

Análisis de Rendimiento

Preparado con materiales de:

Carlos Vicente

Servicios de Red/Universidad de Oregon

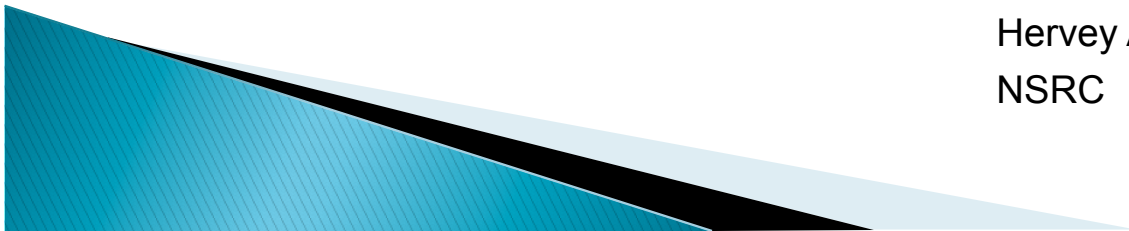
Presentación:

Carlos Armas

Roundtrip Networks

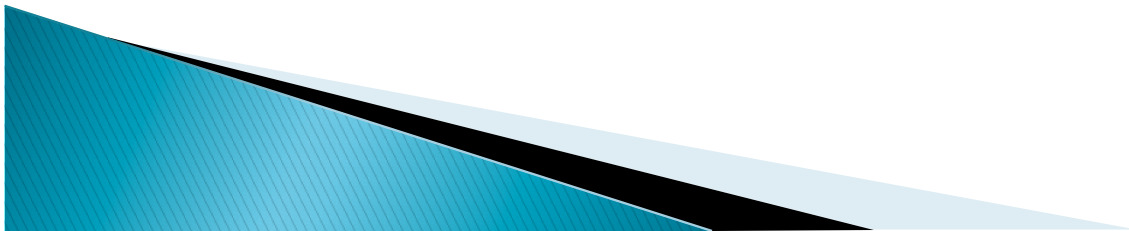
Hervey Allen

NSRC



Contenido

- ▶ Planificación de la gestión del rendimiento
- ▶ Métricas
 - Red
 - Sistemas
 - Servicios
- ▶ Ejemplos de mediciones



Planificación

▶ Propósito

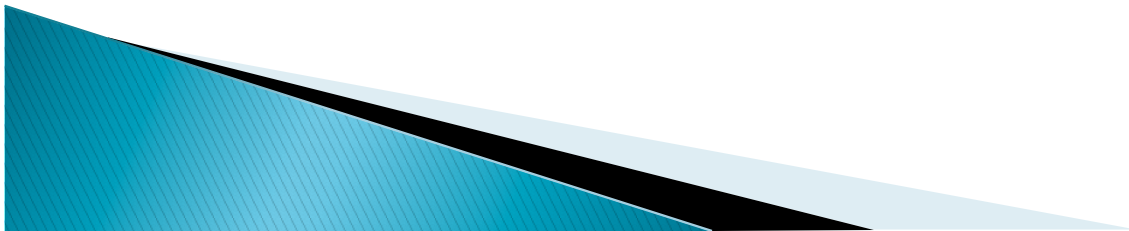
- Establecer estado estable (línea base)
- Diagnóstico de fallas,
- Anticipar crecimiento de demanda

▶ A quién va dirigida la información?

- Gerencia,
- NOC y personal técnico,
- Clientes

▶ Alcance

- Impacto en los dispositivos (medidos y de medición)
- Balance entre cantidad de información y tiempo de recolección



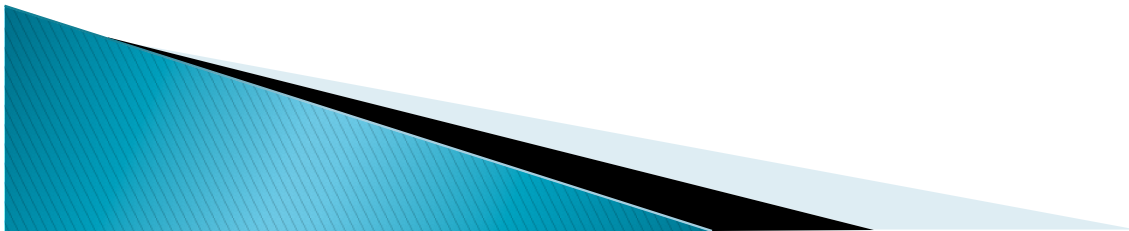
Métricas

- Rendimiento de red
 - Capacidad del canal:
 - nominal y efectiva
 - Utilización del canal
 - Retardo y “parpadeo” (*jitter*)
 - Pérdida de paquetes y errores
- Rendimiento de sistemas
 - Disponibilidad
 - Memoria, utilización y carga de CPU
 - Utilización de dispositivos de entrada/salida
- Rendimiento de servicios
 - Disponibilidad
 - Tiempo de acceso y carga



