Traffic Analysis

- Introducción -

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

Programa de Tecnologías para la gestión distribuida de la información

Contenido

- Ejemplos introductorios
- Introducción al *Teletraffic Engineering*
- *Traffic Measurement*
- *Internet Traffic*
- Descripción del tráfico

Ejemplos introductorios

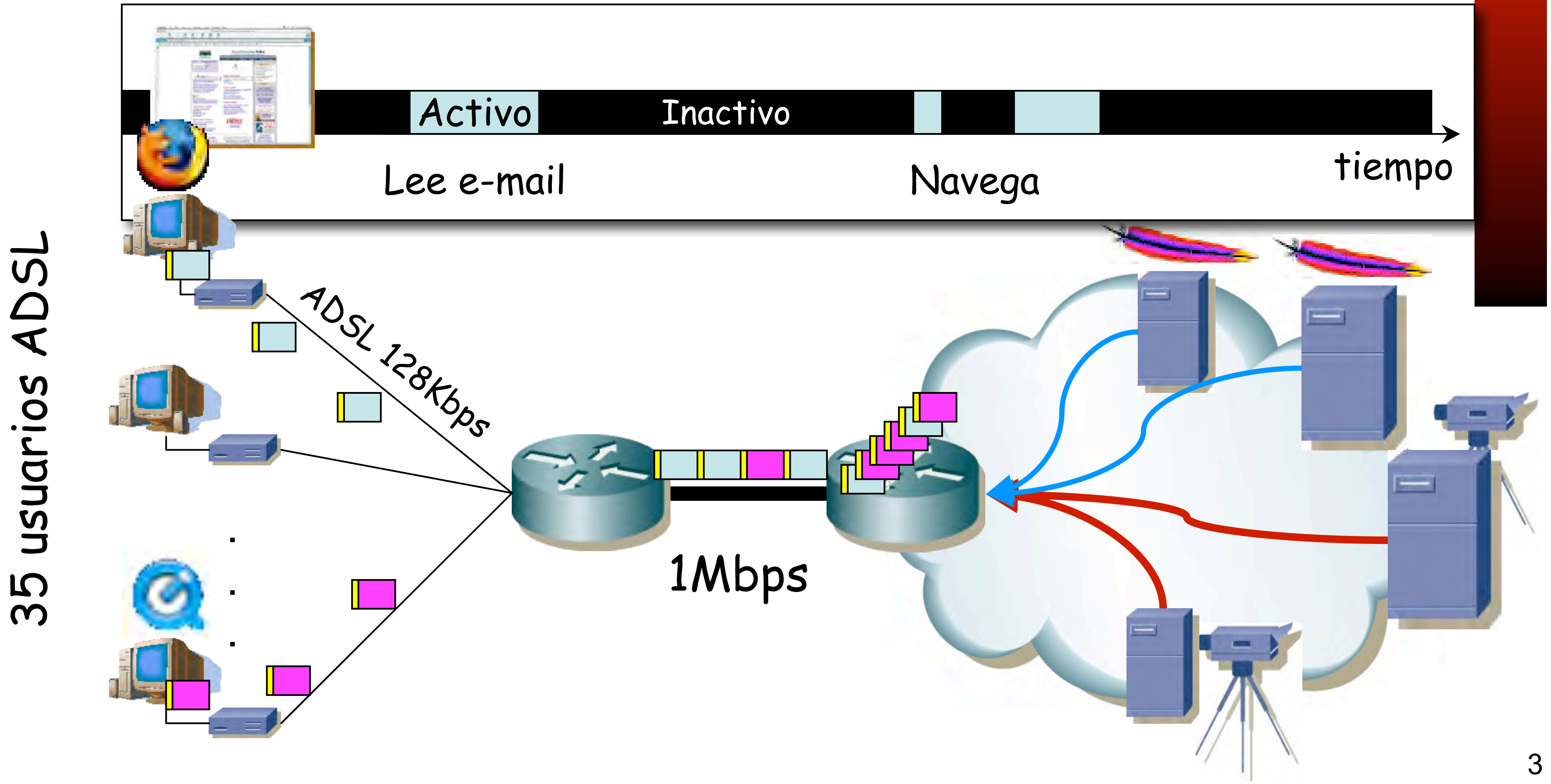
Ejemplo de red

Cada usuario:

- Recibe de un servidor a 100Kbps cuando está activo
- Activo cada uno un 10% del tiempo

10 usuarios a 100Kbps = 1Mbps

¿ Cuál es la probabilidad de que más de 10 usuarios reciban tráfico a la vez ?



Ejemplo de red

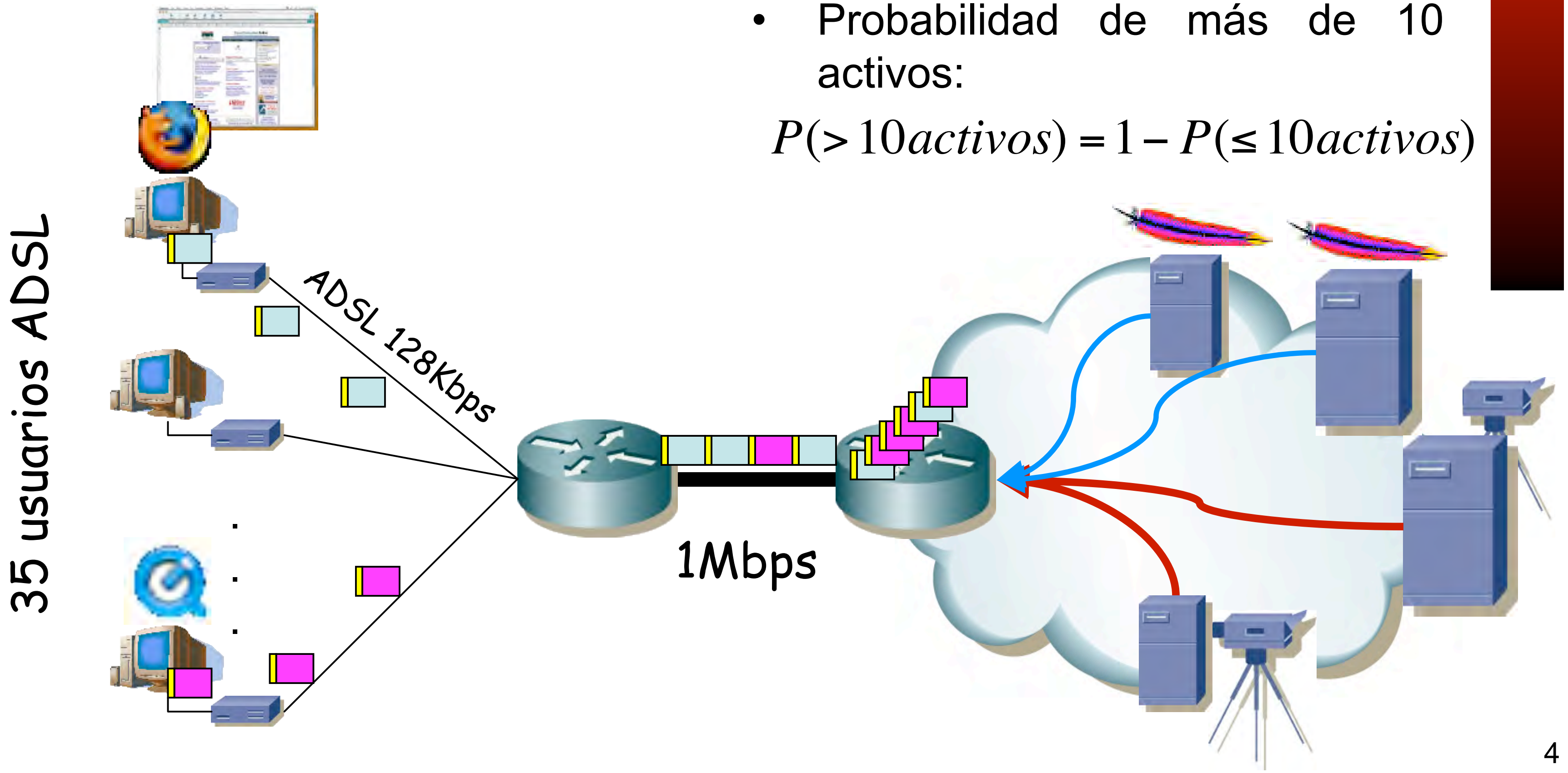
¿ Cuál es la probabilidad de que más de 10 usuarios reciban tráfico a la vez ?

- Usuario activo un 10% del tiempo
- Supongamos pues que en un momento cualquiera:

$$P(usuario\_activo) = 0.1 = p$$

- Probabilidad de más de 10 activos:

$$P(>10activos) = 1 - P(\leq 10activos)$$

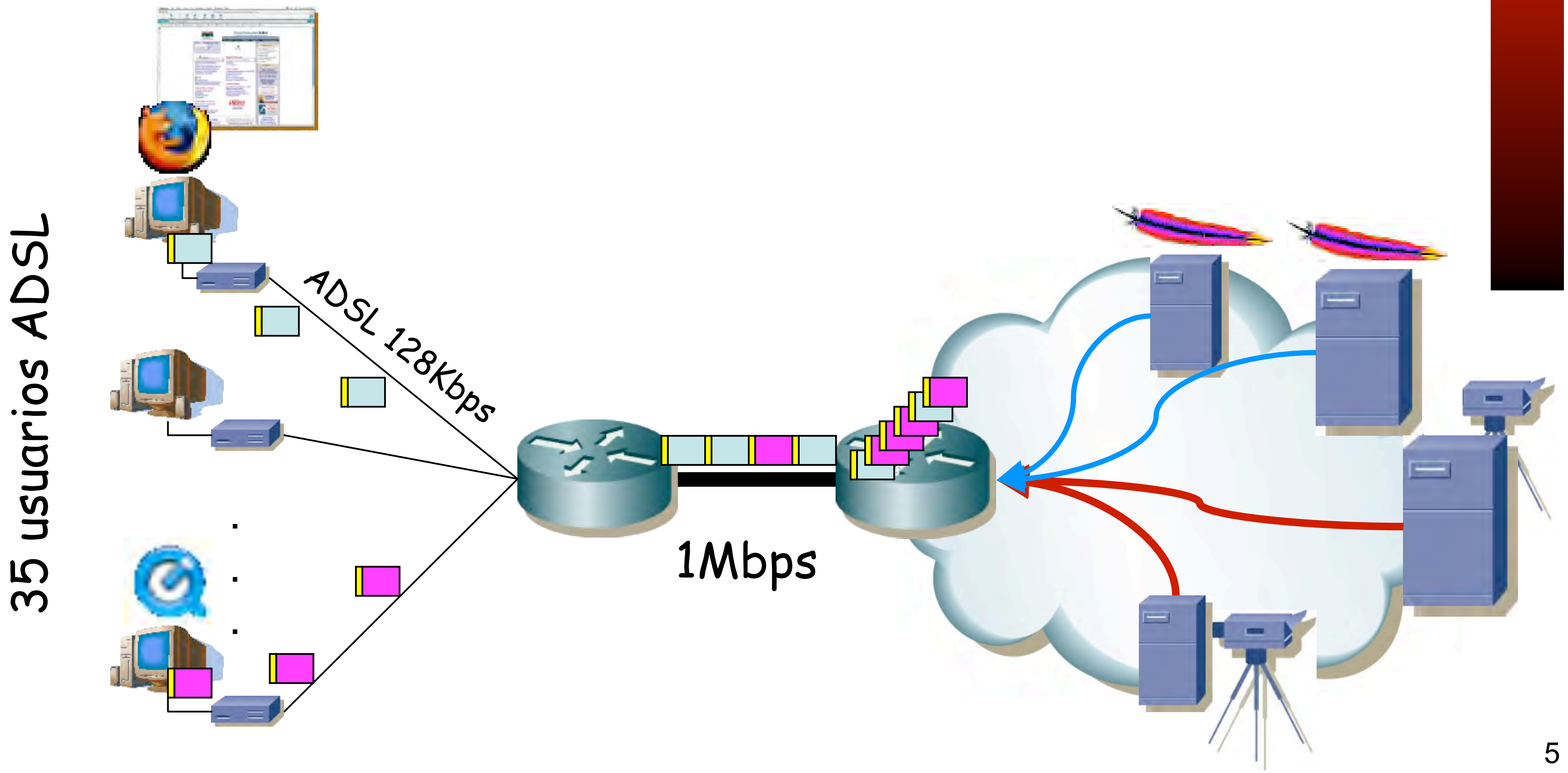


Ejemplo de red

¿ Cuál es la probabilidad de que más de 10 usuarios reciban tráfico a la vez ?

$$P(>10activos) = 1 - P(\leq 10activos)$$

$$P(\leq 10activos) = P(0\_activos) + P(1\_activo) + \ldots + P(10\_activos) = \sum_{i=0}^{10} P(i\_activos)$$



Ejemplo de red

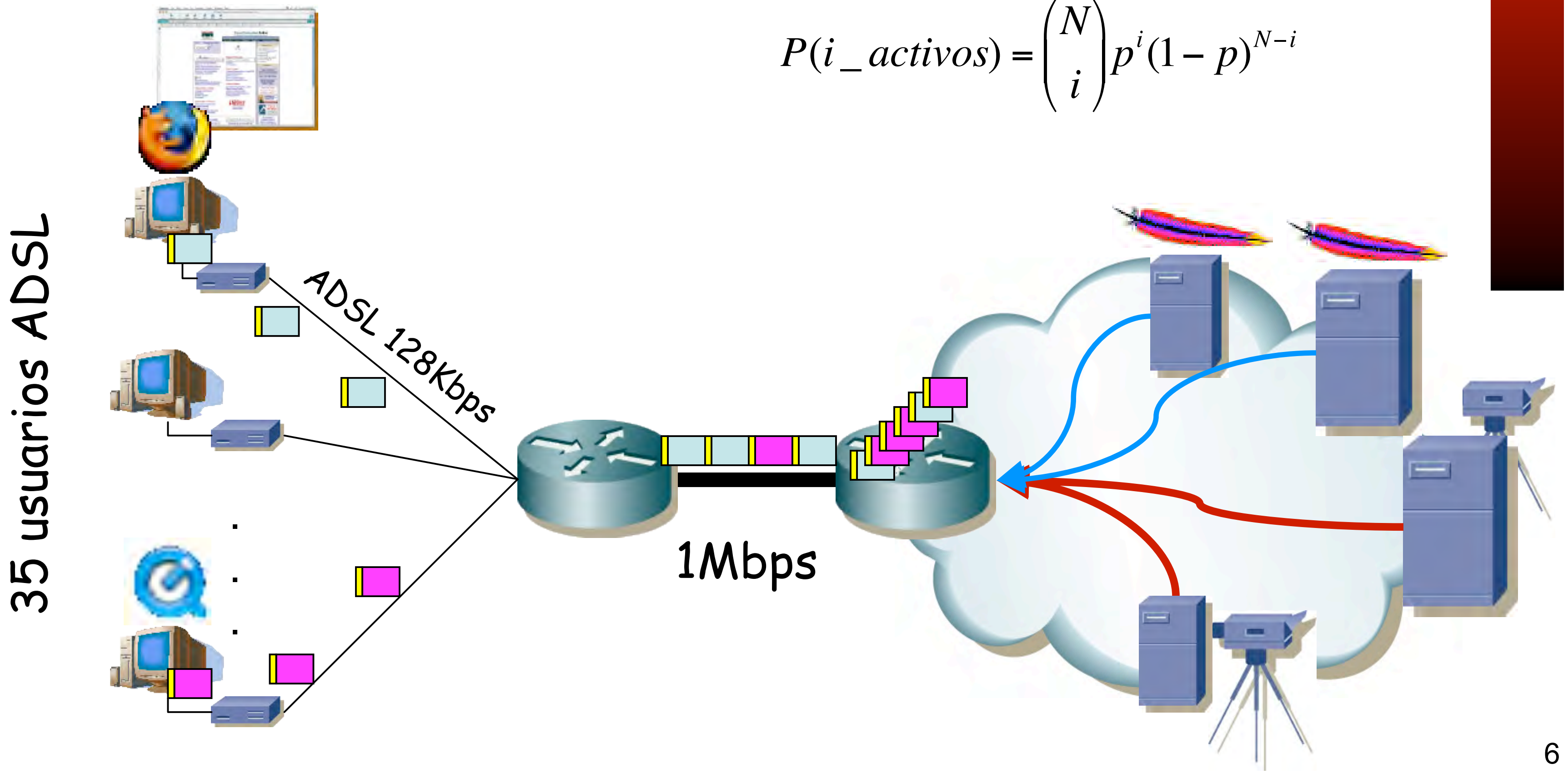
¿ Cuál es la probabilidad de que más de 10 usuarios reciban tráfico a la vez ?

