

# QoS: Introducción

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

Grado en Ingeniería en Tecnologías de  
Telecomunicación, 3º



¿QoS?



# ¿ Qué es esto de la calidad ?

## Para el usuario final experimentado

- Es normal que una llamada con un ordenador tenga diferente calidad que una por teléfono fijo o que una por móvil
- ¡ Aunque todas se cursen por la misma red !
- Es simplemente aquello a lo que está acostumbrado

(...)



# ¿ Qué es esto de la calidad ?

## Para el usuario final

- Pero si nunca ha usado un móvil esperará una calidad similar a la PSTN y se quejará
- Lo mismo si nunca ha usado VoIP

(...)



# ¿ Qué es esto de la calidad ?

## Para el usuario final

- La calidad es relativa a las expectativas
- Lo mismo con el precio, si está acostumbrado a una tarifa plana o gratis le extrañará pagar

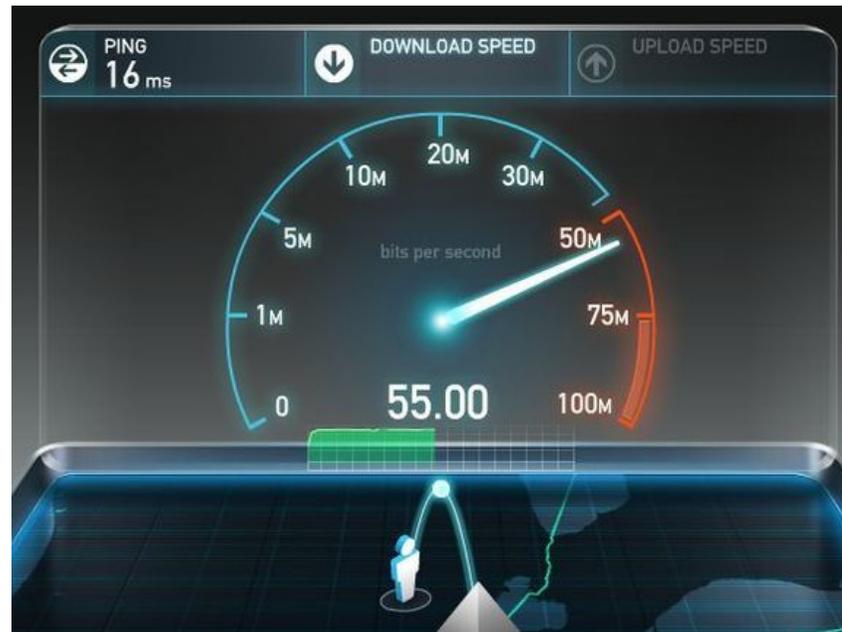
(...)



# ¿ Qué es esto de la calidad ?

## Para el técnico

- Habilidad de la red de *diferenciar* a unos determinados tipos de tráfico, probablemente de unos servicios concretos
- Controlar ciertos parámetros estadísticos:
  - Bandwidth, pérdidas, retardo, jitter... quejas de usuarios
  - Más absolutos y medibles
- ¿En qué se basa? (...)