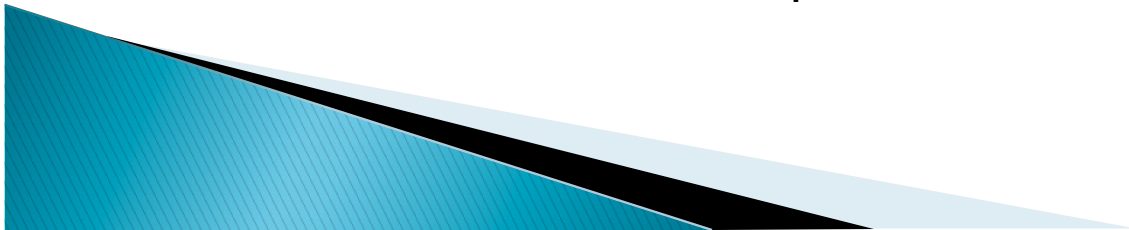
Métricas de rendimiento de red más comunes

- Relativas al tráfico:
 - Bits por segundo
 - Paquetes por segundo
 - Paquetes *unicast* vs. paquetes *no-unicast*
 - Errores
 - Paquetes descartados
 - Flujo por segundo
 - Tiempo de ida y vuelta (RTT)
 - Dispersión del retardo (parpadeo o “jitter”)



Capacidad Nominal del Canal

- Máxima cantidad de datos transmitidos por unidad de tiempo
 - Ejemplo: bits por segundo, paquetes por minuto
- Depende de:
 - I. Ancho de banda del medio físico
 - I. Cable
 - II. Ondas electromagnéticas
 - III. Fibra óptica
 - IV. Líneas de cobre
 - II. Capacidad de procesamiento de elementos transmisores
 - III. Eficiencia de los algoritmos de acceso al medio
 - IV. Mecanismos de codificación de canal
 - v. Mecanismos de compresión de datos



Capacidad efectiva del canal

► Fracción de la capacidad nominal

$\text{CapacidadEfectiva} = N \times \text{CapacidadNominal}$
(Siempre $N < 1$)

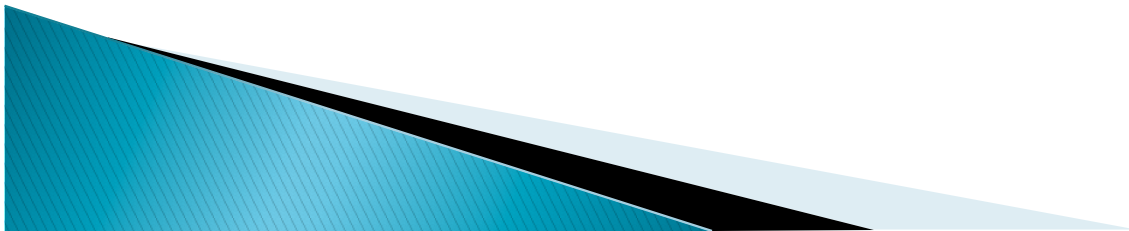
► Afectado por muchos factores

- Carga adicional de procesamiento en las varias capas OSI
- Limitaciones de procesamiento en dispositivos
 - Memoria, CPU, otros
- Eficiencia del protocolo de transmisión
 - Control de flujo
 - Enrutamiento



Utilización del canal

- ▶ Fracción de la capacidad nominal de un canal que está siendo realmente utilizada
 - Planificación:
 - Qué tasa de crecimiento tiene la demanda?
 - Cuándo comprar más capacidad?
 - Dónde invertir en actualizaciones?
 - Resolución de problemas:
 - Detectar puntos de baja capacidad (“cuellos de botella”)