$$P(0\_activos) = (1-p)^N$$

$$P(1\_activo) = Np(1-p)^{N-1}$$

$$P(2\_activos) = \frac{N(N-1)}{2} p^2 (1-p)^{N-2}$$

$$P(i\_activos) = \binom{N}{i} p^i (1-p)^{N-i}$$



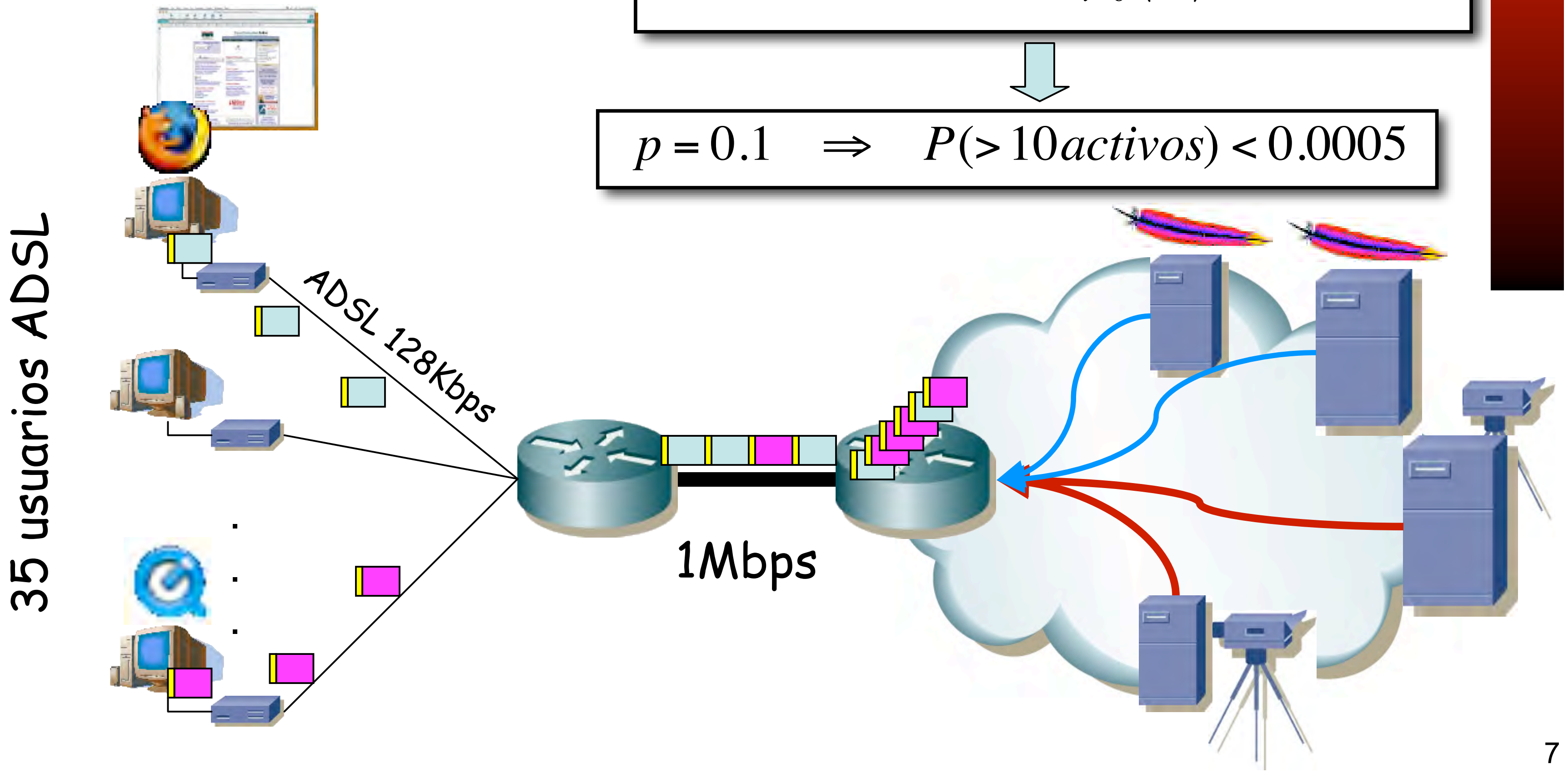
Ejemplo de red

¿ Cuál es la probabilidad de que más de 10 usuarios reciban tráfico a la vez ?

$$P(\leq 10 activos) = \sum_{i=0}^{10} \binom{N}{i} p^i (1-p)^{N-i}$$

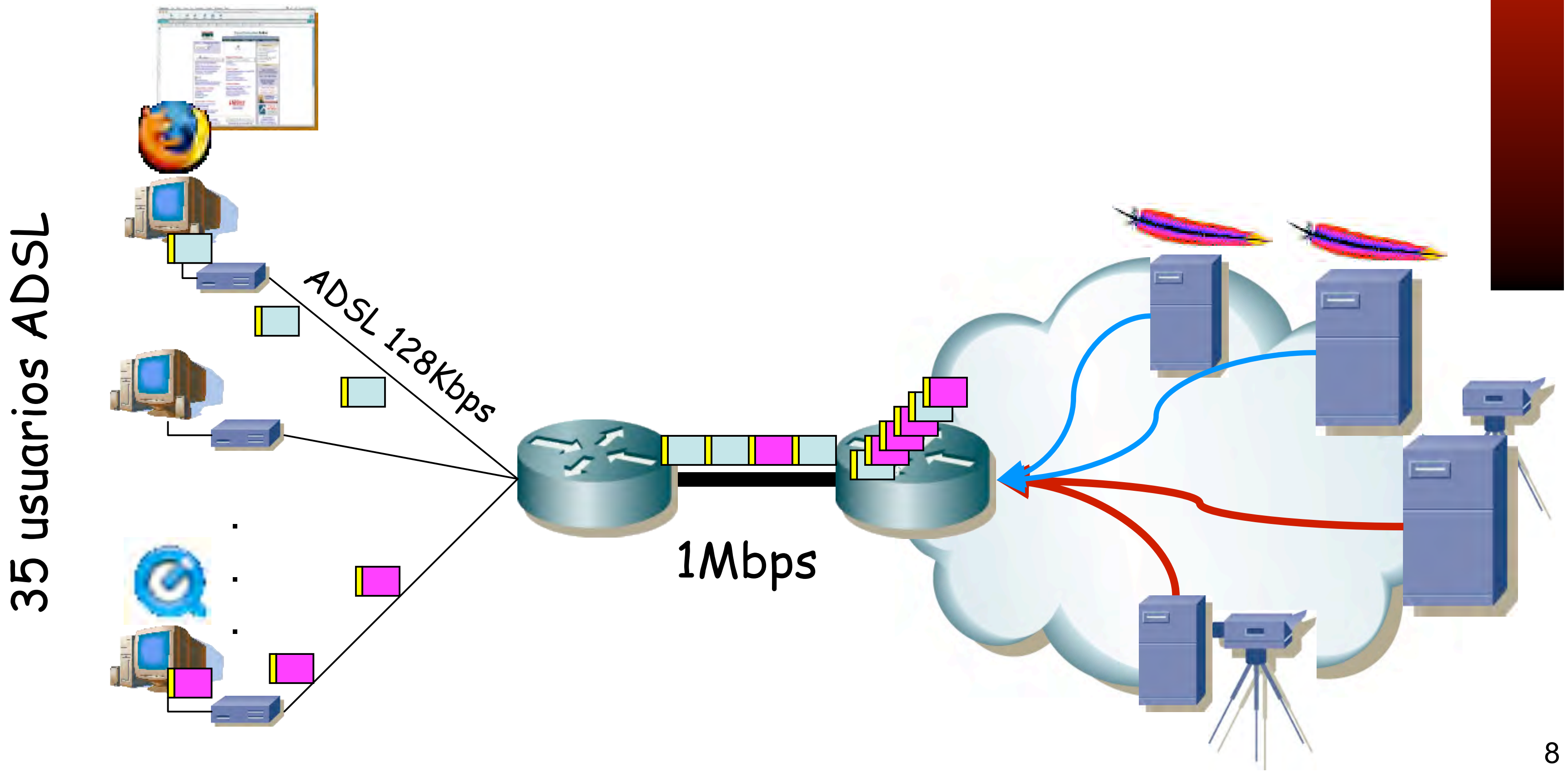
$$P(> 10 activos) = 1 - \sum_{i=0}^{10} \binom{N}{i} p^i (1-p)^{N-i}$$

$$p = 0.1 \quad \Rightarrow \quad P(> 10 activos) < 0.0005$$



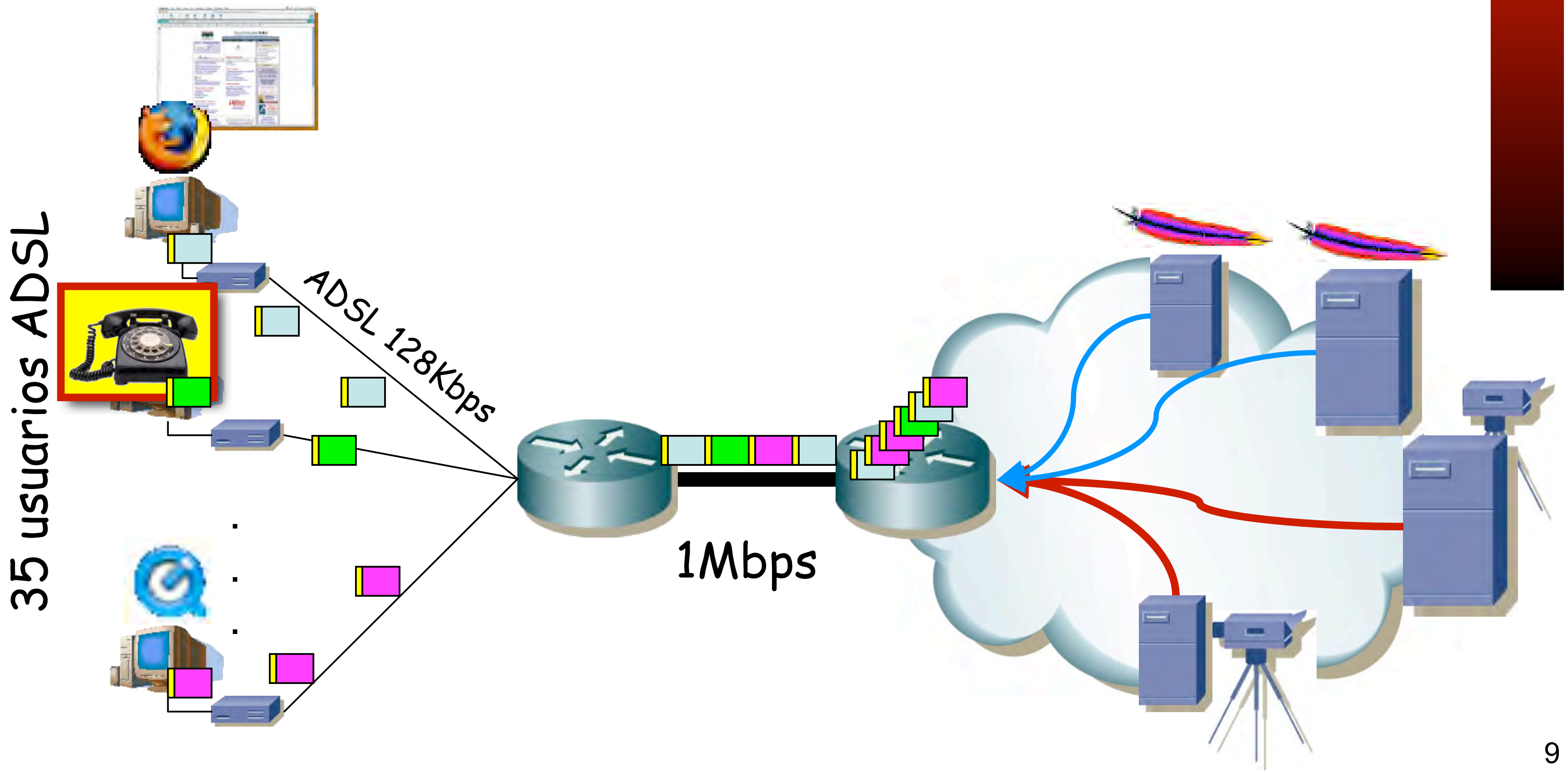
Ejemplo de red

- 35 usuarios x 128 Kbps/usuario = 4,48Mbps
- 4,48Mbps > 1Mbps
- Congestión en enlace de acceso sin dar 128Kbps a todos los usuarios
- *Sobresuscripción* (overbooking)



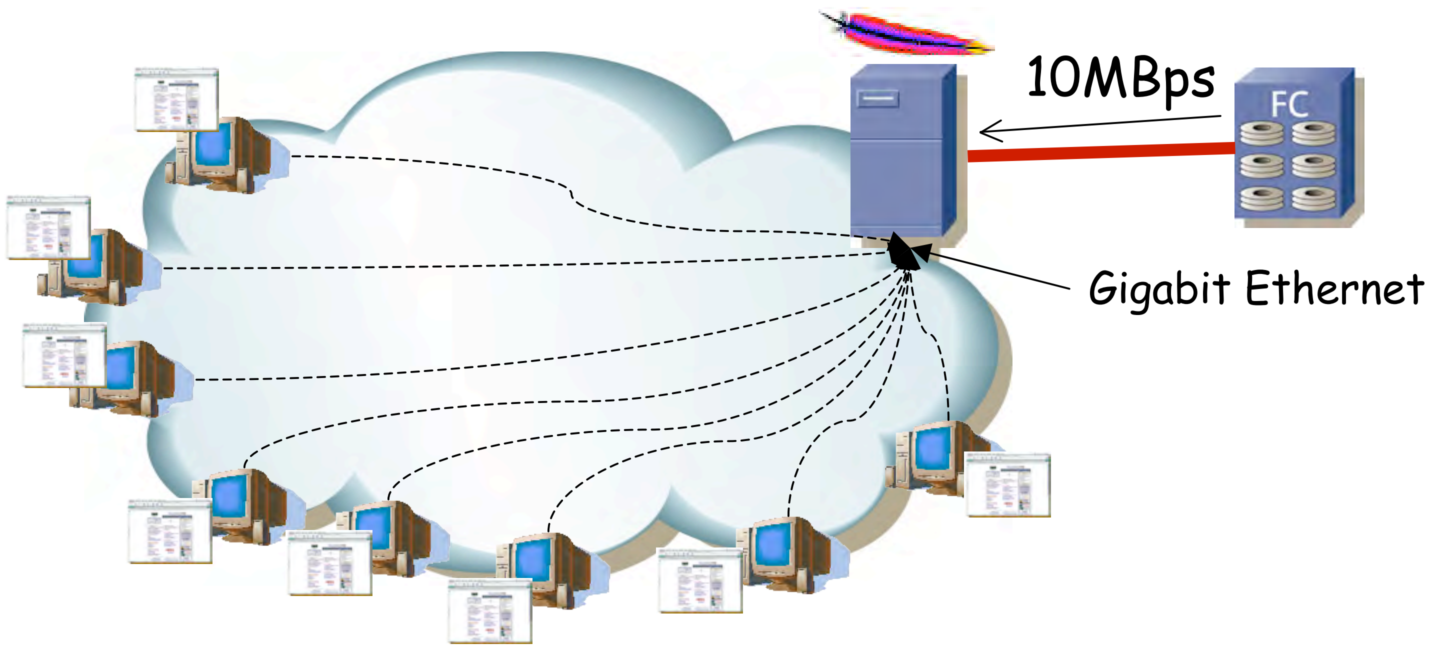
Ejemplo de red

- Si ahora un usuario quiere emplear una aplicación de voz
- Pérdidas, excesivo retardo
- Necesitaríamos implementar un mecanismo de QoS
- Pero también en ese caso tendríamos el problema, solo que con las fuentes con prioridad



Ejemplo de aplicación

- Peticiones a un servidor web
- 1.000 peticiones por segundo ⇒ nueva petición cada 0,001 segs
- Tamaños de los ficheros 100KBytes
- Discos sirven a 10MBps (80Mbps) ⇒ 1 fichero servido en 0,01 seg ⇒ 100 ficheros servidos por segundo
- ¡ Hay 10 veces más peticiones por segundo que las que soportan los discos !