# ¿ Qué es esto de la calidad ?

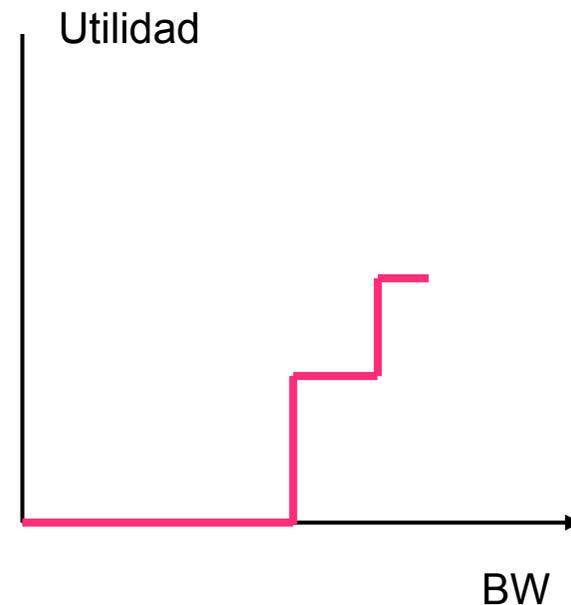
## Para el técnico

- Se basa en un reparto “injusto” pero controlado
  - Ofrecer recursos a clases de alta prioridad *a costa de* las de baja
- Formalizados en SLAs (*Service Level Agreements*)
  - Acuerdo entre proveedor de servicio (la red) y suscriptor (el cliente)
  - Dentro varios SLs (*Service Level Specifications*)
  - Especifica la calidad de servicio que garantizará el proveedor
  - La red mantendrá su promesa mientras los flujos de usuario se mantengan dentro de su especificación de tráfico
  - Especifica las medidas que se tomarán si se incumple
  - ¡ Gran cantidad de parámetros posibles según el servicio !



# Usuario: Utilidad

- Las aplicaciones son sensibles a pérdidas, capacidad, retardo, variación en el retardo
- Por debajo de un umbral puede no ser útil el tráfico
- Ofrecer garantías de prestaciones para
  - Que el usuario esté satisfecho
  - Que los recursos se usen de forma óptima



# ¿ Quién necesita QoS ?

- Dos tipos de aplicaciones/tráfico:
  - Elástico
    - Se ajusta ante grandes cambios en retardo y throughput
    - Sigue manteniendo la funcionalidad de la aplicación
  - Inelástico
    - Si no se cumplen unos requisitos de calidad la utilidad se vuelve 0



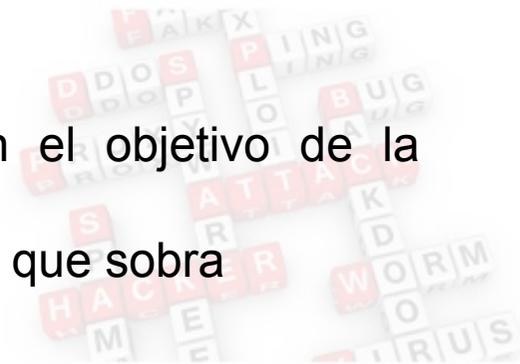
# Requisitos de QoS de las aplicaciones

Aplicación	Fiabilidad	Retardo	Jitter	Ancho de Banda
Correo electrónico	Alta (*)	Alto	Alto	Bajo
Transferencia de ficheros	Alta (*)	Alto	Alto	Medio (**)
Acceso Web	Alta (*)	Medio	Alto	Medio
Login remoto	Alta (*)	Medio	Medio	Bajo
Audio bajo demanda	Media	Alto	Medio	Medio
Vídeo bajo demanda	Media	Alto	Medio	Alto
Telefonía	Media	Bajo	Bajo	Bajo
Vídeoconferencia	Media	Bajo	Bajo	Alto

- (\*) La fiabilidad alta en estas aplicaciones se consigue automáticamente al utilizar el protocolo de transporte TCP
- (\*\*) Transferencia de ficheros: si es interactiva el usuario espera que tarde proporcionalmente al tamaño, luego depende del BW

# ¿ Quién necesita QoS ?

- Voz (IP telephony, radio?)
- Vídeo (streaming, videoconferencia)
- Ciertas aplicaciones de datos (generalmente elásticas)
  - *Transactional Data/Interactive Data* (SAP, Oracle...)
  - *Bulk Data* (backups, replicación en redes de contenidos...)
  - *Locally Defined Mission-Critical Data* (mayor que *transactional*)
- Resto:
  - *Best Effort*
  - Dejar BW para él
  - Gran cantidad de aplicaciones en una empresa (centenares)
  - Probablemente no se puedan clasificar todas, ¡no ahogarlas!
- ¿Queda algo?: *Scavenger Service*
  - *Less than BE*
  - Tráfico no deseado: DoS, Worms, etc
  - Web surfing a destinos no relacionados con el objetivo de la empresa
  - Si no se descarta se cursa solo en la capacidad que sobra