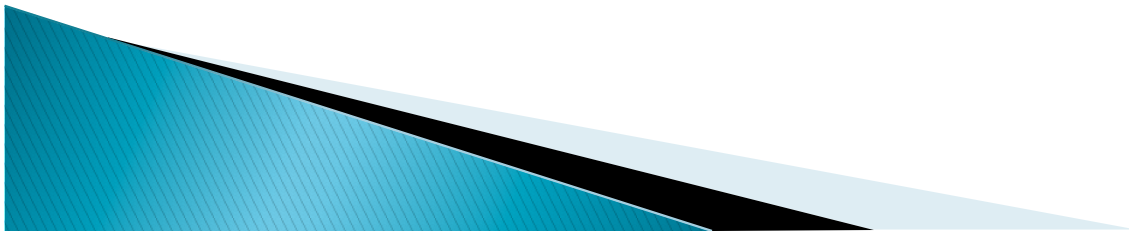


95-Percentile

Percentile: Valor máximo de cierto porcentaje de una muestra

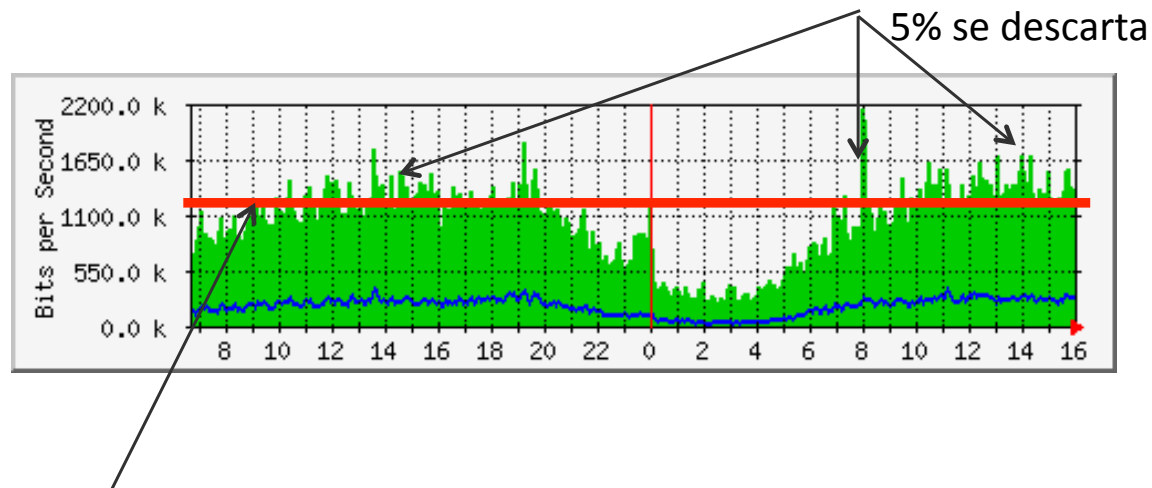
95-percentile de uso del canal:

- Se usa comunmente como medida de la utilización del canal
 - En el 95% de las muestras, el valor observado es igual o menor que este valor.
 - El 5% de las muestras restantes se descartan, suponiendo anomalías de uso.
-
- Importancia en las redes de datos
 - Medida de la utilización sostenida del canal
 - Muchos ISPs la utilizan como medida para facturación de servicios