Teletraffic Engineering

Teletraffic Engineering

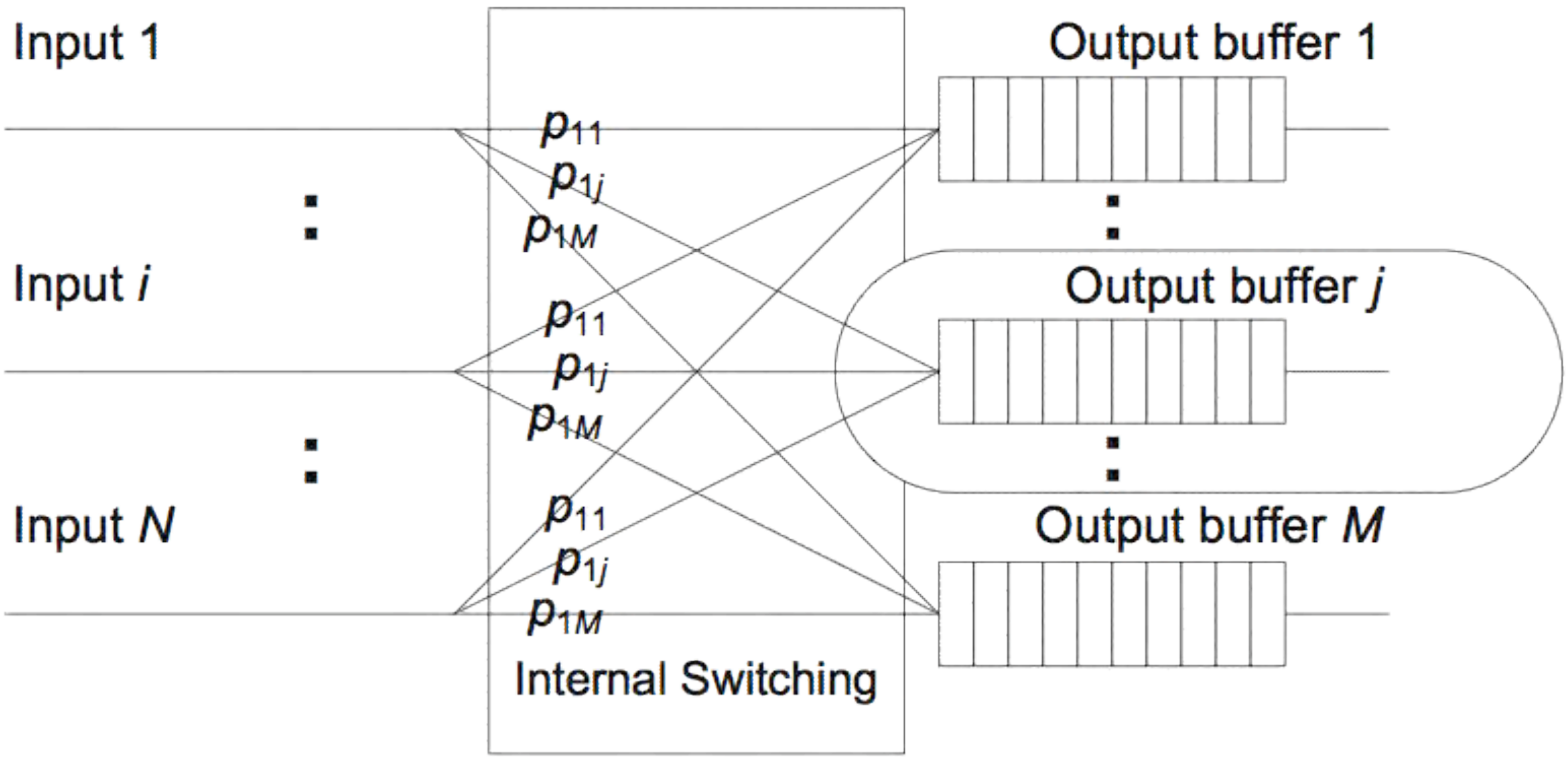
- “*The application of probability theory to the solution of problems concerning planning, performance evaluation, operation, and maintenance of telecommunication systems*”
- Nos interesa su aplicación a systemas de comunicaciones (de datos y telefónicos)
- Pero aplica también a tráfico rodado y aéreo, cadenas de montaje, distribución, gestión de almacén, etc.
- Multidisciplinar:
  - Teoría de probabilidad
  - Teoría de procesos estocásticos
  - Estadísticas
  - Queueing theory
  - Simulación
  - Modelado de tráfico
  - Optimización
  - Fiabilidad

Teletraffic Theory

- Para proveedores de servicio
  - ¿Cómo distribuir los puntos de acceso al servicio?
  - ¿Cuánto servidores se necesitan para satisfacer las peticiones de los usuarios?
- Operadoras de red
  - ¿Cómo distribuir la carga de red?
  - ¿Cuánto buffering requieren los conmutadores/routers?
  - ¿Cuáles son las capacidades óptimas?
- Fabricantes
  - ¿Cómo utilizar de mejor manera los recursos del equipamiento de conmutación/enrutado?
  - ¿Qué mejoras se debería hacer al mismo?
- Usuarios finales
  - ¿Qué calidad de servicio voy a obtener de la red?

Problema clásico

- Dimensionar el buffer de un ruter dada una carga:
  - Determinar tamaño del buffer y velocidad del enlace de salida
  - Asumir encaminamiento estático



Problema clásico: Procedimiento

- Tráfico de entrada a cada enlace
  - Medirlo
  - Determinar estadísticas importantes
  - Elegir un modelo apropiado
  - Ajustar los parámetros del modelo
- Superposición de procesos en la cola del puerto de salida
  - Usar propiedades de los modelos (ej: Poisson+Poisson=Poisson)
- Analizar la cola bajo una carga determinada
- Determinar el buffer y velocidad de enlace requeridos
  - Formular y resolver esta tarea inversa

Problema clásico: Resolución

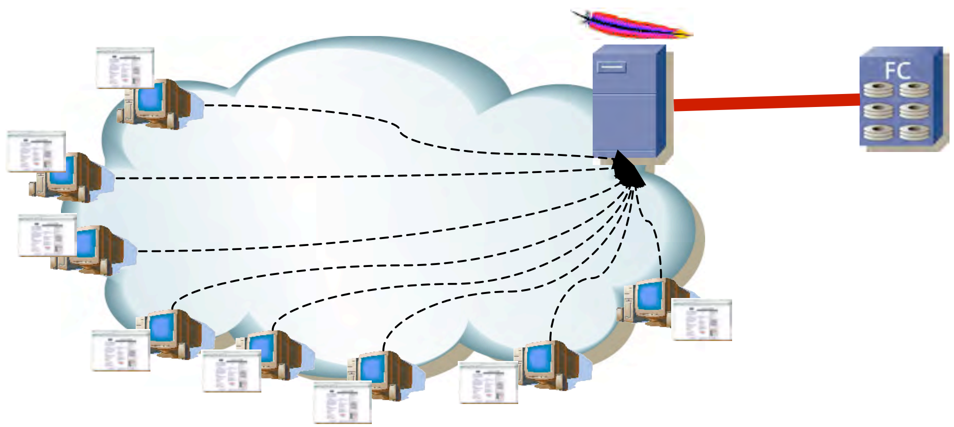
- Analítica
  - Modelado de tráfico
  - Teoría de colas
  - Optimización
- Simulación
  - Modelado de tráfico
  - Simulación
  - Ventaja: menos hipótesis restrictivas
  - Desventaja: no es adecuada para optimizaciones

Terminología y modelos

Terminología

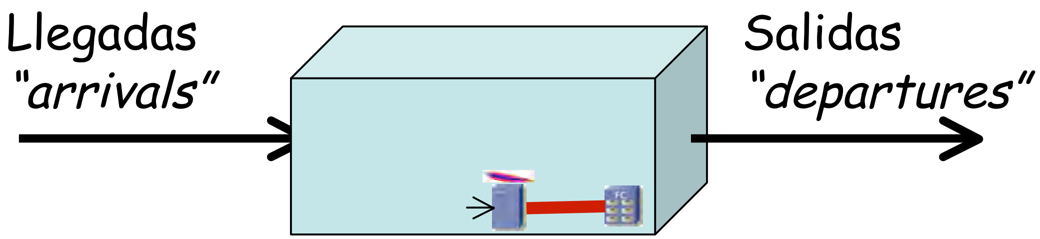
- Peticiones recibidas: *"Tasa de llegadas"* = *"Arrival rate"* = λ (1.000 peticiones/seg)
- Hemos supuesto que cada 0,001 segs una nueva: llegadas *Deterministas*
- Los discos sirven a 80Mbps
- Los tamaños de los ficheros son constantes (100KBps)
- *Capacidad del servidor:* Velocidad/Tamaños = 100 peticiones/seg = μ
- El sistema es estable si y solo si:

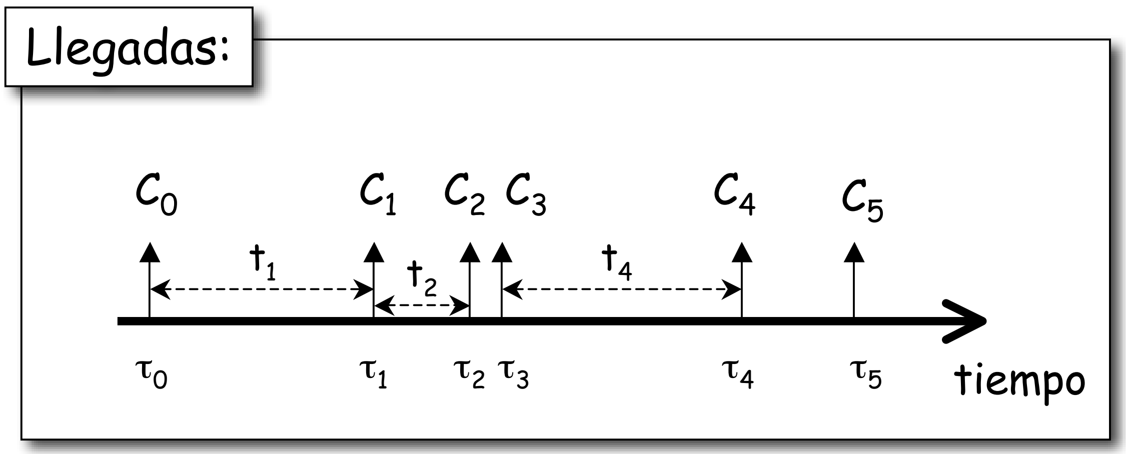
Tasa de llegadas < *Capacidad del servidor* ($\lambda<\mu$)

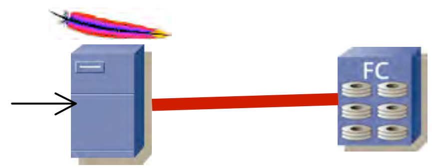


Terminología

- **C_n** : Llegada n-ésima
- **τ_n** : instante de la llegada C_n
- **t_n** : tiempo entre la llegada C_{n-1} y la C_n (**$t_n = \tau_n - \tau_{n-1}$**)
- x_n : tiempo de servicio de la llegada C_n







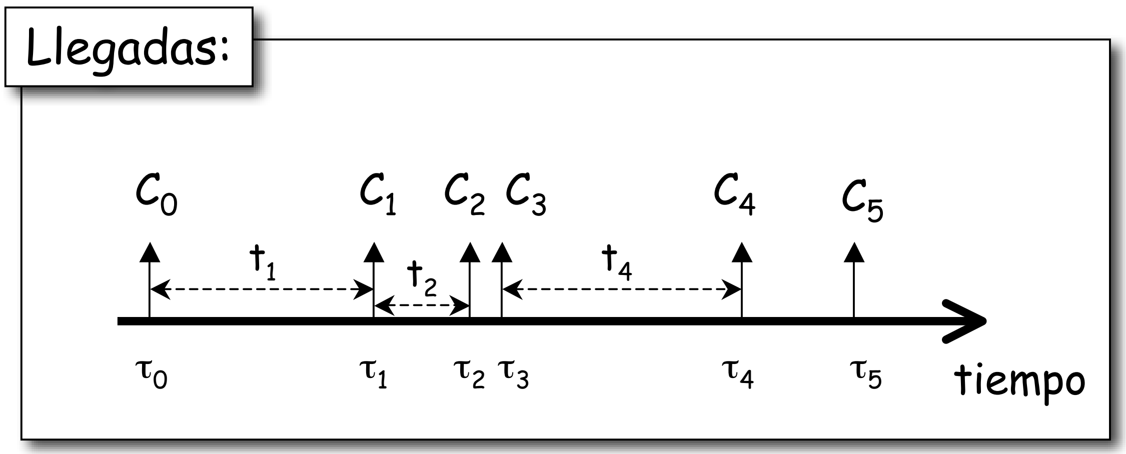
Aleatoriedad en las llegadas

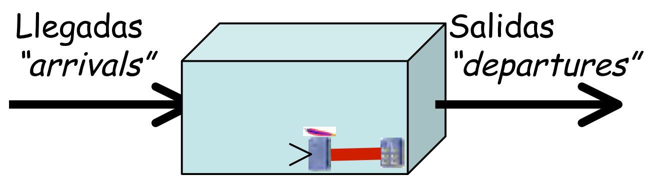
- Generalmente los instantes de llegada serán aleatorios
- Cada $\mathbf{t_n}$ es una variable aleatoria
- Forman un proceso estocástico de tiempo discreto
- Suposiciones más habituales (i.i.d.)
  - Todas las vv.aa. $\mathbf{t_n}$ siguen la misma distribución $\underline{\mathbf{t}}$
  - Todas las $\mathbf{t_n}$ son vv.aa. independientes
- **A**(t) función de distribución del tiempo entre llegadas:

  A(t) = P(tiempo entre llegadas ≤ t)

$$\lambda = \frac{1}{E[\underline{t}]}$$

nº medio de llegadas por unidad de tiempo





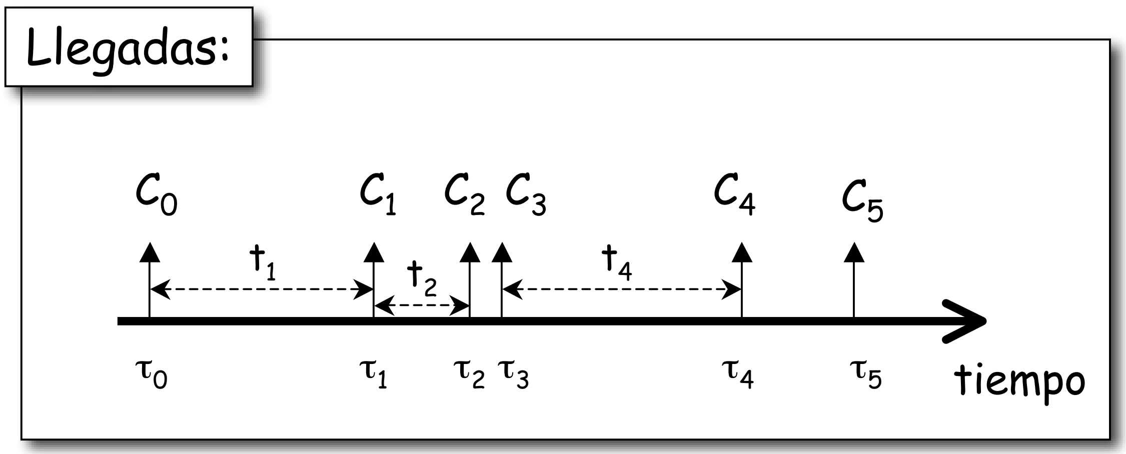
Aleatoriedad en los servicios

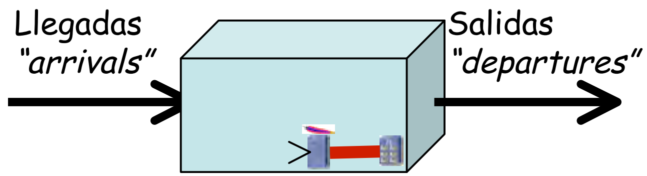
- Generalmente las duraciones de los servicios serán aleatorias
- Cada $\mathbf{x_n}$ es una variable aleatoria
- Forman un proceso estocástico de tiempo discreto
- Suposiciones más habituales (i.i.d.)
  - Todas las vv.aa. $\mathbf{x_n}$ siguen la misma distribución $\underline{\mathbf{x}}$
  - Todas las $\mathbf{x_n}$ son vv.aa. independientes
- **B**(x) función de distribución del tiempo entre llegadas:

  B(x) = P(tiempo de servicio ≤ x)

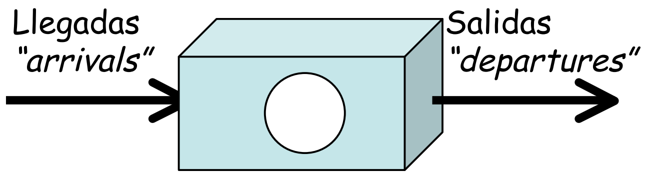
$$\mu = \frac{1}{E[\underline{x}]}$$

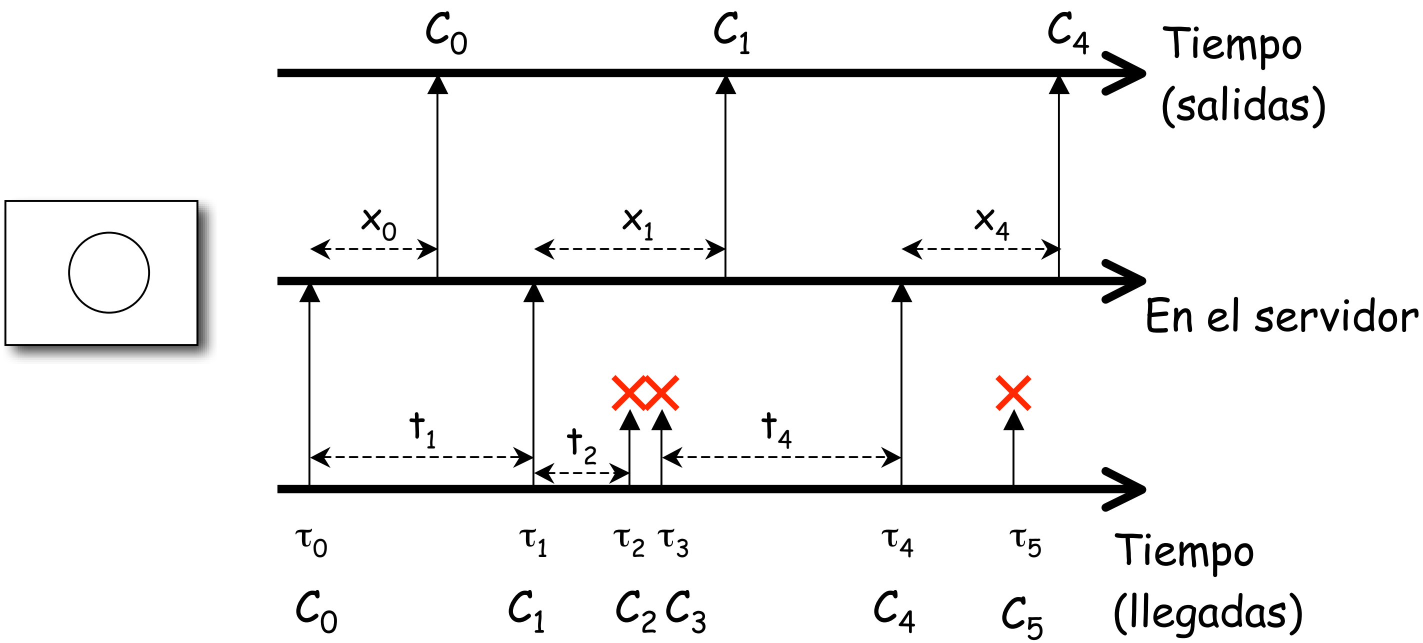
nº medio de usuarios servidos por unidad de tiempo cuando el servidor está permanentemente ocupado