# ¿ Qué necesitan ?

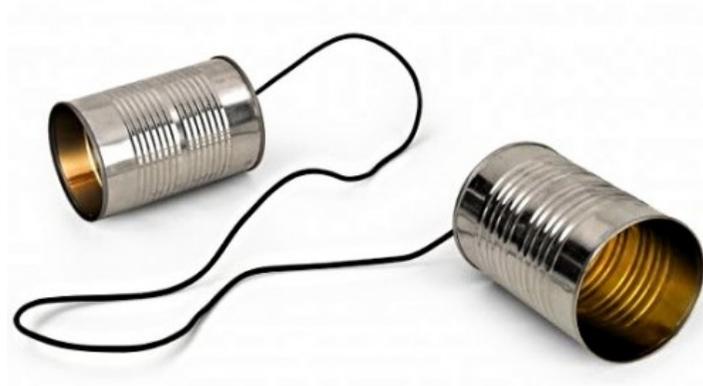
- Que sea predecible el comportamiento de la red
- Garantizar (depende de la aplicación):
  - Bandwidth (Throughput)
  - Delay
  - Variación en el retardo (jitter)
  - Pérdidas



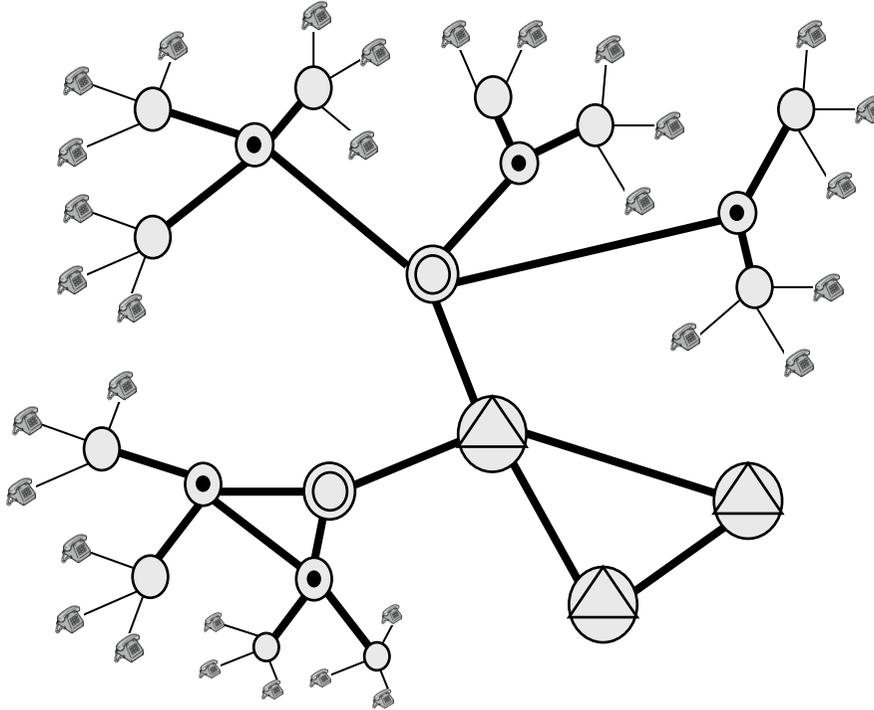
# Cuestión

- ¿Cuáles de los siguientes ejemplos de aplicaciones serían elásticas respecto al throughput?
  - Una llamada de VoIP
  - Navegación Web
  - Descarga de ficheros por FTP
  - Dropbox
  - Netflix

