos en la red para ofrecer un servicio de reserva de recursos (ancho de banda, porcentaje de pérdidas) a las aplicaciones
- Requiere nuevo software en los terminales y en la red (routers)

#### Filosofía de servicios diferenciados (Differentiated services):

Requiere menos cambios en la red, proporcionando clases de servicio distintas

#### Laissez-faire

- No requiere prácticamente cambios en la red
- Los ISPs proporcionarán más velocidad de transmisión (ancho de banda) cuando haya demanda en el mercado
- Soluciones: distribución de contenidos en los extremos, multidifusión en redes solapadas (redes overlay entre pares)
  - Cambios en el nivel de aplicación, que no requieren cambiar la red (routers)

# Evolución de Internet para soportar mejor las aplicaciones multimedia (III)

- Por defecto Internet utiliza un modelo de servicio best effort:
  - Todos los paquetes son tratados por igual, independientemente de la aplicación o protocolo a que pertenezcan.
- Alternativa: varias clases de tráfico a las que se da un servicio diferente:
  - Se divide el tráfico en clases
  - La red trata clases de tráfico distintas de manera distinta (como la clase 1<sup>a</sup> y 2<sup>a</sup> en un tren, o 1<sup>a</sup>, business class y turista en un avión)

## Evolución histórica del campo ToS en la cabecera IP (I)

- Históricamente, el campo ToS (tipo de servicio, type of service) de la cabecera IP estaba pensados precisamente para un modelo basado en clases de servicio.
- Con los años ha ido cambiando la semántica de este campo:

				TOS				
	7	6	5	4	3	2	1	0
RFC 791	precedence			D (delay)	T (throuput)	R (reliability)	unused	
RFC 1349	precedence			D (delay)	T (throuput)	R (reliability)	C (cost)	0
RFC 2474	DSCP (DiffServ Code Point)						ECN (congestion)	

## Evolución histórica del campo ToS en la cabecera IP (II)

- RFC 791: Especificación del Protocolo IP (1981).
  - *Precedence*: Prioridad del paquete, con ámbito de una organización. Jamás se utilizó
  - Bits D, T, R: Indicaciones del tipo de servicio requerido.
- RFC 1349: Aclaraciones sobre el campo ToS (1992).
  - Precedence: Prioridad del paquete, no se habla de él en la RFC.
  - Bits D, T, R, C: Indicaciones del tipo de servicio requerido. Se añade el bit C. Sólo se explica el sentido de la activación de 1 sólo de los 4 bits, y se establecen como deberían usarse en los protocolos de aplicación existentes.
- RFC 2474: Especificación de DiffServ (1998). Reemplaza el campo ToS por el campo DS (Differenciated Services field).
  - *DSCP*: Permite etiquetar paquetes de forma diferente para darles un servicio en función del valor de este campo.
  - ECN: Bits sin usar en la RFC 2474, definidos en la RFC 3168 para notificación explícita de congestión (no relacionado con calidad de servicio).

## Flujo vs Clase

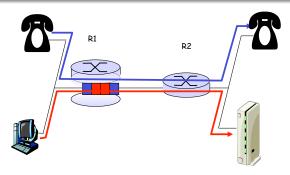
- Flujo: secuencia de datagramas de un determinado usuario que requiere la misma QoS.
  - Un flujo es la entidad más pequeña a la que los routers pueden aplicar una determinada QoS.
  - Un flujo es unidireccional.
- Clase: agrupación de flujos. Todos los flujos dentro de una misma clase reciben la misma QoS.
  - A veces se puede proporcionar QoS a nivel de fujo, a veces simplemente a nivel de una clase que agrupe varios flujos.

- Aplicaciones multimedia en redes
- 2 Evolución Calidad de servicio
- 3 Técnicas para proporcionar múltiples clases de tráfico
- Mecanismos de planificación
- Mecanismos de vigilancia (policing & shaping)
- 6 Referencias

## Escenario: tráfico HTTP y de audio compitiendo

## Ejemplo: telefonía IP a 1Mbps y una transferencia de un fichero por HTTP comparten un enlace a 1'5 Mbps.

- HTTP transmitiendo un fichero ⇒ aplicación elástica tolerante a retardos, no tolerante a pérdidas.
- Las ráfagas de paquetes HTTP pueden congestionar el router causando pérdidas y/o retrasos de paquetes de audio.
- ¿Se le debería dar algún tipo de prioridad al tráfico de audio?