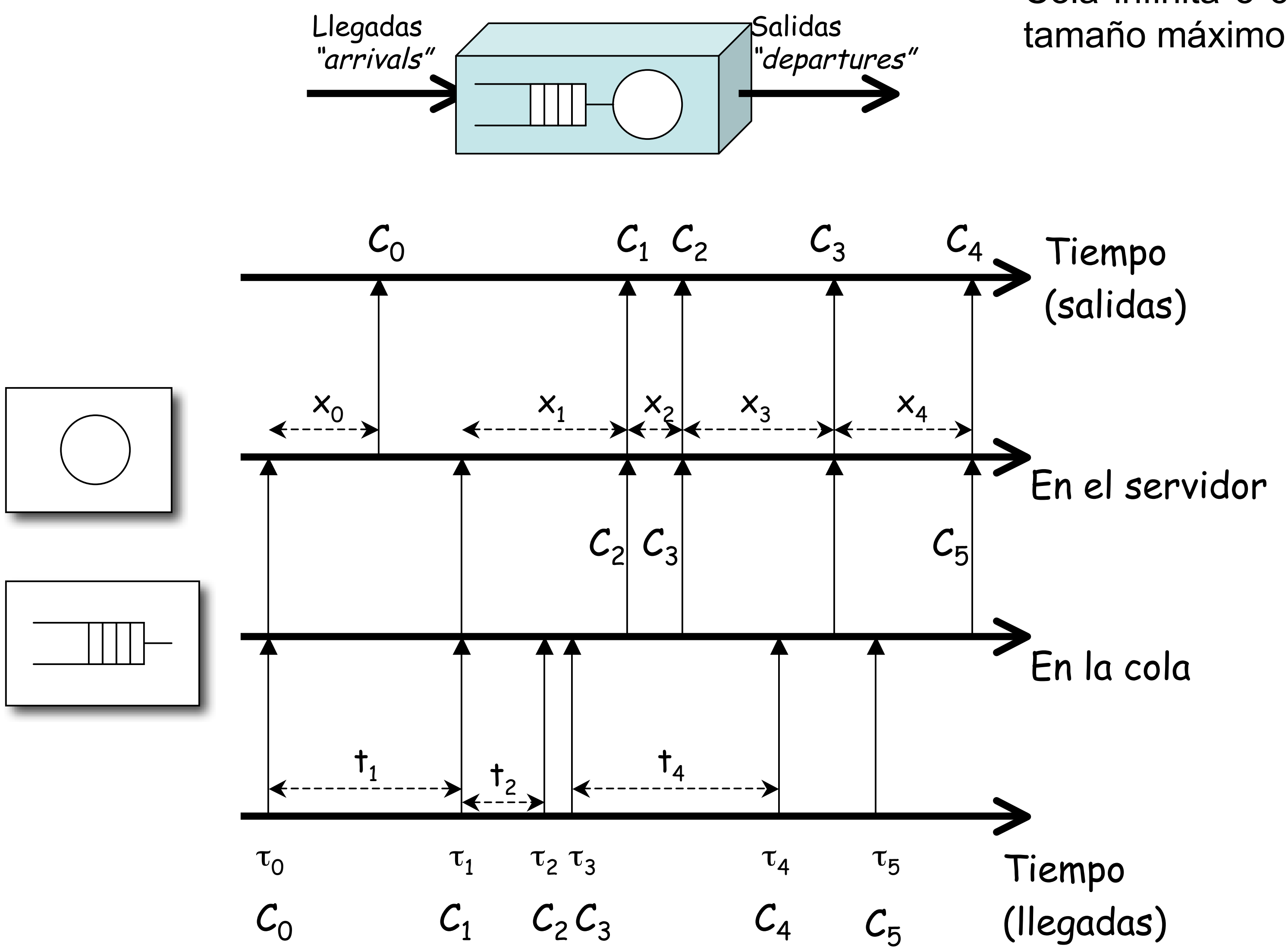
Pérdidas


Llegadas
"arrivals"
Salidas
"departures"

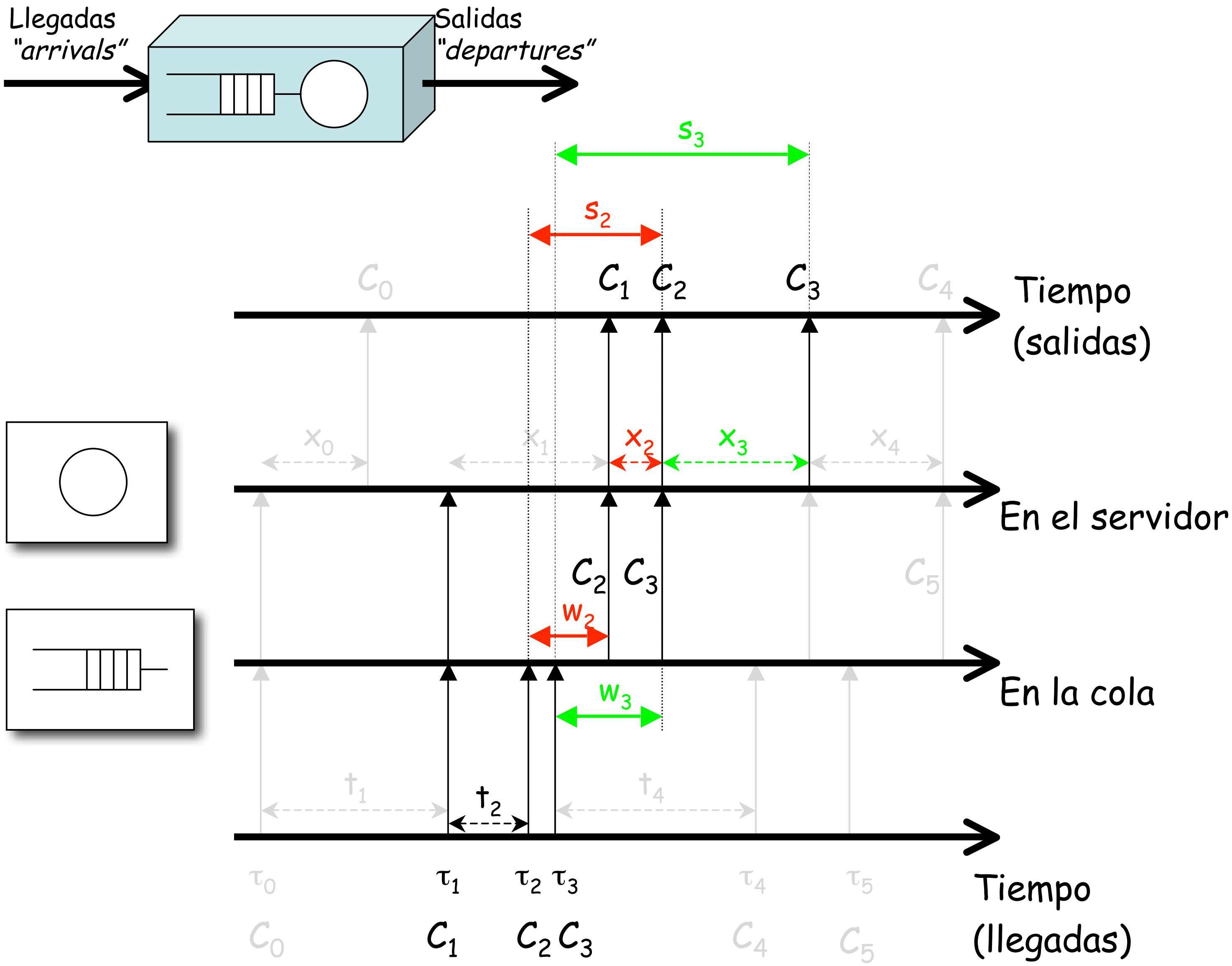

C_0 C_1 C_4
Tiempo
(salidas)
x_0 x_1 x_4
En el servidor
t_1 t_2 t_4
τ_0 τ_1 τ_2 τ_3 τ_4 τ_5
C_0 C_1 C_2 C_3 C_4 C_5
Tiempo
(llegadas)

Con cola

- Cola infinita o con tamaño máximo

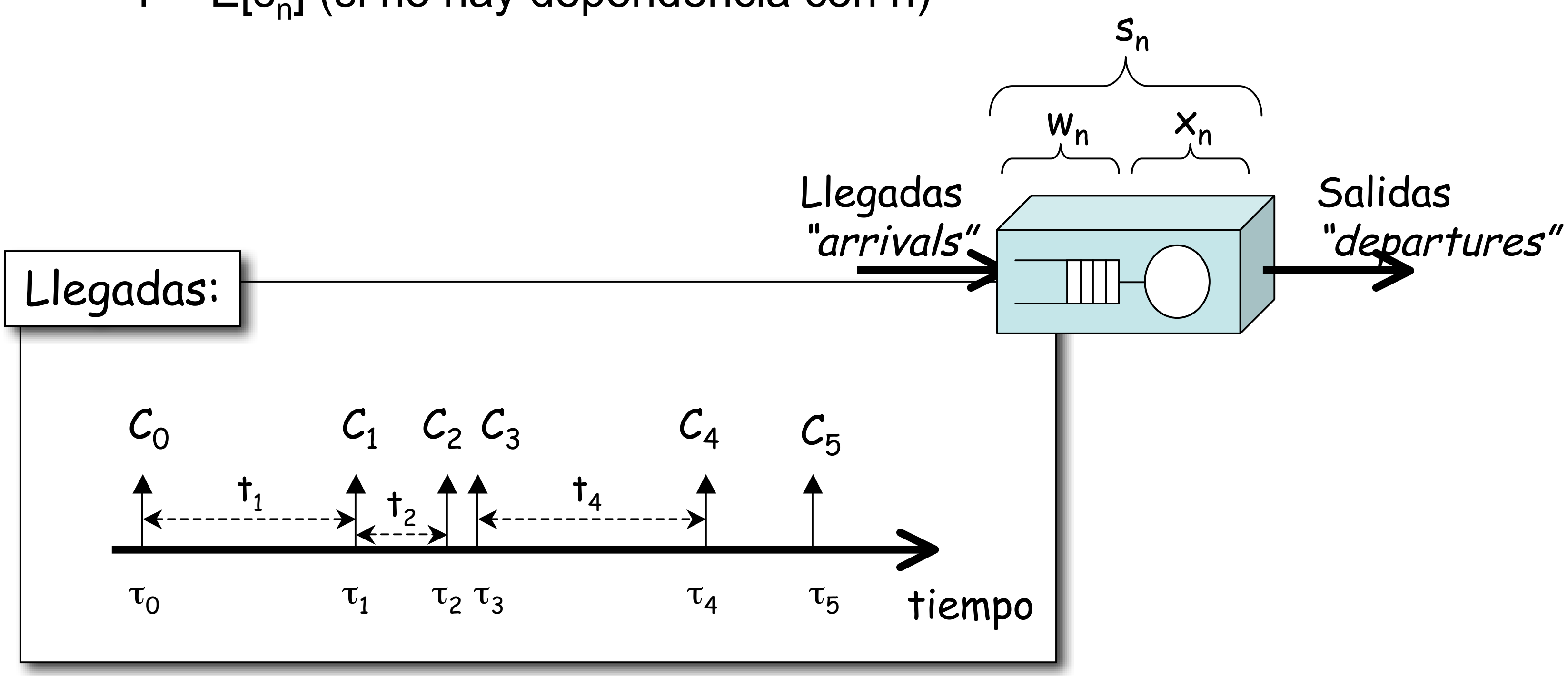


Tiempo en cola y en el sistema



Otros términos y definiciones

- **N(t)** : nº de clientes en el sistema en el instante t
  $\overline{N} = E[N]$
- w_n : tiempo de espera en la cola para C_n
  $W = E[w_n]$ (si no hay dependencia con n)
- s_n : tiempo en el sistema para C_n (tiempo total empleado)
  $s_n = w_n + x_n$
  $T = E[s_n]$ (si no hay dependencia con n)



Otros términos y definiciones

- Intensidad de tráfico

$$u = A_o = \frac{E[\underline{x}]}{E[\underline{t}]} = \frac{E[t\_servicio]}{E[t\_entre\_llegadas]} = \frac{1/\mu}{1/\lambda} = \frac{\lambda}{\mu}$$

- Tasa efectiva de llegadas o rendimiento del sistema

$$\lambda_a = \gamma = \lambda(1 - P_B)$$

$$P_B = P[\text{un usuario no sea servido}]$$

- Tráfico cursado: fracción del tráfico ofrecido servido por el sistema

$$A_c = A_o(1 - P_B) = \frac{\lambda(1 - P_B)}{\mu} = \frac{\lambda_a}{\mu}$$

- Con c servidores idénticos y reparto uniforme del tráfico: intensidad que debe atender cada uno

$$\rho = \frac{A_c}{c} = \frac{\lambda_a}{c\mu}$$

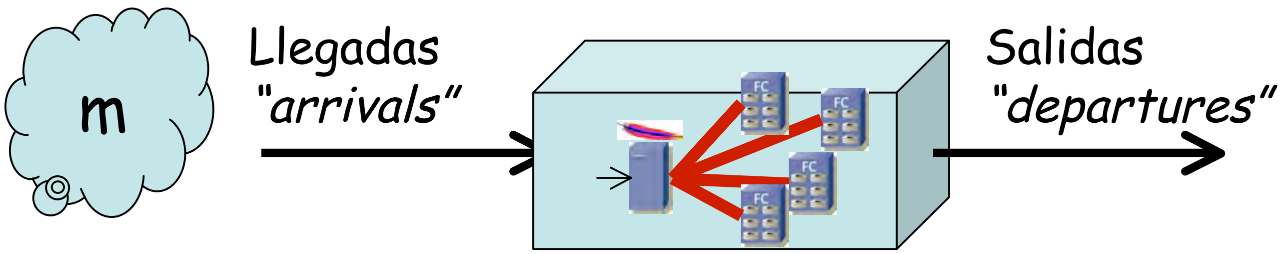
ρ = factor de utilización = % de tiempo que el servidor está ocupado = E[fracción de servidores ocupados]

$$\rho = \frac{\text{velocidad a la que entra trabajo}}{\text{máx. velocidad a la que se puede hacer}} = \frac{\lambda E[\underline{x}] \text{ segs de trabajo/seg}}{c \text{ segs/seg}}$$

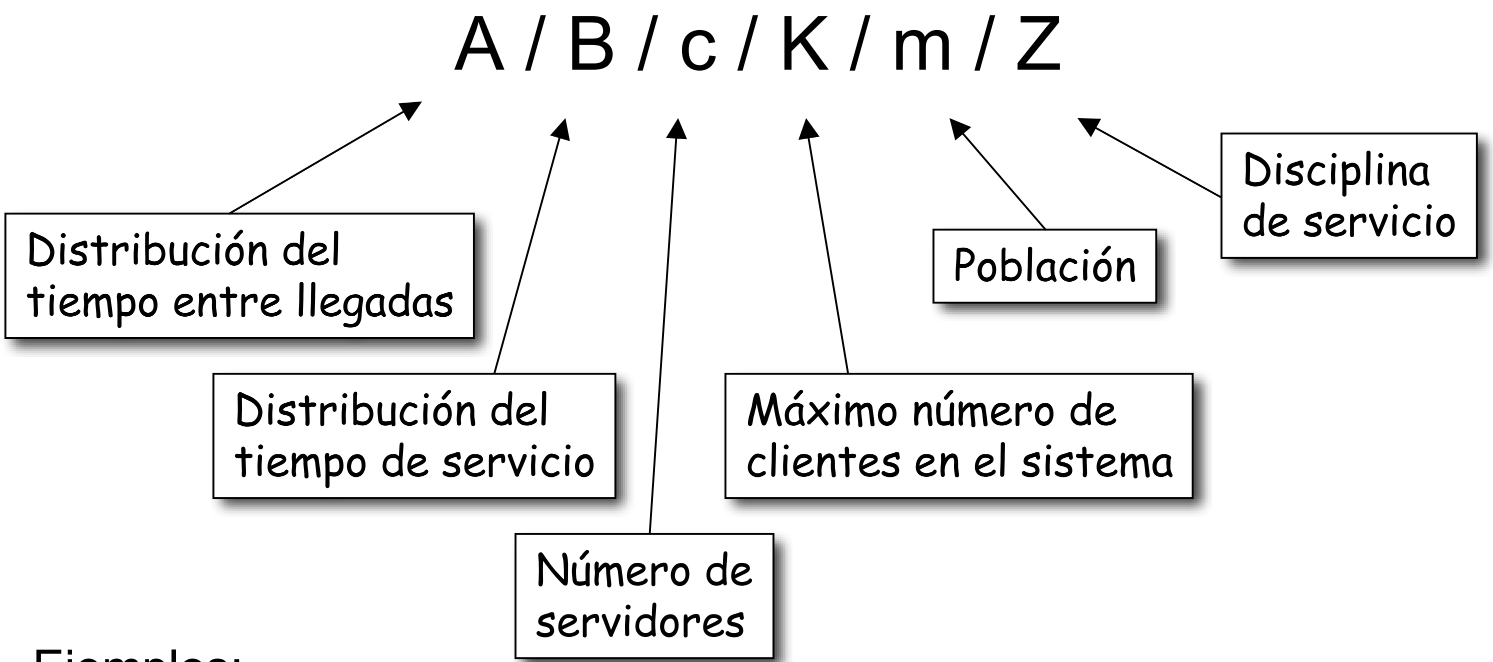
- Fórmula de Little: $\overline{N} = \lambda T$

Otros parámetros

- El número de servidores puede ser distinto de 1
  - ¿Todos sirven de igual manera? (misma distribución del tiempo de servicio)
  - ¿Cómo se selecciona servidor?
- Número máximo de clientes que puede haber en el sistema
- Número máximo de clientes *"en el universo"* (m)
- ¿FIFO? ¿LIFO? ¿Prioridades? ¿Según tamaño? ¿Al azar? ¿Preemptivo?