# Escenarios históricos

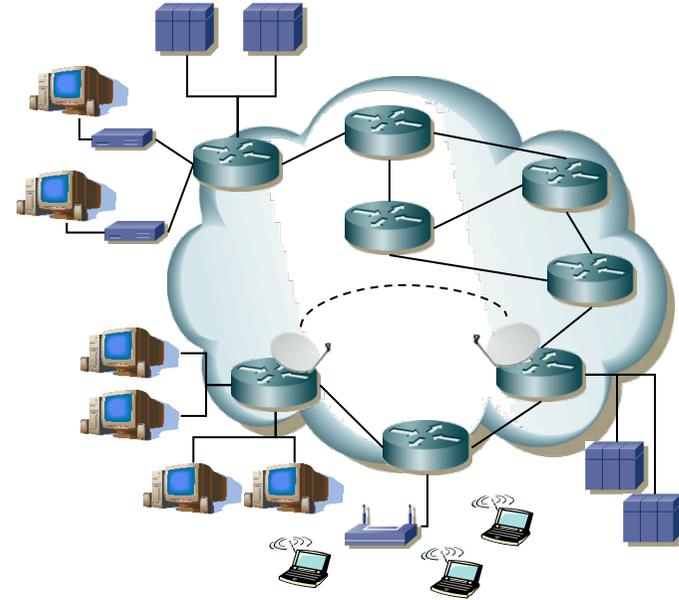


# Escenarios históricos



## PSTN

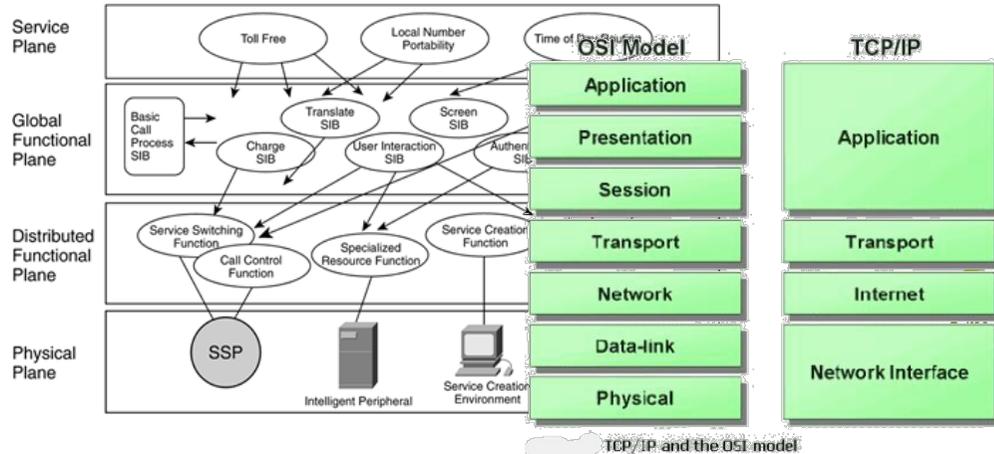
- Conmutación de Circuitos
- BW fijo y garantizado
- Retardo fijo y acotado
- Diseñada para tráfico de voz
- Para datos BW sin usar