## Técnicas para proporcionar múltiples clases de servicio PRINCIPIO 1: marcado de paquetes en clases para un tratamiento diferenciado

#### Principio 1:

- El marcado de los paquetes permite a un router diferenciar entre paquetes pertenecientes a distintas clases de tráfico:
  - Dirección IP origen, dirección IP destino, puerto origen, puerto destino, protocolo de nivel de transporte.
- Las disciplinas de planificación de colas en el router deben permitir tratar de manera diferenciada las distintas clases de tráfico

## Técnicas para proporcionar múltiples clases de servicio PRINCIPIO 2: aislamiento de clases (I)

- Un cliente contrata con su ISP una determinada QoS: a esto se llama SLA (Service-Level Agreement).
- ¿Qué ocurre si las aplicaciones no se comportan según lo pactado?
- Ej: una aplicación de audio envía a una tasa de bits superior a la declarada, ya sea maliciosamente o por error

#### Principio 2:

Es deseable proporcionar un grado de aislamiento entre las clases de tráfico y entre los flujos, de manera que una clase o un flujo no se vea afectado de forma adversa por otro que tiene un comportamiento erróneo.

## Técnicas para proporcionar múltiples clases de servicio PRINCIPIO 2: aislamiento de clases (II)

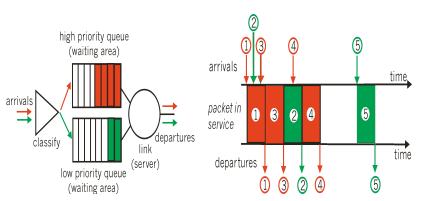
#### ¿Cómo proporcionar aislamiento?

- Las decisiones de vigilancia (policing) en los routers del borde permiten aislar el tráfico con granularidad de flujos:
  - monitorización para garantizar que la velocidad agregada de un flujo individual no excede cierto valor, y retardando, descartando, o marcando como descartables los paquetes que no respetan el contrato
- El mecanismo de planificación en los routers trata de manera distinta las clases o los flujos, asignando anchos de banda distintos a cada clase/flujo

Flujos vs. clases: Por razones de escala, es frecuenta que en la frontera se discrimine según flujos, y en el centro de la organización según clases.

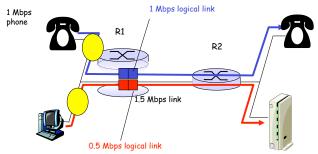
## Técnicas para proporcionar múltiples clases de servicio PRINCIPIO 2: aislamiento de clases (III)

Ejemplo de división del tráfico en clases por prioridad: alta prioridad y baja prioridad. La salida de paquetes se gestiona con un mecanismo de planificación de colas atendiendo primero a la cola de mayor prioridad.



## Técnicas para proporcionar múltiples clases de servicio PRINCIPIO 3: utilización de recursos de manera eficiente

 Dedicar un ancho de banda fijo, no compartible, a un flujo hace un uso ineficiente de la red. Ej. Audio no consume porque están callados y no puede aprovecharse para ficheros transferidos con HTTP.



#### Principio 3:

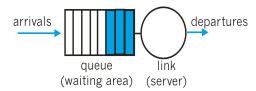
Si bien es bueno proporcionar aislamiento entre clases o flujos, es deseable también utilizar los recursos de la forma más eficiente posible

- Aplicaciones multimedia en redes
- 2 Evolución Calidad de servicio
- Técnicas para proporcionar múltiples clases de tráfico
- Mecanismos de planificación
- 5 Mecanismos de vigilancia (policing & shaping)
- 6 Referencias

- Aplicaciones multimedia en redes
  - 1. Flujos de audio/vídeo almacenados
  - 2. Flujos de audio/vídeo en vivo
  - 3. Flujos de audio/vídeo interactivo en tiempo real
- 2 Evolución Calidad de servicio
- 3 Técnicas para proporcionar múltiples clases de tráfico
- 4 Mecanismos de planificación
  - Planificación FIFO
  - Colas con prioridad
  - Colas de turno rotatorio
  - Colas equitativas ponderadas (WFQ)
- Mecanismos de vigilancia (policing & shaping)
  - Token bucket
  - Token bucket +WFQ
  - Referencias

# Mecanismos de planificación: FIFO

- Disciplinas de planificación de enlace: la forma en que el router escoge el siguiente paquete que será enviado por un enlace de entre los que están almacenados en el buffer de salida
- Planificación FIFO (first in first out): se envía en el orden de llegada a la cola
  - Como cualquier cola (autobús, cine,...)
  - Política de eliminación de paquetes: si un paquete llega a la cola cuando está llena, ¿qué paquete se descarta?. Alternativas:
    - el último que llega pierde
    - e eliminar en función de la prioridad de cada paquete
    - aleatoriamente



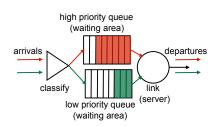
- - 1. Flujos de audio/vídeo almacenados
  - 2. Flujos de audio/vídeo en vivo
  - 3. Flujos de audio/vídeo interactivo en tiempo real