Notación de Kendall



- Ejemplos:
  - *"M"* : distribución exponencial : $P(X < x) = 1 - e^{-\beta x}$
  - *"D"* : determinista
  - *"G"* : distribución genérica

  - *"M/M/1"* : tiempos entre llegadas exponenciales, tiempos de servicio exponenciales, 1 solo servidor
  - *"D/D/2/5"* : tiempos entre llegadas deterministas, tiempos de servicio deterministas, 2 servidores, máximo de 3 clientes en cola

¿Podemos resolverlos?

- Podemos querer calcular:
  - Tiempo entre una llegada y que se completa su servicio:
    - Tiempo medio, distribución…
    - Nos sirve para saber si el sistema da un servicio de “suficiente” calidad
  - Número de peticiones encoladas ante una nueva llegada
    - Número medio, distribución…
    - Nos ayuda a saber cuánta cola deberíamos poner
  - Cómo son las salidas
    - ¿Siguen siendo deterministas/exponenciales/lo_que_sea?
    - Necesario conocerlo si detrás hay otro sistema que tome nuestros servicios como nuevas llegadas

- Podemos calcularlo:
  - Para algunos sistemas sencillos (D/D/c, M/M/c/K, etc)
  - Aproximaciones, cotas, para sistemas un poco más complejos
  - Simulación para sistemas “reales”

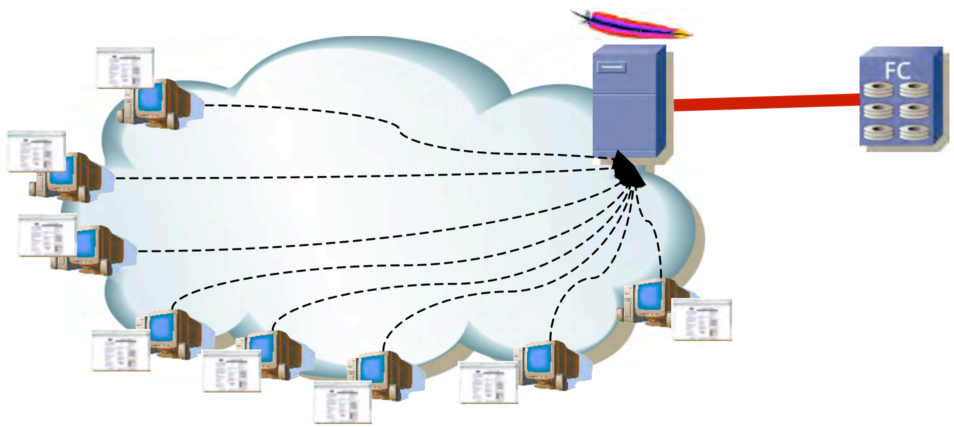
Ejemplo de aplicación

- Llegadas las tomamos equiespaciadas (determinista)
- Los tiempos de servicio siempre iguales (determinista)
- Los discos solo pueden servir 1 fichero a la vez

D/D/1

- ¿Cola? ¿Cuántas peticiones puede almacenar sin servir el servidor?
- Por ejemplo supongamos que es un servidor web multihilo que se bloquea el hilo al recibir una petición
- Tantos servidores como hilos (h)

D/D/1/h

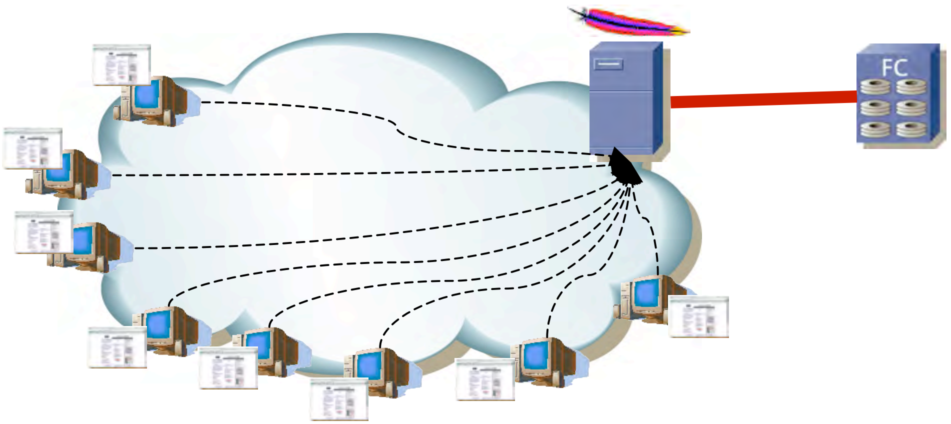


Ejemplo de aplicación

D/D/1/h

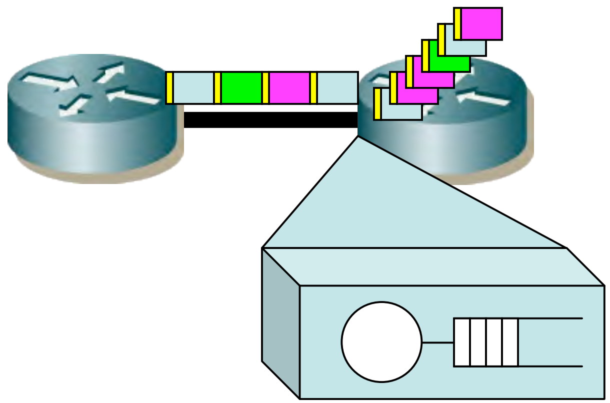
¿Son hipótesis realistas?

- No hace falta cola: si no es capaz de servir antes de la llegada siguiente ($\lambda>\mu$) el sistema será inestable
- Las llegadas serán normalmente aleatorias y dependerán principalmente del comportamiento de los usuarios
- Los tiempos de servicio dependen del tamaño de los ficheros y no son todos iguales
- También los tiempos dependerán de la red (ej. TCP)



Ejemplo de red

- Las llegadas son de los paquetes
- El tiempo de servicio es el tiempo de enviar el paquete
- El tiempo depende del tamaño del paquete
- La cola es el tamaño del buffer de memoria de paquetes

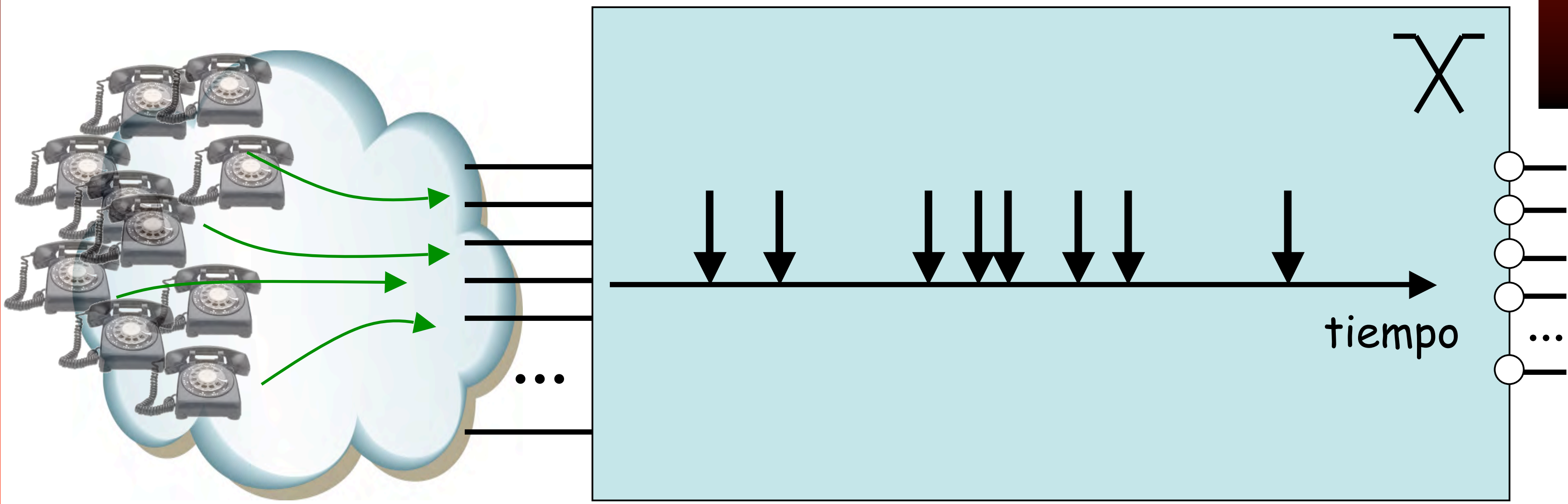


- ¿Cómo son los tiempos de llegada de paquetes a un router?
- ¿Cómo son los tamaños de los paquetes?
- ¿Es igual para cualquier router? (al menos el tipo de distribución)

¿Con qué solemos trabajar?

Proceso de llegadas

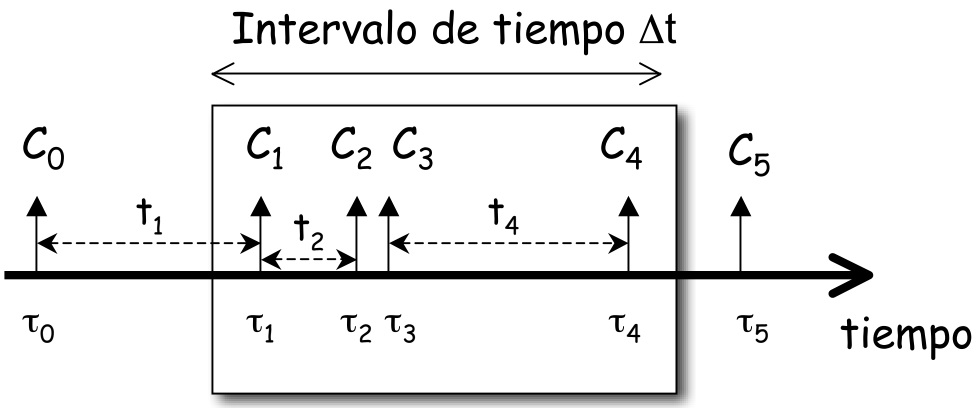
- Hipótesis fundamental en teoría clásica: llegadas independientes
- Tasa media de llegadas de llamadas de una gran población de fuentes (usuarios) independientes: λ



Número de llegadas

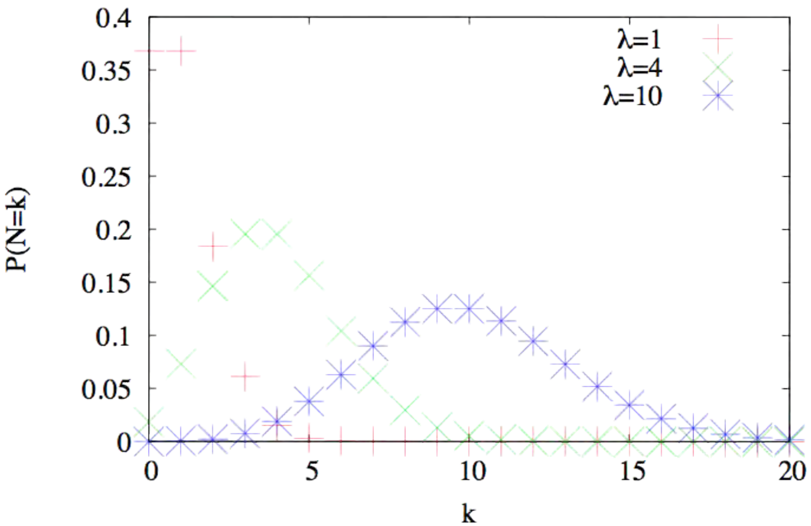
- Hipótesis:
  - En un intervalo suficientemente pequeño solo puede producirse una llegada
  - La probabilidad de una llegada en un intervalo suficientemente pequeño es directamente proporcional a la longitud del mismo (probabilidad $\lambda\Delta t$)
  - La probabilidad de una llegada en un intervalo es independiente de lo que suceda en otros intervalos
- Se demuestra que el número de llegadas en un intervalo sigue una distribución de Poisson

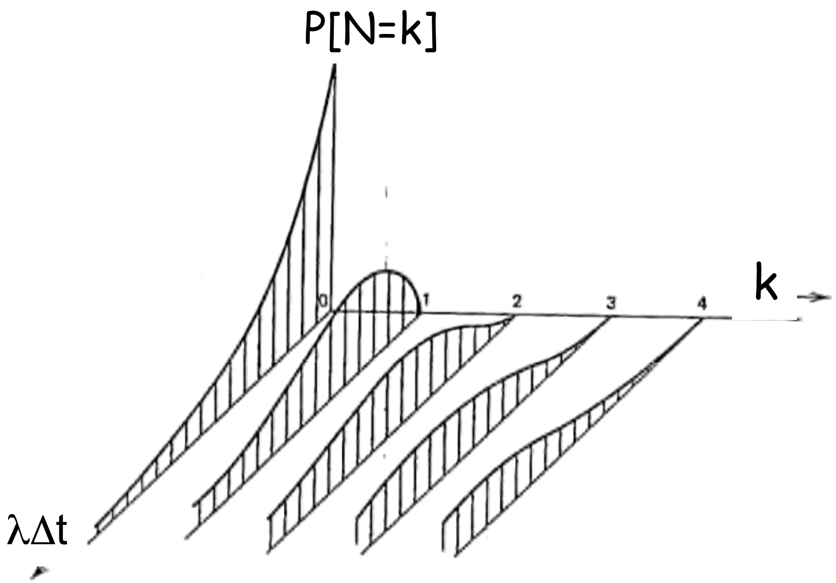
$$P_{\lambda\Delta t}[N = k] = \frac{(\lambda\Delta t)^k}{k!} e^{-\lambda\Delta t}$$



Distribución de Poisson

$$P[N=k]=\frac{(\lambda\Delta t)^k}{k!}e^{-\lambda\Delta t}$$





- Es una función de distribución:

$$\sum_{k=0}^{\infty}P[N=k]=\left(1+\lambda\Delta t+\frac{(\lambda\Delta t)^2}{2}+\frac{(\lambda\Delta t)^3}{6}+...\right)e^{-\lambda\Delta t}=e^{\lambda\Delta t}e^{-\lambda\Delta t}=1$$

- Su valor medio es λΔt :

$$\overline{N}=E[N]=\sum_{k=0}^{\infty}kP[N=k]=\left(0+\lambda\Delta t+(\lambda\Delta t)^2+\frac{(\lambda\Delta t)^3}{2}+\frac{(\lambda\Delta t)^4}{6}...\right)e^{-\lambda\Delta t}=\lambda\Delta t e^{\lambda\Delta t}e^{-\lambda\Delta t}=\lambda\Delta t$$