## 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*

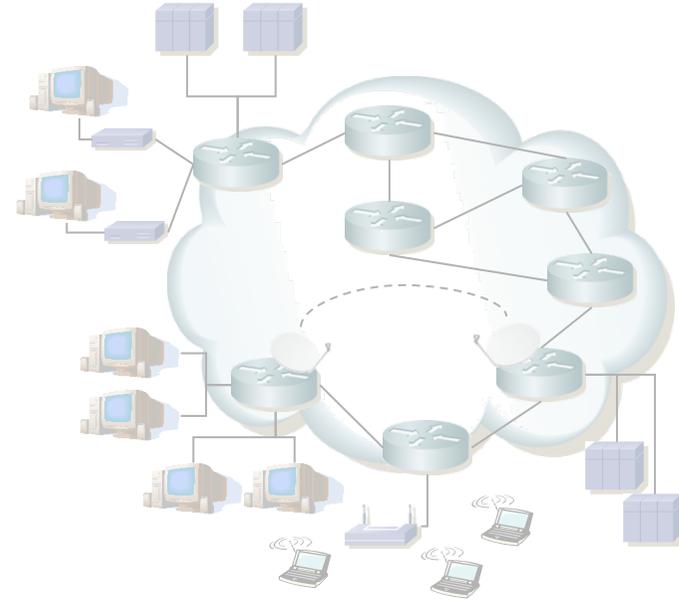
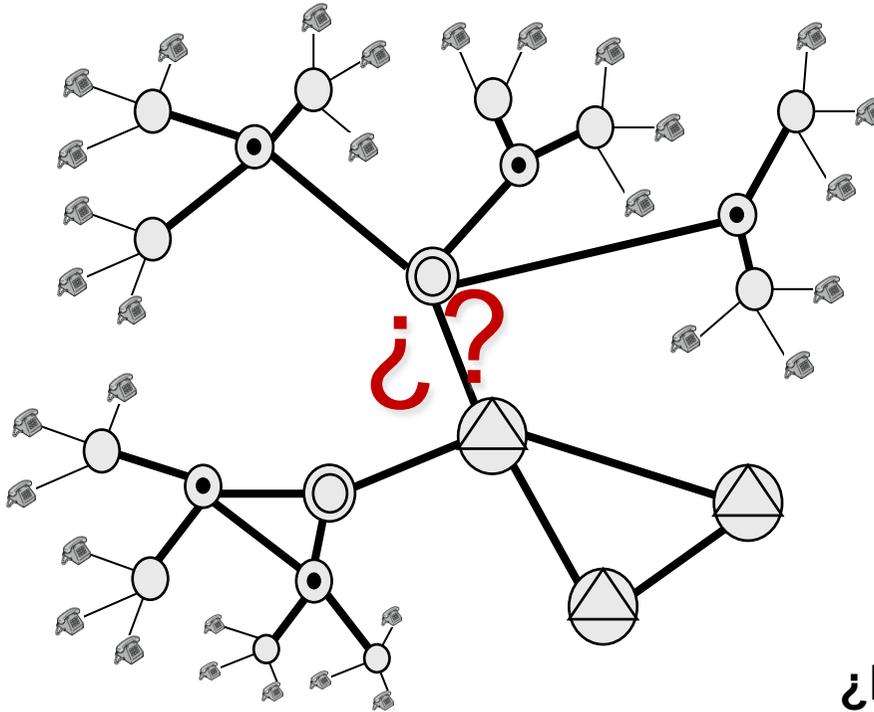
## The Intelligent Network