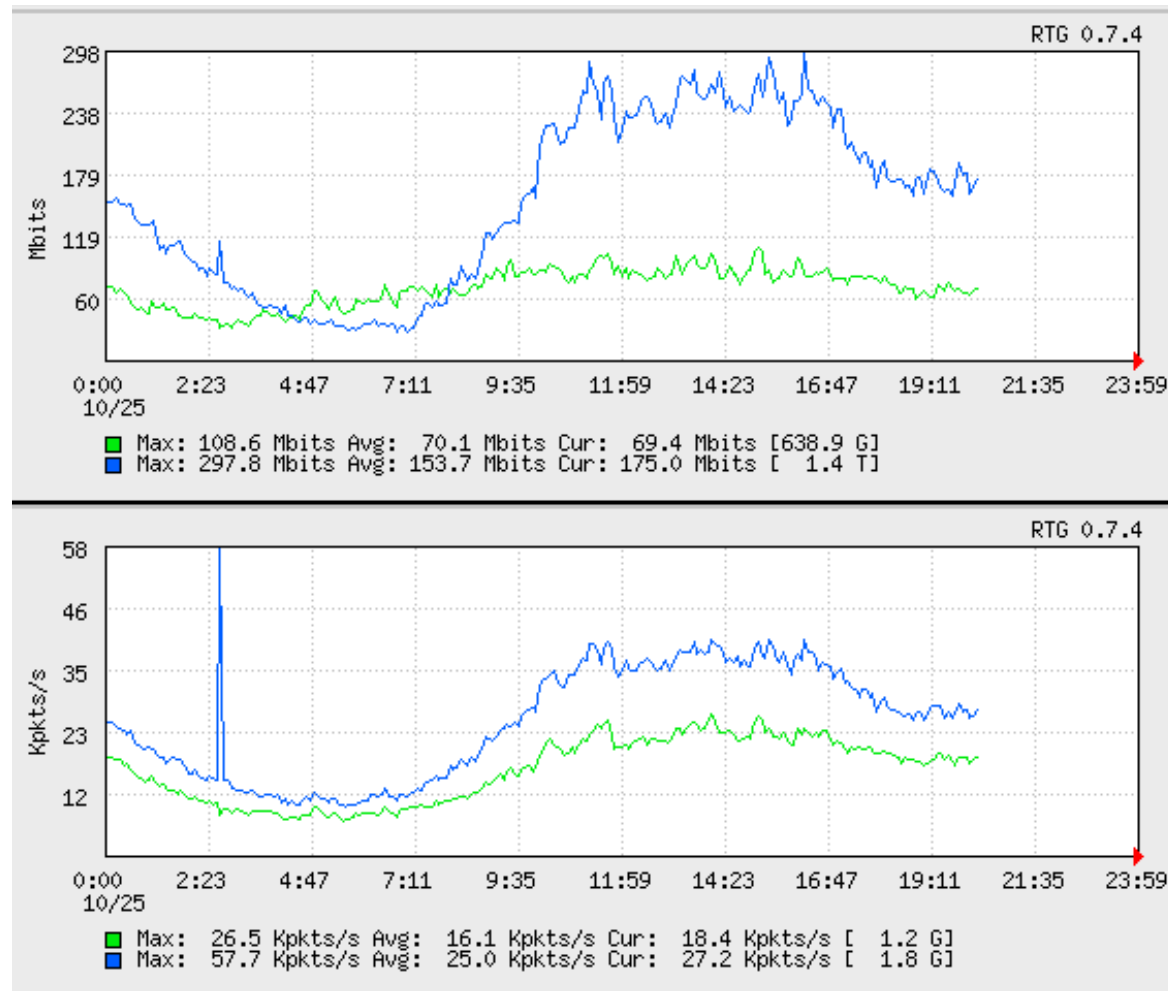
Facturación 95-Percentile

- Proveedor toma muestras de consumo de ancho de banda
- Tan frecuente como cada 5 minutos
- Al final del mes, el 5% de los valores más altos se descarta
- El más alto de los valores restantes (N) es el uso medido a 95-percentile
 - Puede ser valor de entrada o salida (se mide en ambas direcciones)
- Se factura al cliente: $N \times \text{Costo por unidad}$
 - (ejemplo: 40 Mbit/s $\times$ \$30.00 Mbits = \$120.00 al mes)



El valor más alto despues de descartar el 5% de los valores más altos se convierte en la medida de uso de ancho de banda a facturar