Tiempos entre llegadas

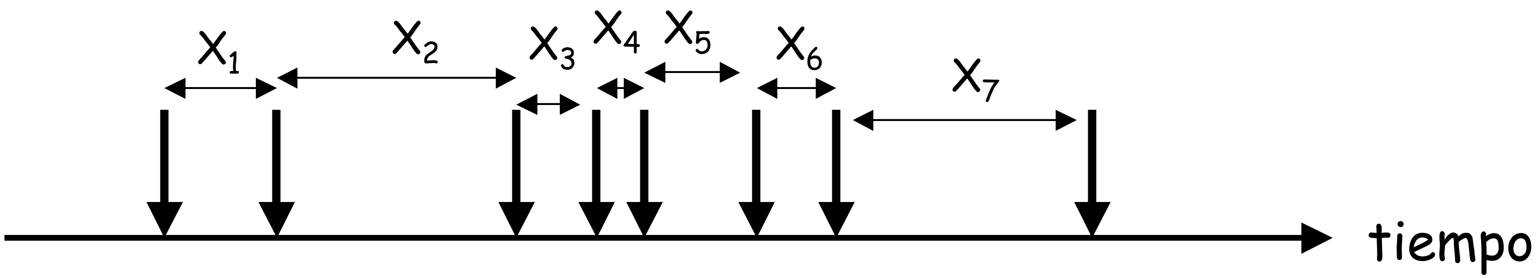
- Se demuestra que: si el número de eventos que ocurren en un intervalo sigue una distribución de Poisson los tiempos entre llegadas de eventos siguen una distribución exponencial
- El tiempo entre llegadas sigue una v.a. exponencial de parámetro λ
- X_i variables aleatorias independientes e idénticamente distribuidas (i.i.d.) ('X')

$$\boxed{p_X(t) = \lambda e^{-\lambda t} \quad \text{(t>0)}}$$

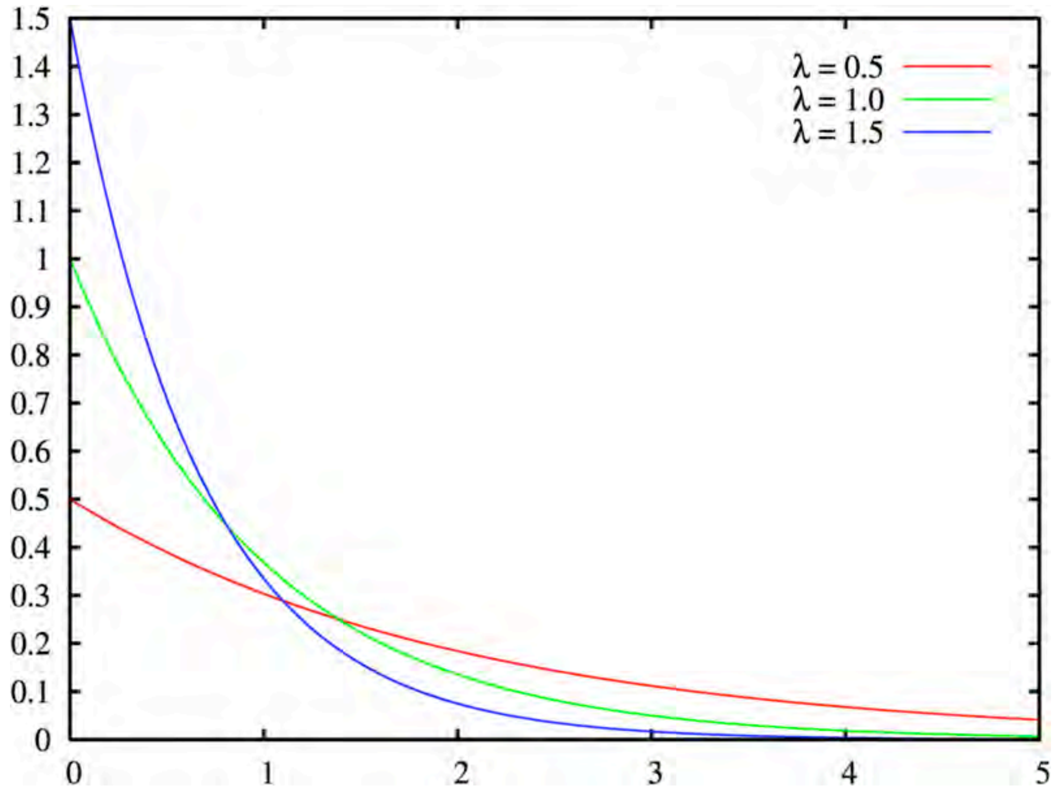
$$P[X < t] = 1 - e^{-\lambda t}$$

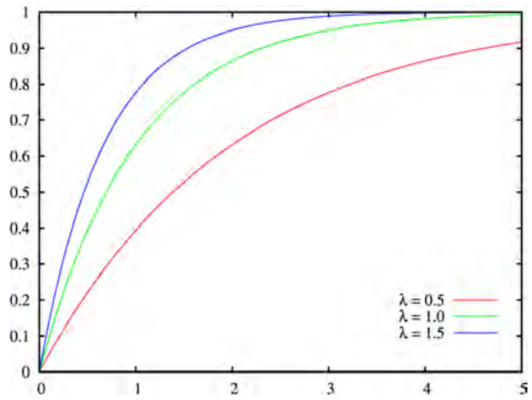
- Media: $E[X] = \int_0^{\infty} t\lambda e^{-\lambda t} = 1/\lambda$

- Tiempo medio entre llegadas $1/\lambda \Rightarrow$ en media λ llegadas por segundo

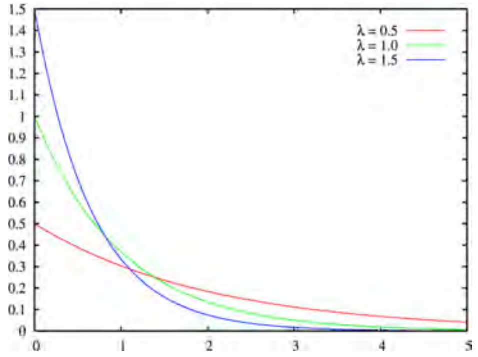


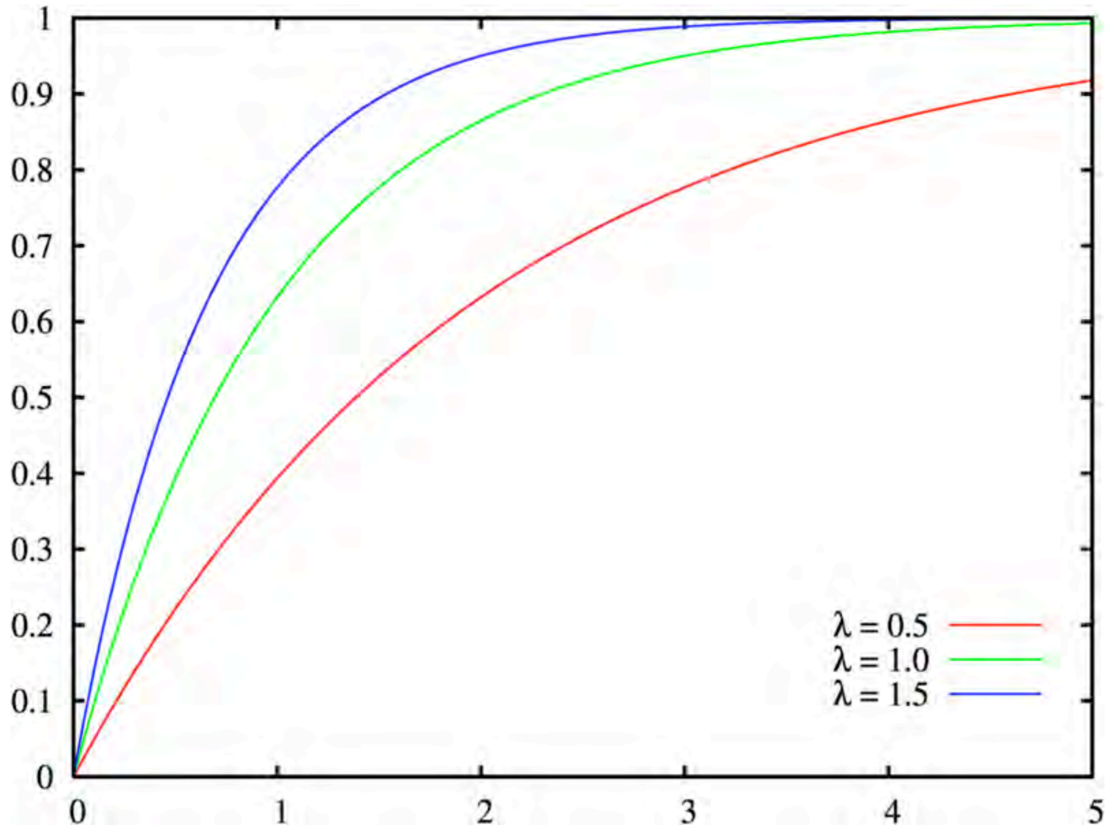
Ejemplo


λ = 0.5
λ = 1.0
λ = 1.5
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5

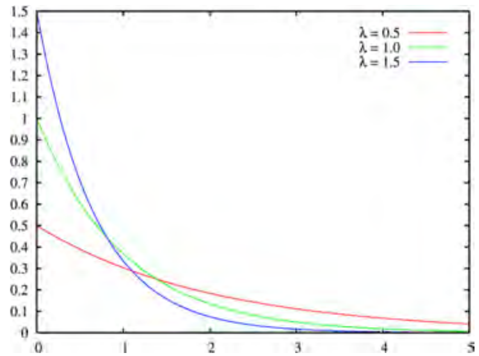

λ = 0.5
λ = 1.0
λ = 1.5

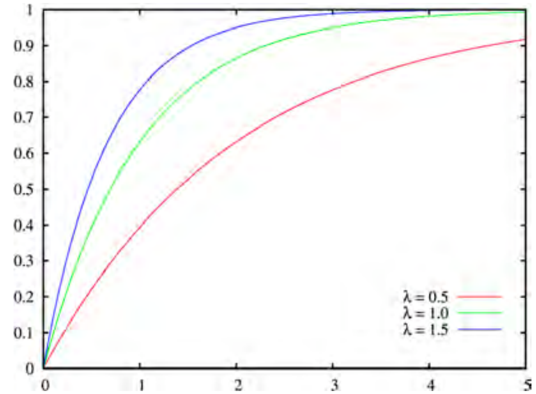
Ejemplo

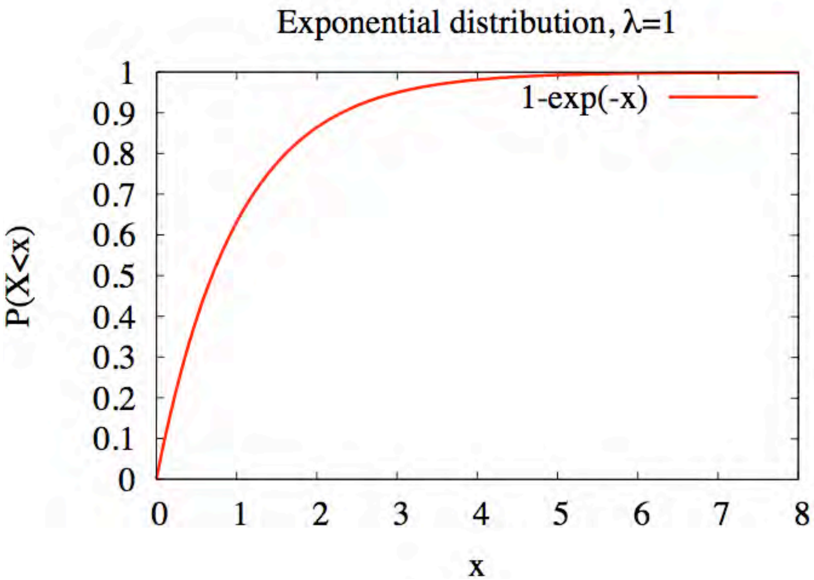


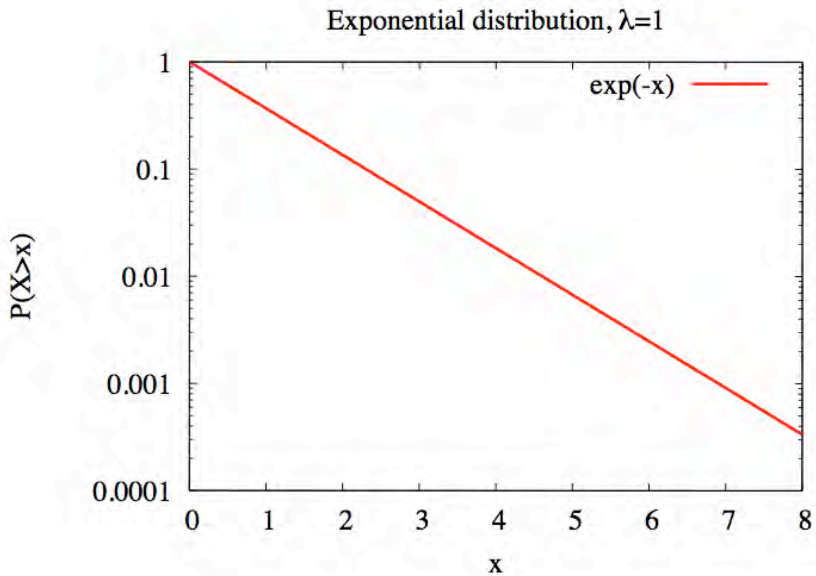


Ejemplo


λ = 0.5
λ = 1.0
λ = 1.5


λ = 0.5
λ = 1.0
λ = 1.5


Exponential distribution, λ=1
1-exp(-x)
P(X<x)
x


Exponential distribution, λ=1
exp(-x)
P(X>x)
x

¿Llegadas de *Poisson?*

- ¿Los tiempos entre llegadas de paquetes son i.i.d. y siguen una distribución exponencial (*M*)?
- ¿Equiespaciadas (determinista)?
  - Fuentes CBR
- ¿Los tiempos entre llegadas son independientes?
  - ¿Las peticiones de páginas web? El usuario suele navegar en “sesiones”
  - ¿Los paquetes IP? Si vienen varios paquetes suele haber bastante probabilidad de que vengan más $\Rightarrow$ correlación no nula

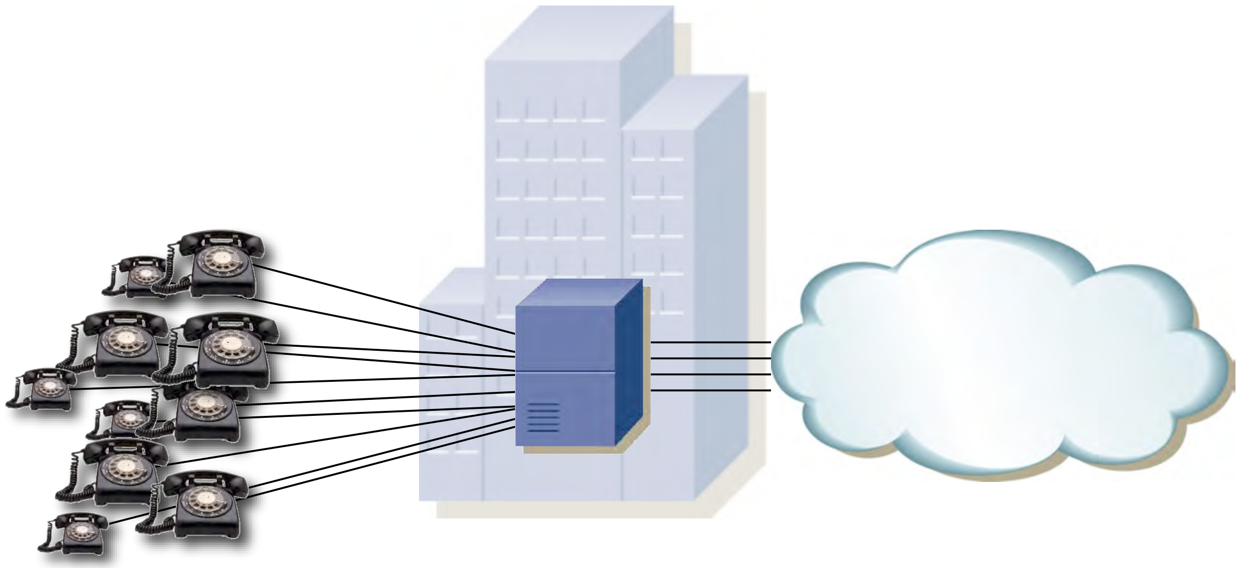
Tiempos de servicio

- Equivalentemente tamaños de paquetes y de ficheros
- ¿Deterministas?
  - Celdas ATM ok
  - ¿Ficheros servidor web?
  - ¿Paquetes IP?
- ¿Exponenciales?
  - Es una distribución continua. ¿Aproximado?
  - ¿Los ficheros que piden los usuario tienen esa distribución?
  - ¿Los tamaños de los paquetes IP?
- ¿Independientes?