- 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



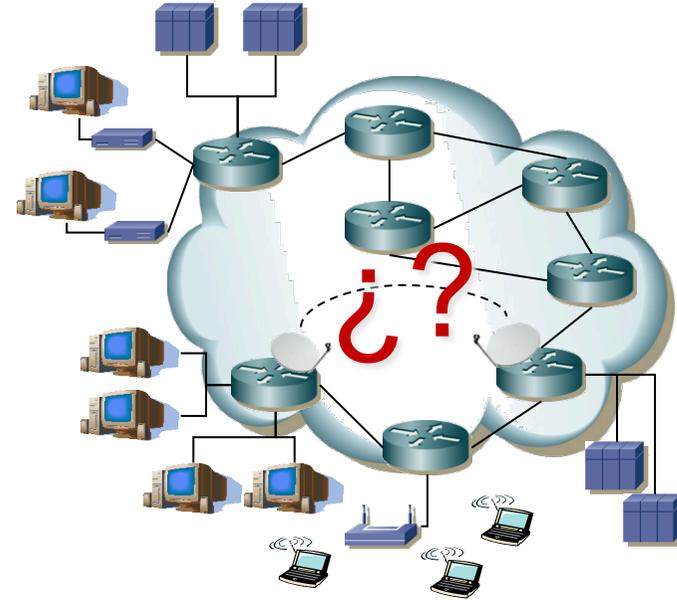
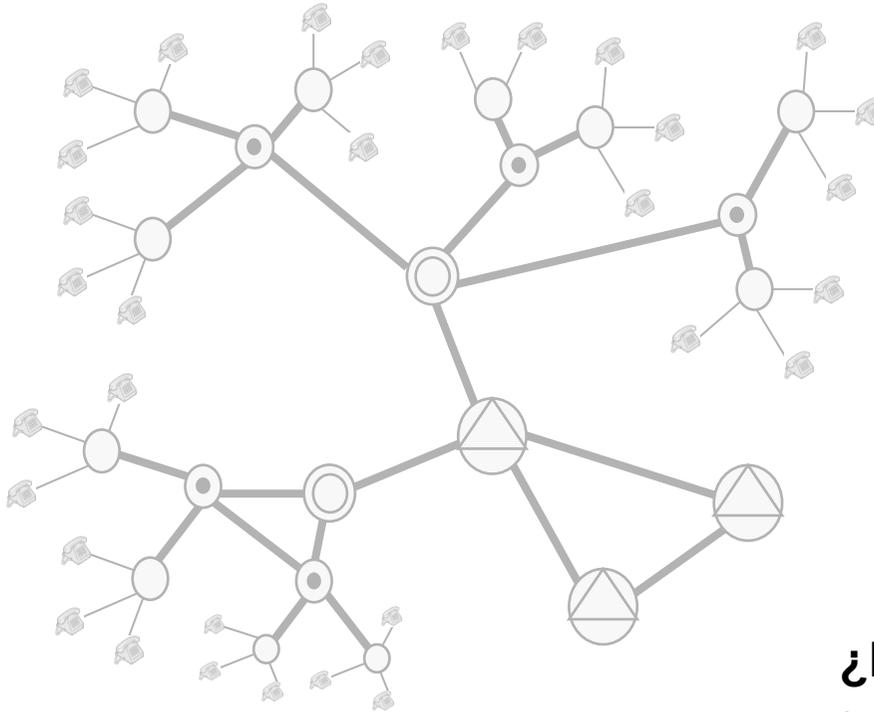
## ¿ Tener dos infraestructuras ?

- 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 ?

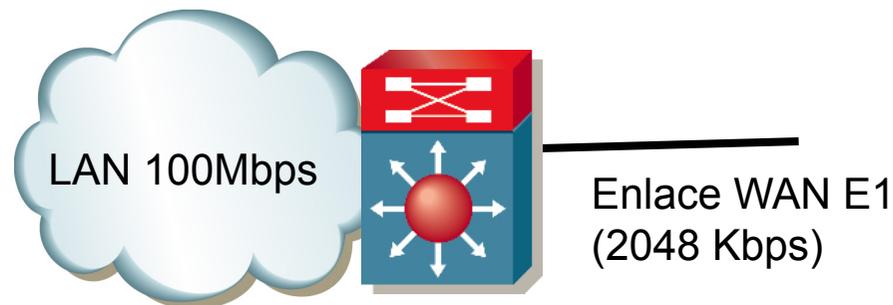
- Deberá tener soporte de QoS
- Aprovechar multiplexación estadística
- Circuitos virtuales
- ATM
- No extendida