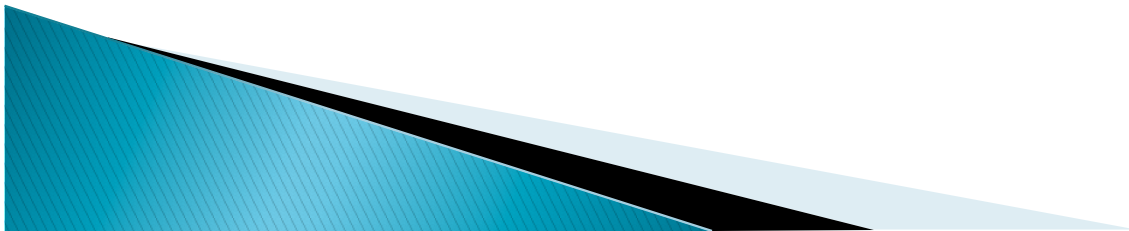
bps vs. pps => tamaño del paquete



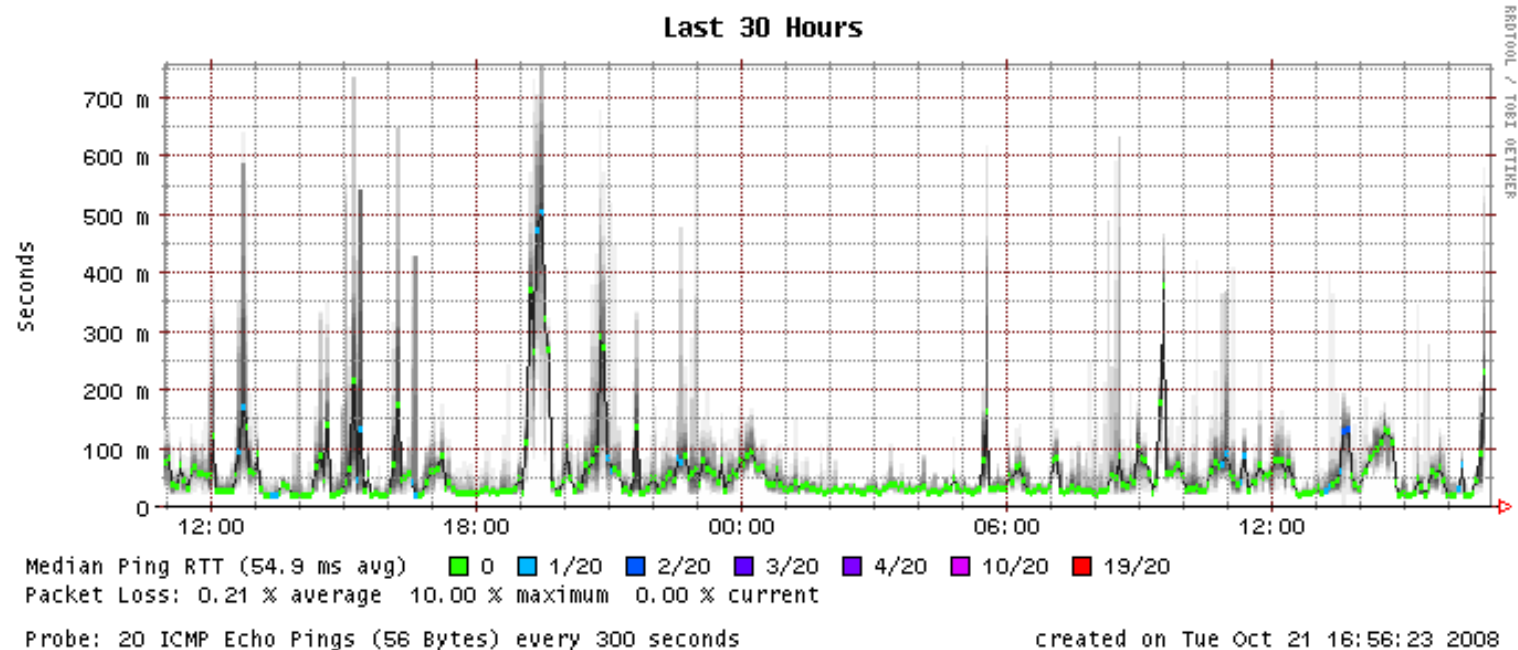
Retardo extremo-a-extremo

El tiempo transcurrido en transmitir un paquete de fuente a destino final

- producido por una aplicación, entregado al sistema operativo,
 - pasado a la tarjeta de red, codificado,
 - transmitido por el medio físico,
 - recibido por un equipo intermedio (switch, router), analizado,
 - retransmitido en otro medio...etc., etc.
-
- La medición más común es de ida y vuelta (RTT)
 - El utilitario *ping* se usa para medir esta variable

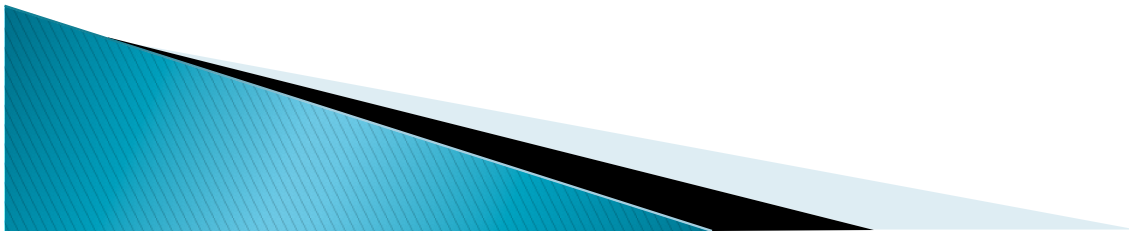


Medición histórica de Retardo



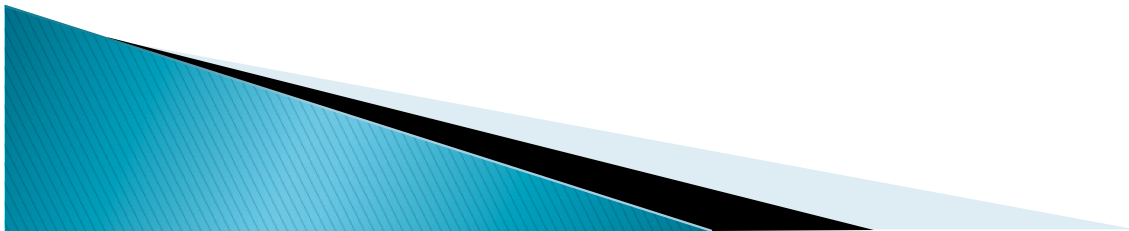
Tipos de Retardo

- Componentes del retardo extremo a extremo:
 - Retardo de Procesamiento
 - Retardo de Colas
 - Retardo de Transmisión
 - Retardo de Propagación