M/M/c/c

Servicio telefónico (Erlang)

- Gran número de usuarios
- Tiempo entre llamadas a una central: exponenciales independientes
- Duración de las llamadas: exponencial
- c líneas externas
- Nos permite calcular la probabilidad de que una llamada encuentre una línea libre
- Estupendo, pero un buen día…

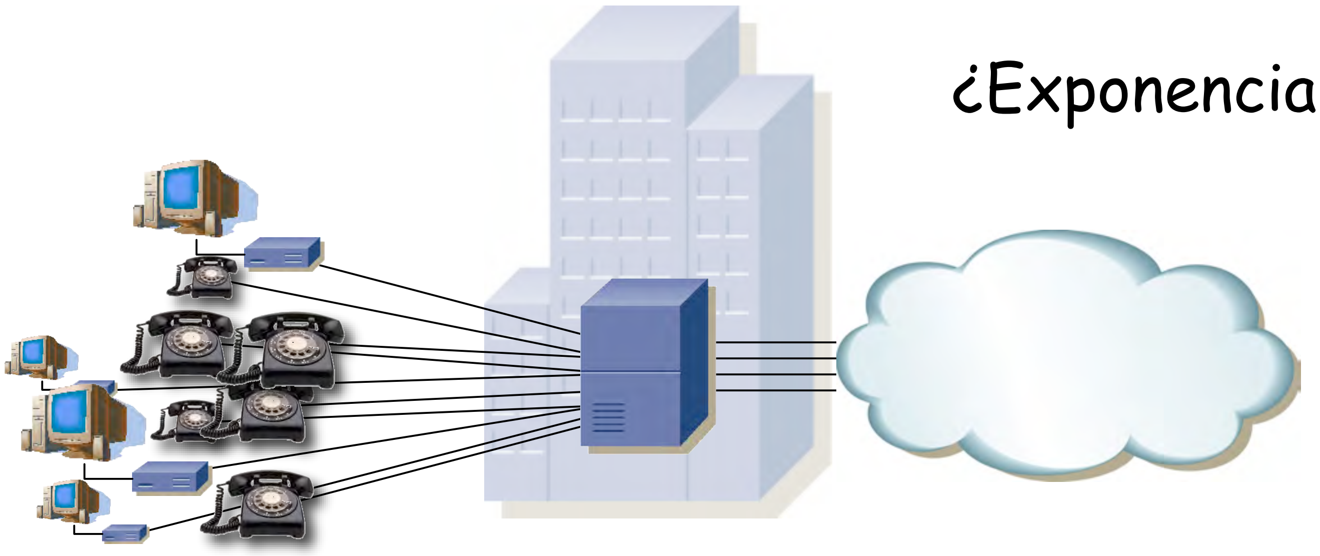


M/M/c/c

Servicio telefónico (Erlang)

- Gran número de usuarios
- Tiempo entre llamadas a una central: exponenciales independientes
- Duración de las llamadas: exponencial
- c líneas externas
- Nos permite calcular la probabilidad de que una llamada encuentre una línea libre
- Estupendo, pero un buen día…

¿Exponenciales?



Traffic Measurement

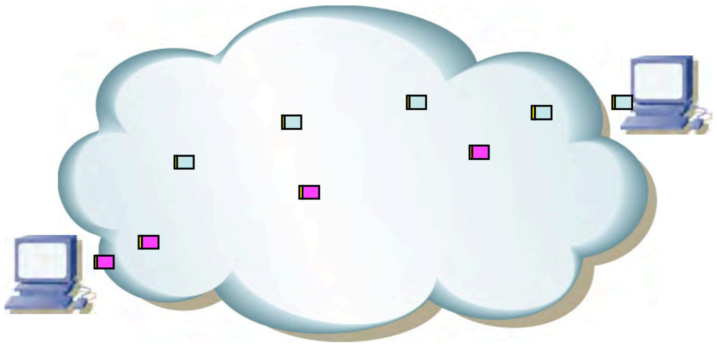
Utilidad

- *Network troubleshooting*
  - Equipo defectuoso puede afectar al funcionamiento de la red
  - Las medidas ayudan a localizarlo
- Depuración de protocolos/aplicaciones
  - Las medidas pueden desvelar problemas ocultos
- Caracterización de tráfico
  - ¿Cuál es la carga actual y la tendencia?
  - Medida necesaria
- Evaluación de prestaciones
  - Prestaciones de un router o aplicación
  - Medida necesaria para caracterizar la carga

Medidas activas y pasivas

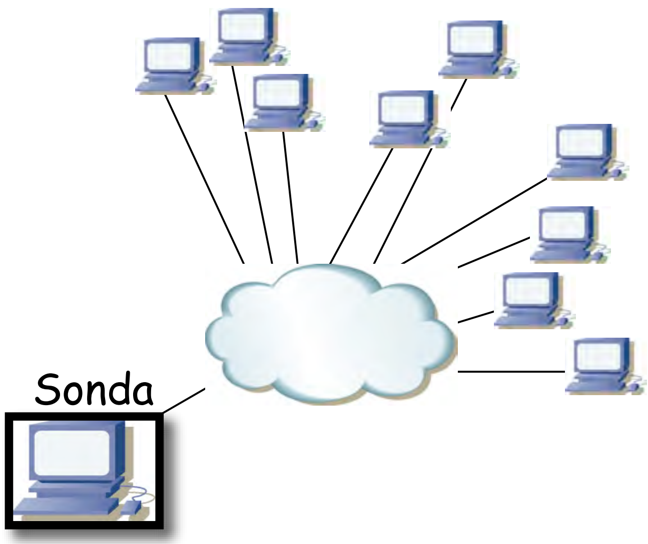
Activas

- Los paquetes son generados por una herramienta
- Se busca medir con ellos características de la red
- Bien enviándolos a otro nodo controlado
- O bien esperando una respuesta
- O bien con un medidor pasivo intermedio
- Software+PC o hardware específico
- Ejemplos: ping, traceroute, pathchar, etc.



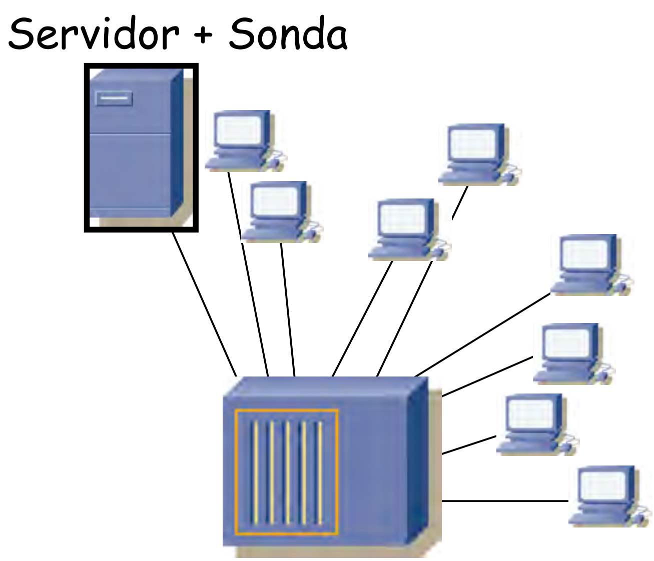
Pasivas

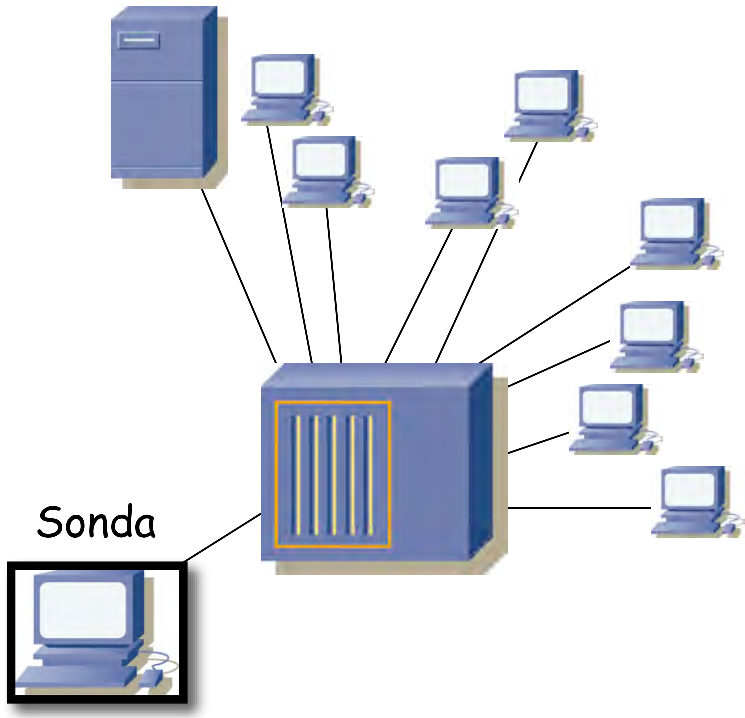
- Herramienta de monitorización observa el tráfico
- En un punto de la red o varios
- Software+PC o hardware específico
- Ejemplos: tcpdump, wireshark, etc.



¿Cómo se captura el tráfico?

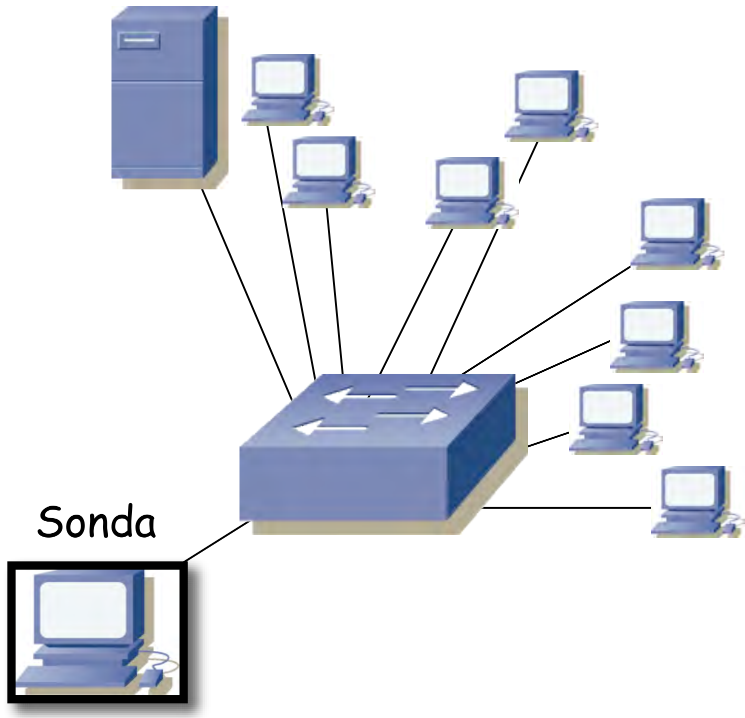
- Hub y *sniffer* (tcpdump, programa ad-hoc, etc)
- Trazas a disco
- Análisis en tiempo real u offline





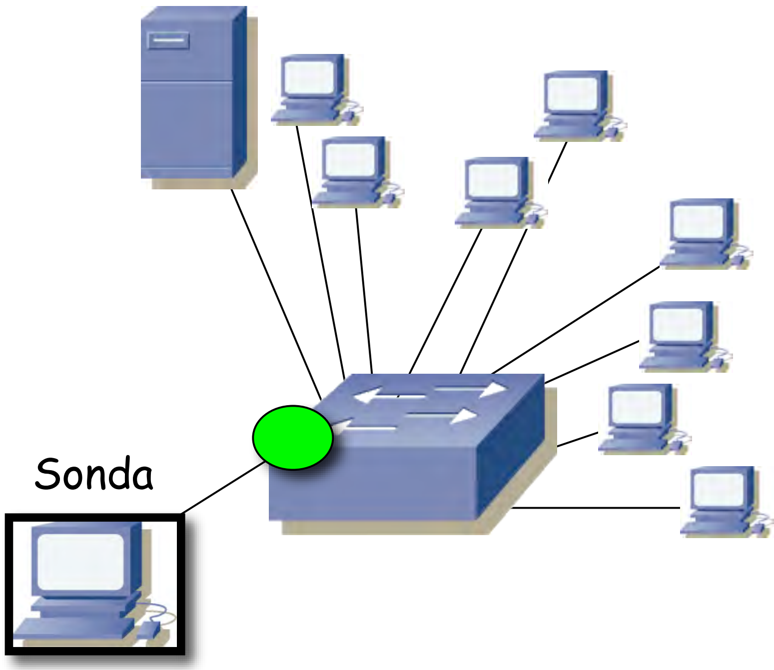
¿Cómo se captura el tráfico?

- Switch con *SPAN* y *sniffer*
- Trazas a disco
- Análisis en tiempo real u offline



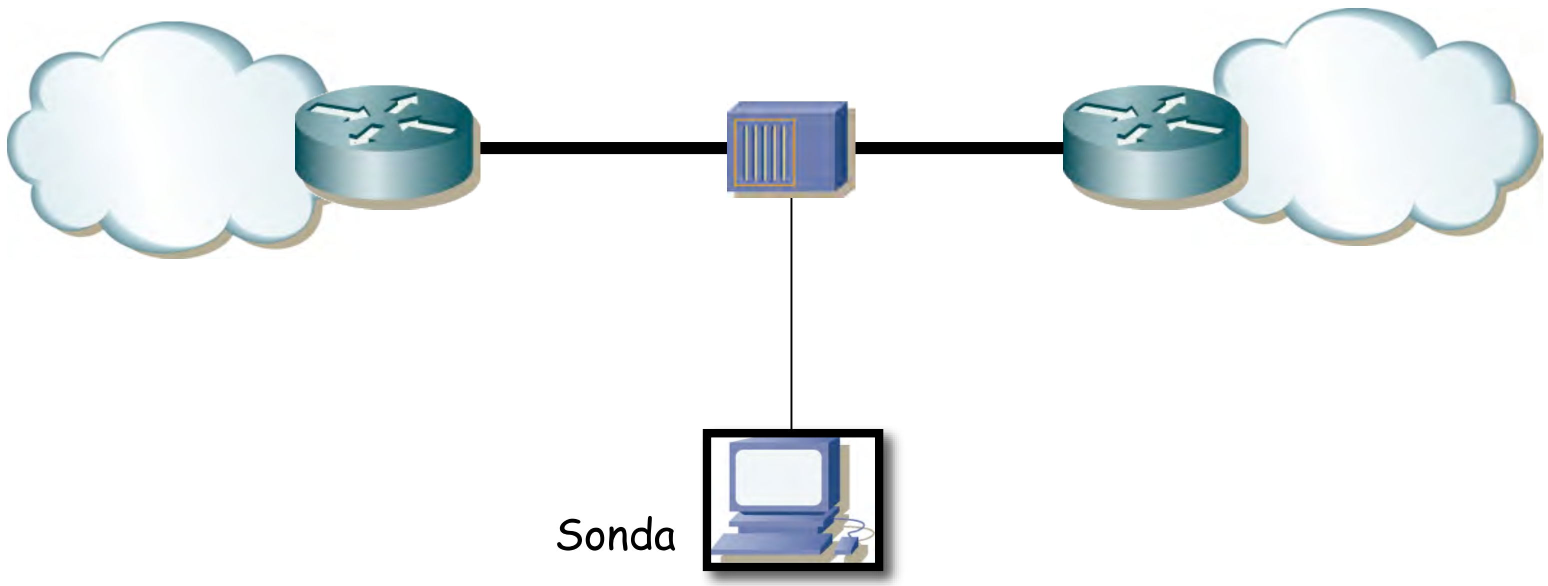
¿Cómo se captura el tráfico?

- Conmutador puede recoger estadísticas
- Normalmente resumidas
- La sonda recogería periódicamente esos resúmenes
- También aplica a routers
- Ejemplo: NetFlow



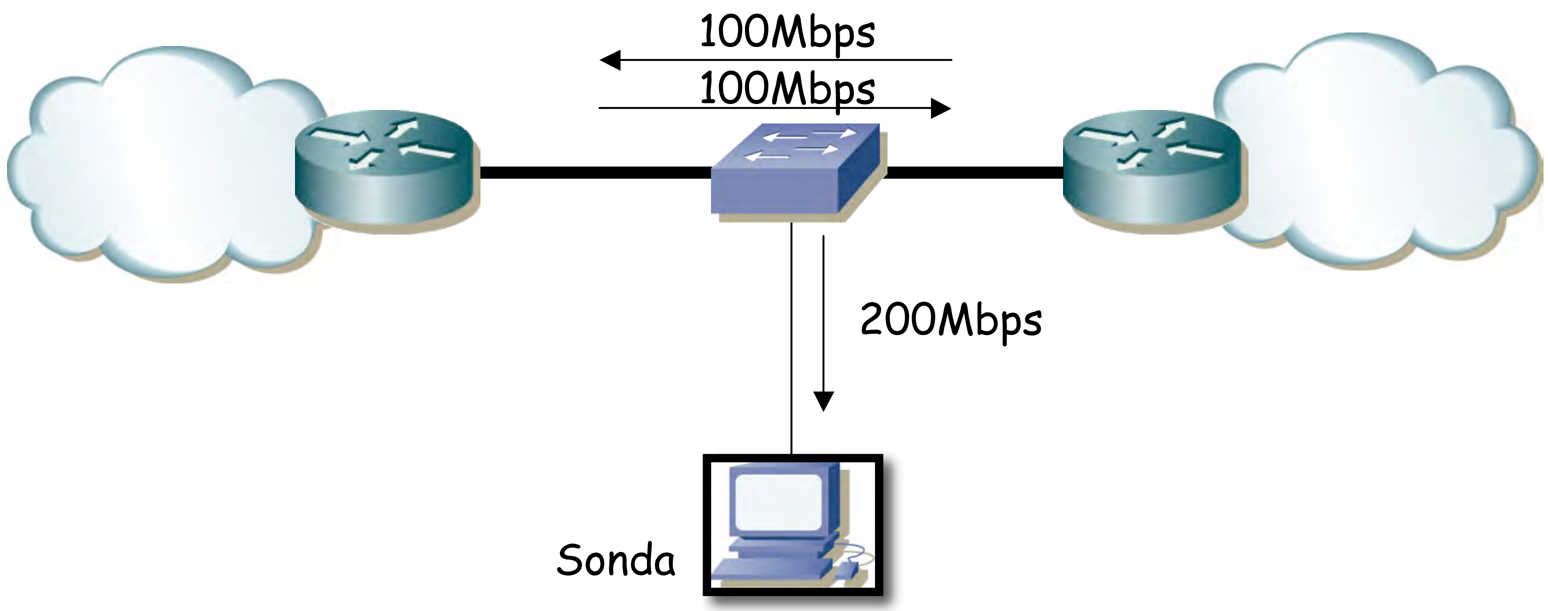
¿Cómo se captura el tráfico?