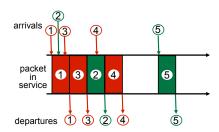
- Mecanismos de planificación
  - Planificación FIFO
  - Colas con prioridad
  - Colas de turno rotatorio
  - Colas equitativas ponderadas (WFQ)
- - Token bucket.
  - Token bucket +WFQ

GS<sub>v</sub>C - 2024 Calidad de Servicio

## Mecanismos de planificación: Colas con prioridad

- Se transmite el paquete más prioritario de los encolados
- Cada clase/flujo tiene una prioridad distinta
  - para el router la clasificación de los paquetes puede realizarse según cómo venga marcado, o en función de otras cabeceras como origen/destino IP, números de puerto, protocolo.
  - se eligen paquetes para transmitir de la cola de más alta prioridad no vacía. Dentro de cada cola, FIFO



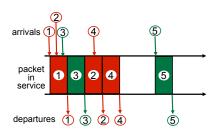


- Aplicaciones multimedia en redes
  - 1. Flujos de audio/vídeo almacenados
  - 2. Flujos de audio/vídeo en vivo
  - 3. Flujos de audio/vídeo interactivo en tiempo real
- 2 Evolución Calidad de servicio
- 3 Técnicas para proporcionar múltiples clases de tráfico
- 4 Mecanismos de planificación
  - Planificación FIFO
  - Colas con prioridad
  - Colas de turno rotatorio
  - Colas equitativas ponderadas (WFQ)
- Mecanismos de vigilancia (policing & shaping)
  - Token bucket
  - Token bucket +WFQ
  - Referencias

# Mecanismos de planificación: Colas de turno rotatorio

## **Round Robin Scheduling**

- diferentes colas almacenan diferentes clases de tráfico
- de entre todas las colas que tengan paquetes, se elige por turno rotatorio un paquete de cada una
- no hay prioridades, todas las colas con las diferentes clases de tráfico se tratan por igual.

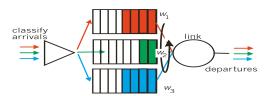


- Aplicaciones multimedia en redes
  - 1. Flujos de audio/vídeo almacenados
  - 2. Flujos de audio/vídeo en vivo
  - 3. Flujos de audio/vídeo interactivo en tiempo real
- 2 Evolución Calidad de servicio
- 3 Técnicas para proporcionar múltiples clases de tráfico
- 4 Mecanismos de planificación
  - Planificación FIFO
  - Colas con prioridad
  - Colas de turno rotatorio
  - Colas equitativas ponderadas (WFQ)
- Mecanismos de vigilancia (policing & shaping)
  - Token bucket
  - Token bucket +WFQ
  - Referencias

# Mecanismos de planificación: Colas equitativas ponderadas (WFQ)

## Weighted Fair Queuing (WFQ)

- A cada clase/flujo i se le asigna un peso wi y recibe un ancho de banda proporcional a su peso en cada turno. Si R paquetes/s es el ancho de banda global del enlace, el flujo i recibirá un ancho de banda de wi: R wi paquetes/s
- Aún cuando haya tráfico en otras colas se garantiza que se conservan los pesos de cada clase (proporción de ancho de banda)



- Aplicaciones multimedia en redes
- 2 Evolución Calidad de servicio
- 3 Técnicas para proporcionar múltiples clases de tráfico
- Mecanismos de planificación
- 5 Mecanismos de vigilancia (policing & shaping)
- 6 Referencias

41

# Mecanismos de vigilancia

- Objetivo: limitar el tráfico para que no supere los parámetros declarados/comprometidos
- ¿Qué aspectos de la tasa de paquetes de un flujo deben ser vigilados?
- Tres criterios utilizados habitualmente para vigilar/limitar:
  - (a largo plazo) tasa promedio: cuántos paquetes se pueden enviar por unidad de tiempo en promedio
    - 100 paquetes/s y 6000 paquetes por minuto (ppm) tienen la misma tasa promedio. Probablemente hay que establecer definir un periodo de duración para una cierta tasa promedio.
  - Tasa de pico: Ej: 6000 paquetes por minuto (ppm) de media y 150 pps de tasa de pico
  - Tamaño máximo de una ráfaga: número máximo de paquetes que se pueden enviar "de golpe": consecutivamente, sin solución de continuidad, en un intervalo lo más pequeño posible

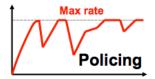
# Mecanismos de vigilancia: policing & shaping

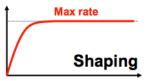
#### Policing:

- Descarta el tráfico que supera los parámetros comprometidos.
- Se suele utilizar en interfaces de entrada y de salida

#### Shaping:

- Encola el tráfico que supera los parámetros comprometidos.
- Se suele utilizar sólo en interfaces de salida
- Retardar un paquete en vez de descartarlo



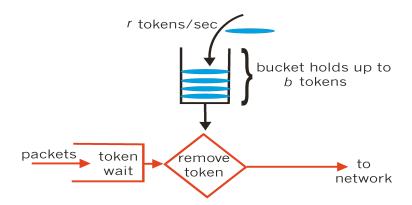


- Aplicaciones multimedia en redes
  - 1. Flujos de audio/vídeo almacenados
  - 2. Flujos de audio/vídeo en vivo
  - 3. Flujos de audio/vídeo interactivo en tiempo real
- 2 Evolución Calidad de servicio
- 3 Técnicas para proporcionar múltiples clases de tráfico
- 4 Mecanismos de planificación
  - Planificación FIFO
  - Colas con prioridad
  - Colas de turno rotatorio
  - Colas equitativas ponderadas (WFQ)
- Mecanismos de vigilancia (policing & shaping)
  - Token bucket
  - Token bucket +WFQ
  - Referencias

# Mecanismos de vigilancia (policing & shaping): Token bucket (I)

- Cubeta con fichas (token bucket): Abstracción que sirve para caracterizar los límites de vigilancia en función del tamaño máximo de ráfaga y la tasa promedio.
- Los paquetes que llegan se encolan. Para salir un paquete tiene que haber al menos una ficha en la cubeta: se elimina la ficha y se da salida al paquete.
  - La cubeta puede alojar un máximo de b fichas (tokens)
  - Las fichas se generan a una velocidad de r fichas/s
    - si la cubeta está llena y siguen llegando fichas, éstas se pierden
  - en un intervalo de longitud t: el número de paquetes que se admiten es < rt + b
  - la tasa de generación de fichas, *r*, limita la tasa promedio a largo plazo
- Dos cubetas en serie pueden además limitar la tasa de pico

# Mecanismos de vigilancia (policing & shaping): Token bucket (II)



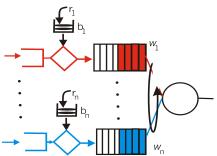
- Aplicaciones multimedia en redes
  - 1. Flujos de audio/vídeo almacenados
  - 2. Flujos de audio/vídeo en vivo
  - 3. Flujos de audio/vídeo interactivo en tiempo real
- 2 Evolución Calidad de servicio
- Técnicas para proporcionar múltiples clases de tráfico
- 4 Mecanismos de planificación
  - Planificación FIFO
  - Colas con prioridad
  - Colas de turno rotatorio
  - Colas equitativas ponderadas (WFQ)
- Mecanismos de vigilancia (policing & shaping)
  - Token bucket
  - Token bucket +WFQ
  - Referencias

GSyC - 2024 Calidad de Servicio

47

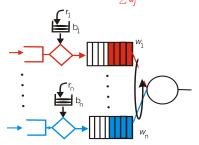
# Mecanismos de vigilancia (policing & shaping): Token bucket + WFQ (I)

 La cubeta con fichas se puede combinar con WFQ (una cubeta a la entrada de cada cola) para proporcionar un límite superior al retardo que un paquete puede estar en la cola de un router



# Mecanismos de vigilancia (policing & shaping): Token bucket + WFQ (II)

- Supongamos que  $b_1$  está llena y llega una ráfaga de  $b_1$  paquetes de esa clase. Los  $b_1$  paquetes pasan a la cola  $WFQ_1$  sin esperar.
- ullet La velocidad de salida de paquetes  $WFQ_1$  será:  $R rac{w_1}{\sum w_j}$
- El tiempo máximo que pasa un paquete de la ráfaga esperando en la cola  $WFQ_1$ , será el tiempo que pase el último paquete de la ráfaga en  $WFQ_1$ . Es decir, el tiempo que necesite  $WFQ_1$  para enviar todos los paquetes que hay en la cola:  $D_{max} = \frac{b_1}{R \sum_{l=1}^{W_1}}$



- Aplicaciones multimedia en redes
- 2 Evolución Calidad de servicio
- Técnicas para proporcionar múltiples clases de tráfico
- 4 Mecanismos de planificación
- 5 Mecanismos de vigilancia (policing & shaping)
- 6 Referencias

### Referencias

- James F. Kurose y Keith W. Ross, Redes de Computadores: un enfoque descendente, Pearson Educación, 5<sup>a</sup> edición.
- Larry L. Peterson, Bruce S. Davie, **Computer networks, a systems approach**, edition 4. Morgan Kaufmann 2007.