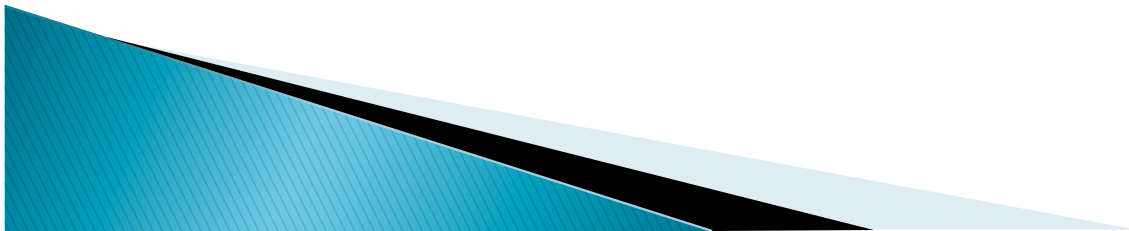
Retardo de Procesamiento

- Tiempo requerido en analizar el encabezado y decidir a dónde enviar el paquete (ej. decisión de enrutamiento)
 - En un enrutador, depende de
 - número de entradas en la tabla de rutas,
 - implementación (estructuras de datos),
 - recursos del dispositivo
- Puede incluir la verificación de errores



Retardo de Colas

- Tiempo en que el paquete espera en un *búfer* hasta ser transmitido
- El número de paquetes esperando en cola dependerá de la intensidad y la naturaleza del tráfico
- Los algoritmos de colas en los enrutadores intentan adaptar estos retardos a ciertas preferencias, o imponer un uso equitativo



Retardo de Transmisión

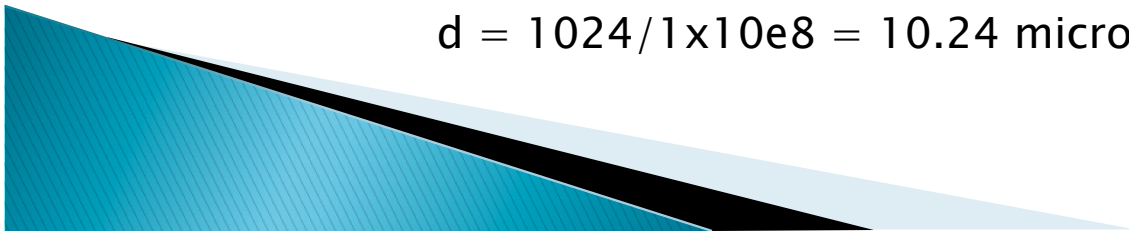
Tiempo requerido para pasar todos los bits de un paquete a través del medio de transmisión

$$d = L/R$$

- R= tasa de bits, (o velocidad de transferencia de datos)
- L=Longitud del paquete,
- d = retardo

Para transmitir 1024 bits utilizando Fast Ethernet (100 Mbps):

$$d = 1024/1 \times 10^8 = 10.24 \text{ microsegundos}$$



Retardo de Propagación

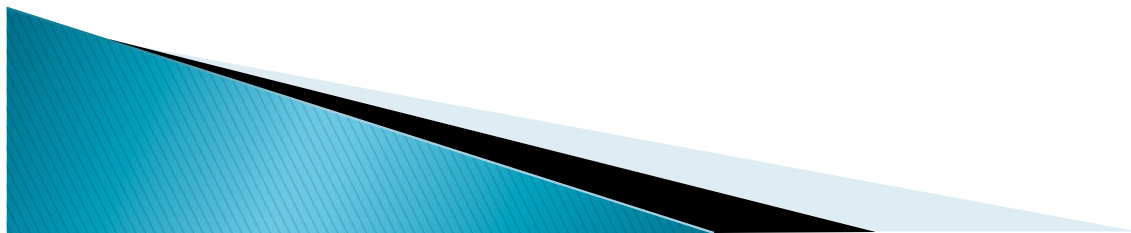
Una vez que el bit de dato entra al medio físico, el tiempo transcurrido en su propagación hasta el final del medio

- ▶ La velocidad de propagación del enlace depende fundamentalmente de la longitud del medio físico
- ▶ Velocidad cercana a la velocidad de la luz en la mayoría de los casos