## ¿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

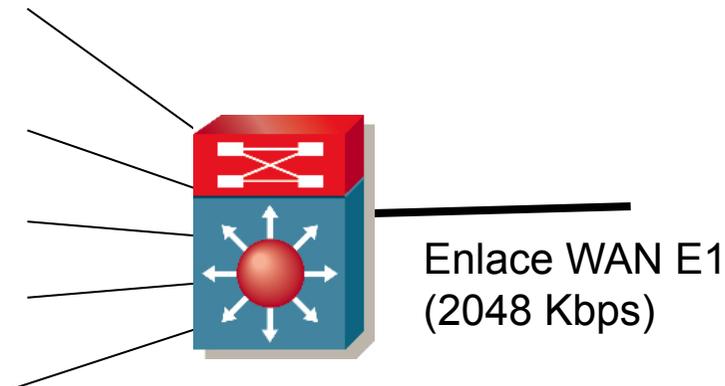
# Escenarios con congestión

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



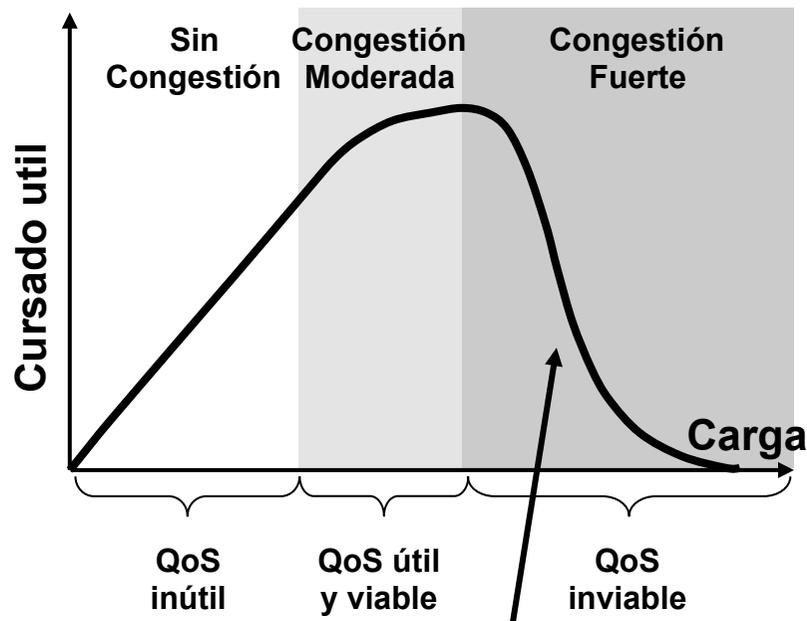
# Escenarios con congestión

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

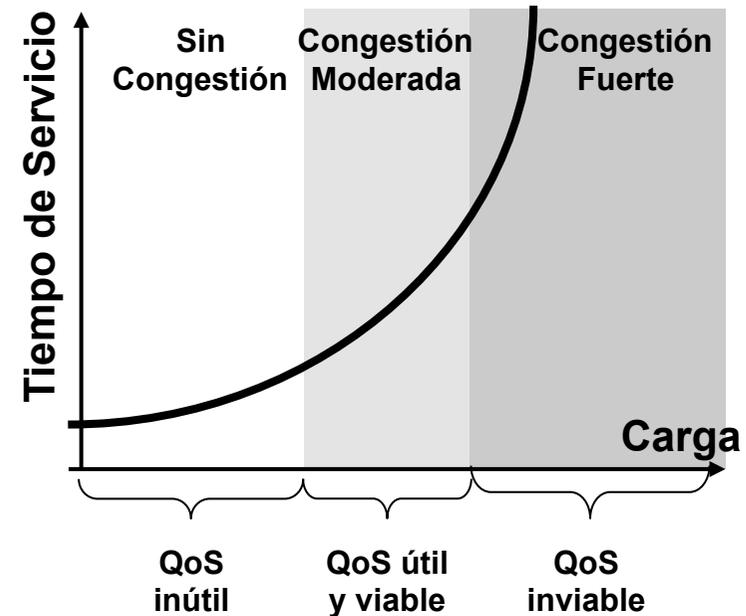


# Congestión y Calidad de Servicio

- 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 !



# 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

