QoS: Introducción

Area de Ingeniería Telemática
<http://www.tlm.unavarra.es>

Grado en Ingeniería en Tecnologías de
Telecomunicación, 3º

Pérdidas y disponibilidad

Pérdidas

- Los paquetes podrían no llegar nunca a su destino

Posibles motivos

- Se corrompió y fue descartado en algún nodo de la red (CRCs)
- BER = Bit Error Rate
- Aproxima a la probabilidad de error de bit p_{err}
- Probabilidad de algún error en un paquete de N bits:

$$p_{epk} = 1 - (1 - p_{err})^N$$

- Asumiendo errores indep. (no ráfagas)
- Sin código “corrector” de errores
- Ejemplo: $p_{err} = 10^{-6}$, $N = 12.000 \rightarrow \rightarrow$
 $p_{epk} \approx 10^{-2}$
- (...)

Pérdidas

- Los paquetes podrían no llegar nunca a su destino

Posibles motivos

- Se descartó en un nodo de la red por desbordamiento de buffer
- (...)

Tamaño del buffer

Pérdidas

- Los paquetes podrían no llegar nunca a su destino

Posibles motivos

- Se descartó por exceder el tiempo en la red
- (...)

Pérdidas

- Los paquetes podrían no llegar nunca a su destino

Posibles motivos

- Fallo de un elemento de red
- Lleva un tiempo recalcular caminos
- (...)

Pérdidas

- Los paquetes podrían no llegar nunca a su destino

Posibles motivos

- Descarte en nodo extremo por desbordamiento de buffer
- Puede ser culpa de la propia aplicación y del tiempo que le lleva procesar los datos recibidos

Ejemplo

Voz y pérdida de paquetes

- Causan cortes y saltos
- Un paquete suele contener en torno a 20ms de muestras de voz
 - Si contiene menos: mayor ratio cabeceras/datos
 - Si contiene más: mayor retardo de formación
- Pérdida de 1 paquete se puede intentar recuperar (interpolación, etc)
- Pérdida de más de 1 paquete crea un corte que se nota claramente (...)

Ejemplo

Paquetes retrasados y jitter (en voz)

- Si un paquete llega demasiado tarde es equivalente a una pérdida
- Si el retardo general end-to-end es demasiado grande se pierde interactividad
- El jitter (variación en el retardo) se puede recuperar con buffer en el receptor
- Si el jitter es demasiado grande el buffer ha de ser tan grande que de nuevo implica demasiado retardo

Availability

- **Network availability**

- Fracción de tiempo en que la conectividad está disponible
- Puede faltar la conectividad por cortes programador o por fallos
- Se tiene en cuenta la disponibilidad de cada elemento y se combinan, según estén en serie o en paralelo
- En serie deben estar todos disponibles (“disponible link 1 Y disponible link 2 Y disponible link 3 ...”)
- En paralelo debe estar alguno disponible (“disponible link 1 O disponible link 2 O disponible link 3 ...”)

Availability

- **Service availability**

- Fracción de tiempo en que el servicio está disponible dentro de los parámetros de SLA del servicio
- Puede medirse independiente de la disponibilidad de red
- En ese caso caso no puede superarla o debe medirse condicionada a disponibilidad
- Puede implicar o no al comportamiento de servidores, según cómo se haya definido el servicio

Congestión y QoS

Escenarios con congestión

- Interconexión LAN-to-WAN
 - Tráfico desde LAN de $> 10\text{Mbps}$ hacia WAN de $< 10\text{Mbps}$
- (...)

Escenarios con congestión

- Interconexión LAN-to-WAN
 - Tráfico desde LAN de $> 10\text{Mbps}$ hacia WAN de $< 10\text{Mbps}$
- Agregación

Congestión y Calidad de Servicio

- Congestión: colas llenas
- Fácil dar QoS si nunca hay congestión
- Para dar QoS con congestión
 - Opción 1: Gestionar los recursos ante congestión para dar un trato diferenciado (congestion control/management)
 - Opción 2: Evitar la congestión (congestion avoidance)

Por efecto de retransmisiones

Solución

- *Throw more bandwidth at the problem !!*
- Así no hay congestión
- No siempre es la solución:
 - Si se le da más BW al usuario habrá mayor demanda
 - No siempre es barato aumentar el BW (ej: acceso, *transit*)
 - Un pico en la demanda degradaría la calidad de servicios sensibles
 - Voz: pequeño BW y bajo retardo pero para lograr bajo retardo puede hacer falta un BW desproporcionado (y caro)
 - ¡¡ Congestión por tráfico *scavenger* !! (DoS, worm... bastan 10 PCs para saturar 1GEth)

¿Solución?

- Hoy en día ya no es suficiente para un ISP con ofrecer un servicio *Best Effort*
- La tecnología para ofrecer QoS ya es asequible y fiable
- ¡ Pero la ingeniería de estas redes no es sencilla !

Escenarios históricos

Escenarios históricos

PSTN

- Conmutación de Circuitos
- BW fijo y garantizado
- Retardo fijo y acotado
- Diseñada para tráfico de voz
- Para datos BW sin usar

Tecnología IP

- Conmutación de datagramas
- *Best effort*
- *No single points of failure*
- TOS no se llega a usar

“BellHeads” vs “NetHeads”

“We reject kings, presidents and voting. We believe in rough consensus and running code” D. Clark

The Intelligent Network

- Red inteligente
- Terminales tontos
- Conoce el servicio
- Complejo dimensionamiento para asegurar el servicio
- Su colaboración es necesaria para el servicio

The Stupid Network

- Red simple
- Terminales inteligentes
- Solo mueve bits
- Se puede congestionar y se resuelve con más capacidad
- Los usuarios extremo controlan en exclusiva el servicio

Escenarios históricos

¿ Tener dos infraestructuras ?

- Más caro
- Más equipos
- Más BW sin usar
- Gestión independiente

¿Por qué no usar simplemente la PSTN?

- No optimizada para datos
- Arquitectura rígida
- Ineficiente en asignaciones de BW
- Inadecuada para sesiones cortas, de tasa variable, multipunto, etc

Escenarios históricos

¿ Usar una red de paquetes ?

- Deberá tener soporte de QoS
- Aprovechar multiplexación estadística
- Circuitos virtuales
- ATM
- No extendida

¿IP?

- Best Effort aun sobre tech. con QoS
- Separar los flujos IP en flujos de tecnología con QoS
- o añadir QoS a IP
- Para ello tratar de forma diferenciada al tráfico de datos
- Hoy solo en el dominio de redes concretas

IP QoS

- Tecnologías de capa 2 ofrecen QoS (ATM, Ethernet...)
- IP es la tecnología de nivel de red extremo a extremo más extendida
- Diferentes tecnologías capa 2 pueden ser empleadas en el camino capa 3
- Lo más razonable es mapear QoS de capa 3 en la QoS de capa 2 de cada tecnología en cada salto
- En vez de mapear capacidades de capa 2 de un salto en el siguiente