$$D_p = d/s$$

d = distancia,

s = velocidad de propagación

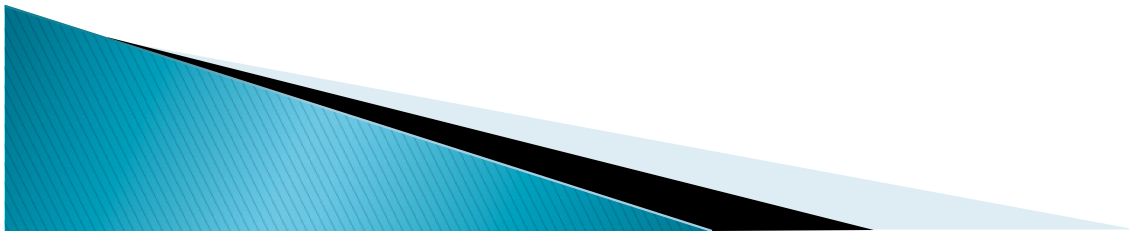


Transmisión vs. Propagación

- Puede ser confuso al principio

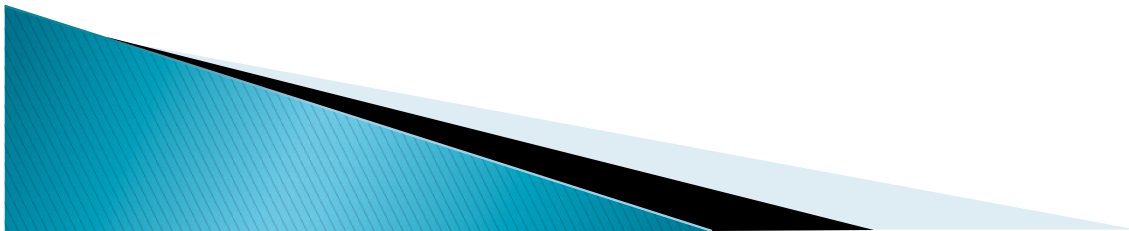
Ejemplo:

- Dos enlaces de 100 Mbps:
 - Via fibra óptica de 1 Km de longitud
 - Via satélite, con una distancia de 30Km entre estación base y satélite
- Para dos paquetes del mismo tamaño, cuál tiene mayor retardo de transmisión? Y propagación?

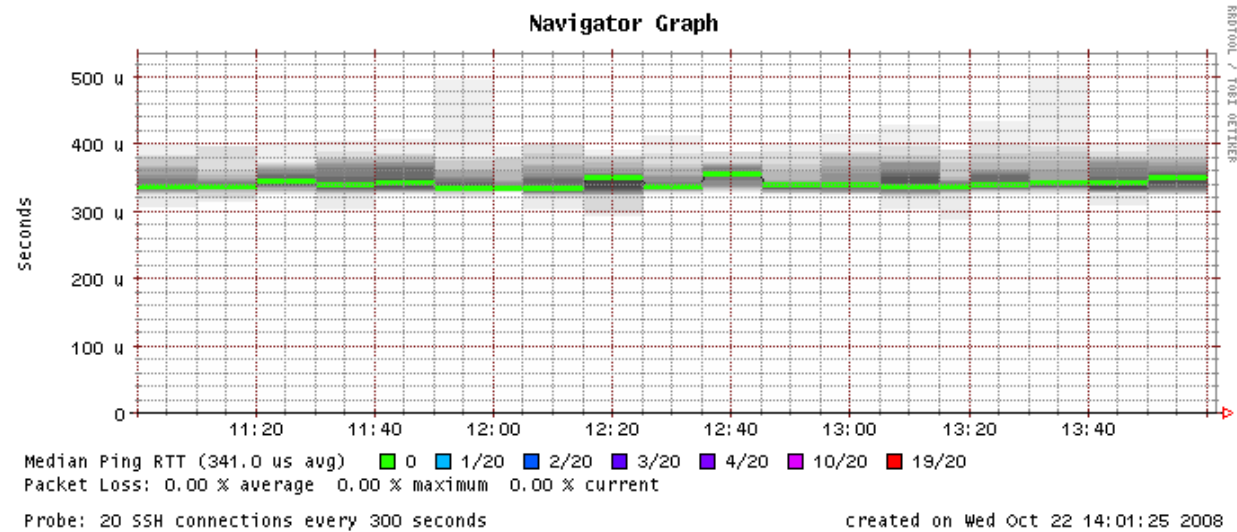
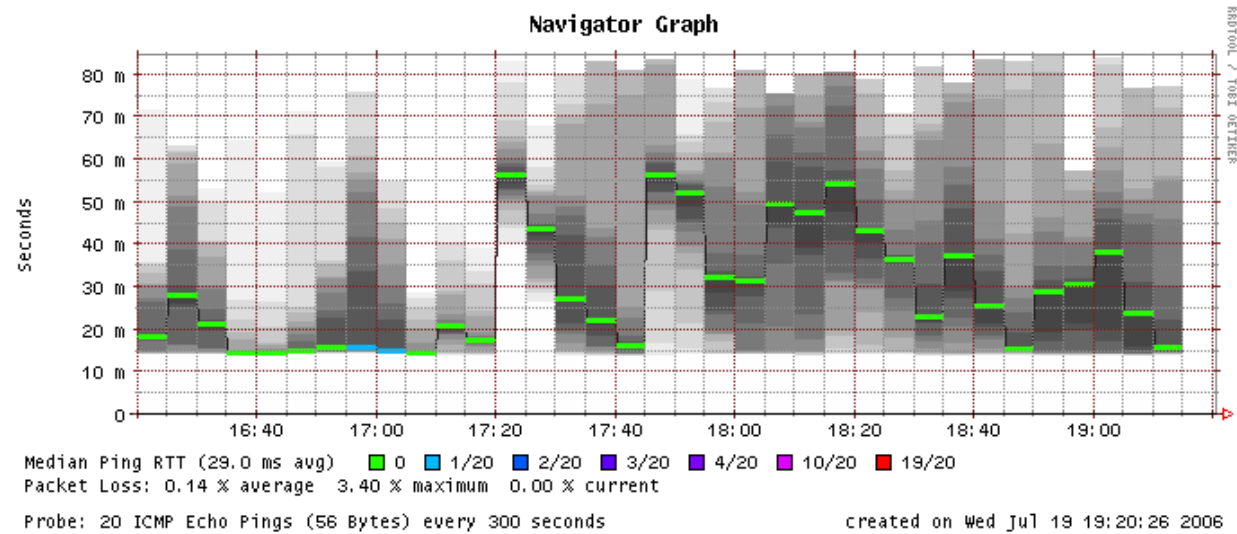


Pérdida de paquetes

- Ocurren por el hecho de que las colas (búfers) no son infinitas
 - Cuando un paquete llega a una cola y ésta está llena, el paquete se descarta.
 - La pérdida de paquetes, si ha de ser corregida, se resuelve en capas superiores (transporte o aplicación)
 - La corrección de pérdidas, usando retransmisión, puede causar aún más congestión si no se ejerce algún tipo de control

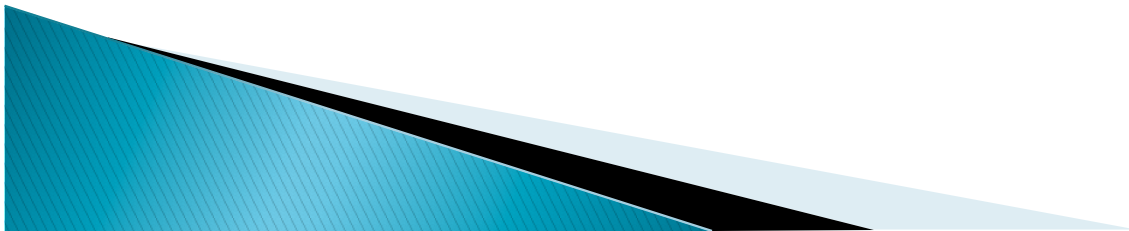


Jitter



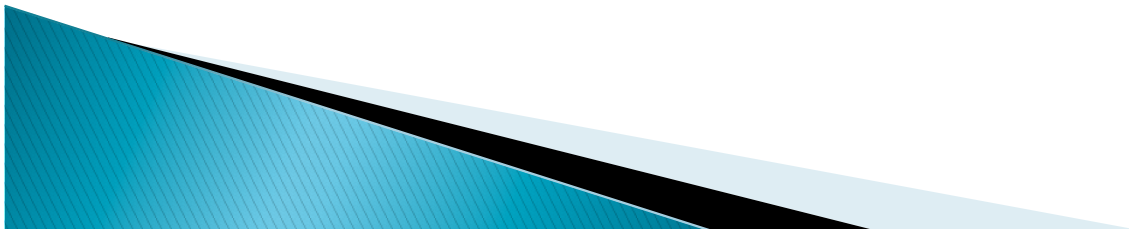
Control de Flujo y Congestión

- Limitar la tasa de envío porque el receptor no puede procesar los paquetes a la misma velocidad que los recibe
- Limitar la tasa de envío del emisor porque existen pérdidas y retardos en el trayecto



Controles en TCP

- IP implementa un servicio no-orientado a conexión
 - No existe ningún mecanismo en IP que resuelva las causas de la pérdida de paquetes
- TCP implementa control de flujo y congestión
 - En los extremos, porque los nodos intermedios en la capa de red no hablan TCP



Flujo vs. Congestión en TCP

Flujo: controlado por los tamaños de ventana (RcvWindow) enviados por el receptor

Congestión: controlado por el valor de ventana de congestión (CongWin)

- Mantenido independientemente por el emisor
- Varía de acuerdo a la detección de paquetes perdidos
 - Timeout o la recepción de tres ACKs repetidos
- Comportamientos:
 - Incremento aditivo / Decremento multiplicativo (AIMD)
 - Comienzo lento (Slow Start)
 - Reacción a eventos de *timeout*
- Ver: RFC 2581, y 2001
- *"Internetworking with TCP/IP" – Douglas Comer*



Diferentes algoritmos de Control de Congestión en TCP