- Enlace entre 2 routers (…)
- Poner un hub
  - Solo para 10/100 Ethernet
  - Podría para Gigabit pero no se fabrican hubs
  - Se pierde el full duplex, puede haber colisiones



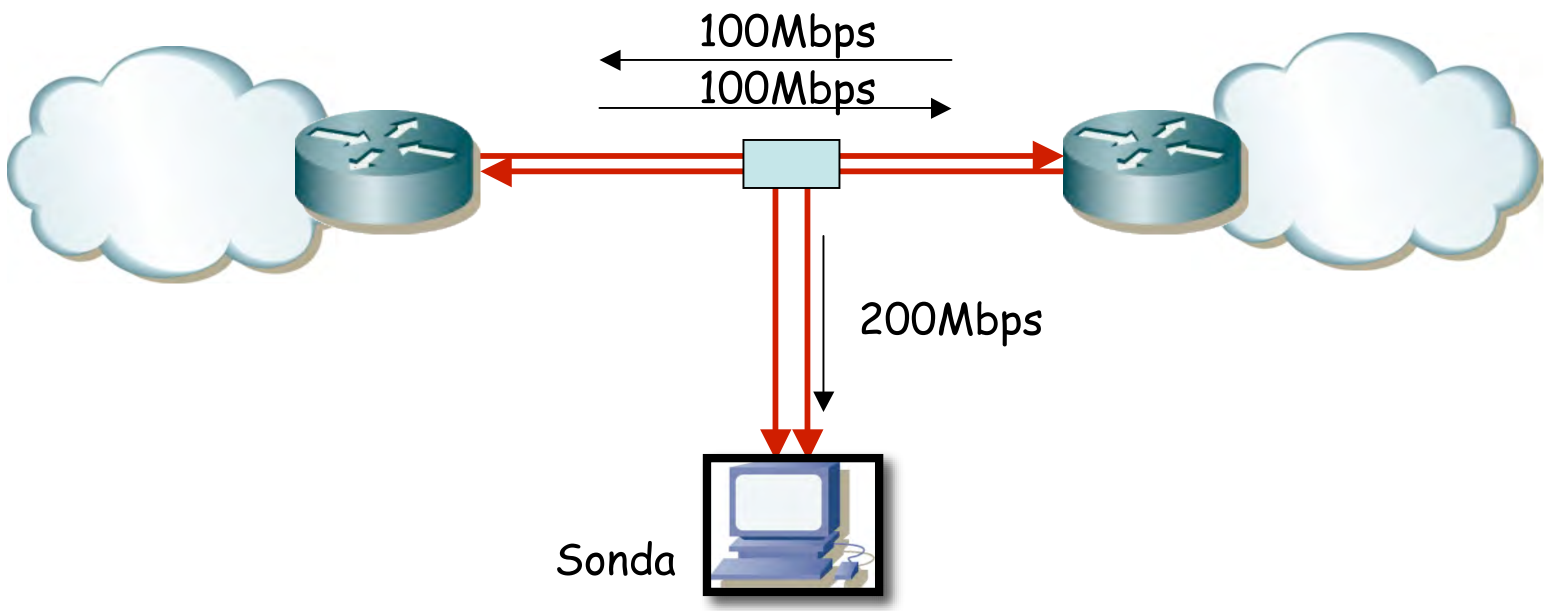
¿Cómo se captura el tráfico?

- Enlace entre 2 routers
- Poner un switch con *SPAN*
  - Con full duplex el tráfico a monitorizar es 2x la velocidad



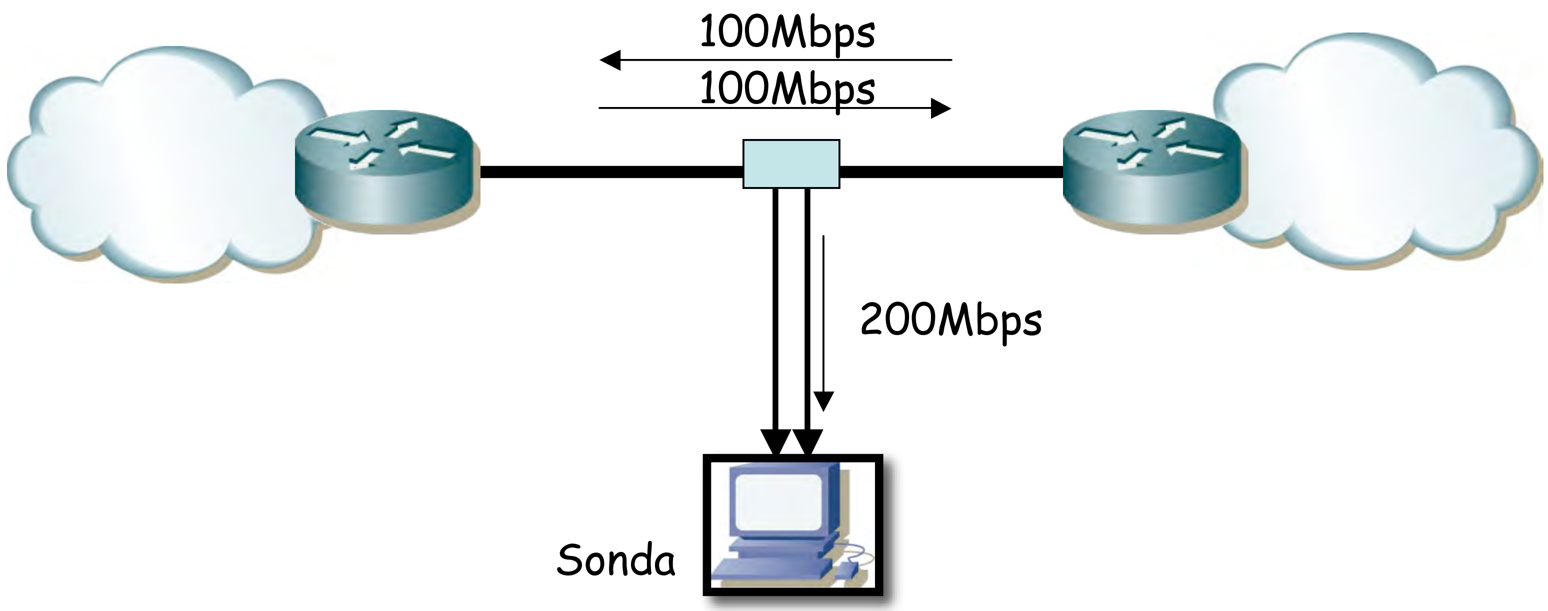
¿Cómo se captura el tráfico?

- Enlace **óptico** entre 2 routers
- Poner un splitter
  - Con full duplex el tráfico a monitorizar es 2x la velocidad
  - Normalmente requiere 2 interfaces de red en la sonda (solo RX)
  - Pasivo (resistente a fallos de alimentación)

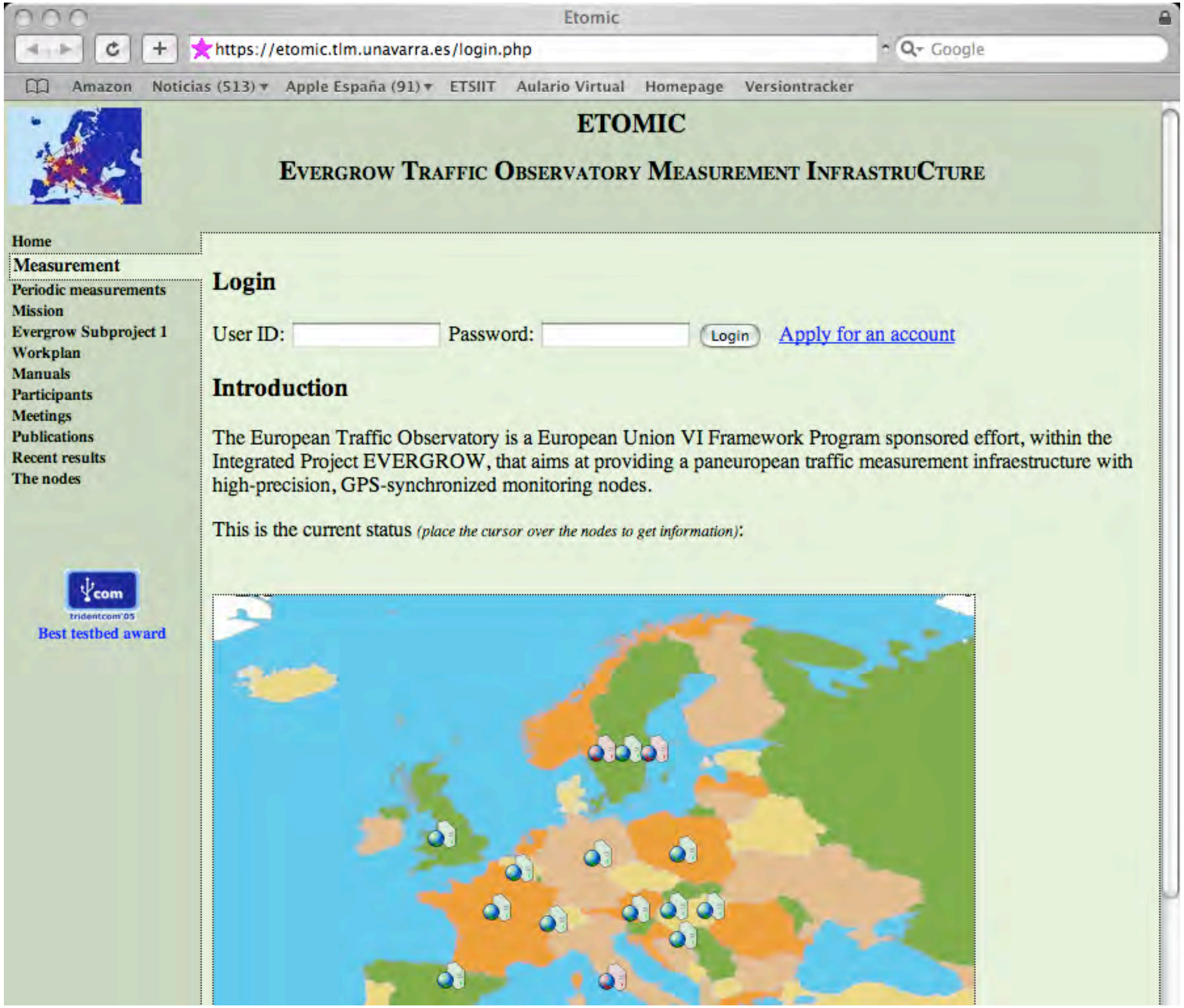


¿Cómo se captura el tráfico?

- También existe el splitter (*Tap*) eléctrico
  - Con full duplex el tráfico a monitorizar es 2x la velocidad
  - Normalmente requiere 2 interfaces de red en la sonda (solo RX)
  - Pasivo (resistente a fallos de alimentación)



Medidas distribuidas



Modelado del tráfico

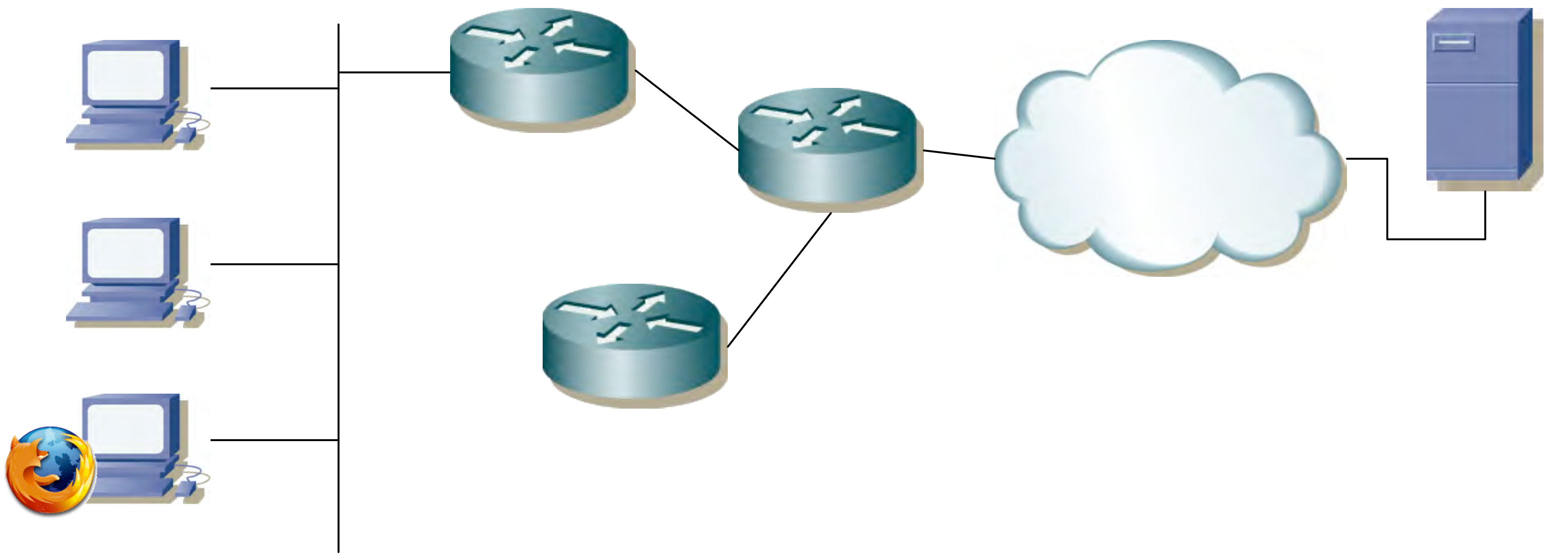
- Determinar la escala que interesa
- Medir en el punto de interés
- Decidir las estadísticas a calcular
- Estimarlas con las observaciones
- Elegir un modelo
- Ajustar los parámetros del mismo
- Comprobar la precisión del mismo

Escala de interés

- Nivel de sesión
  - Ej: peticiones de ficheros a un servidor ftp o web
  - Relacionado con comportamiento humano
- Nivel de flujo
  - El: bytes entre dos nodos de la red
- Nivel de paquete

Punto de medida

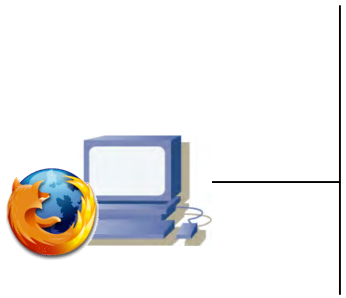
- ¿Dónde queremos modelar el tráfico?

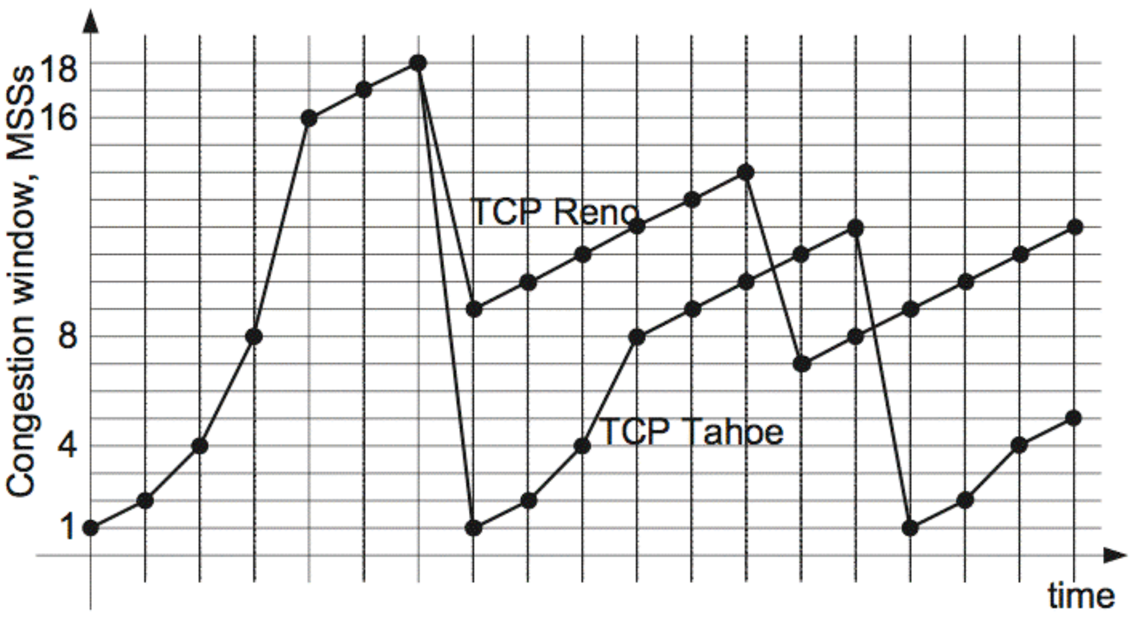


Punto de medida

En la aplicación/host

- Generalmente tráfico generado
- Según la aplicación UDP o TCP
- UDP cambia poco el tráfico frente a cuándo pide enviar la aplicación
- En TCP importantes efectos de la red y protocolos

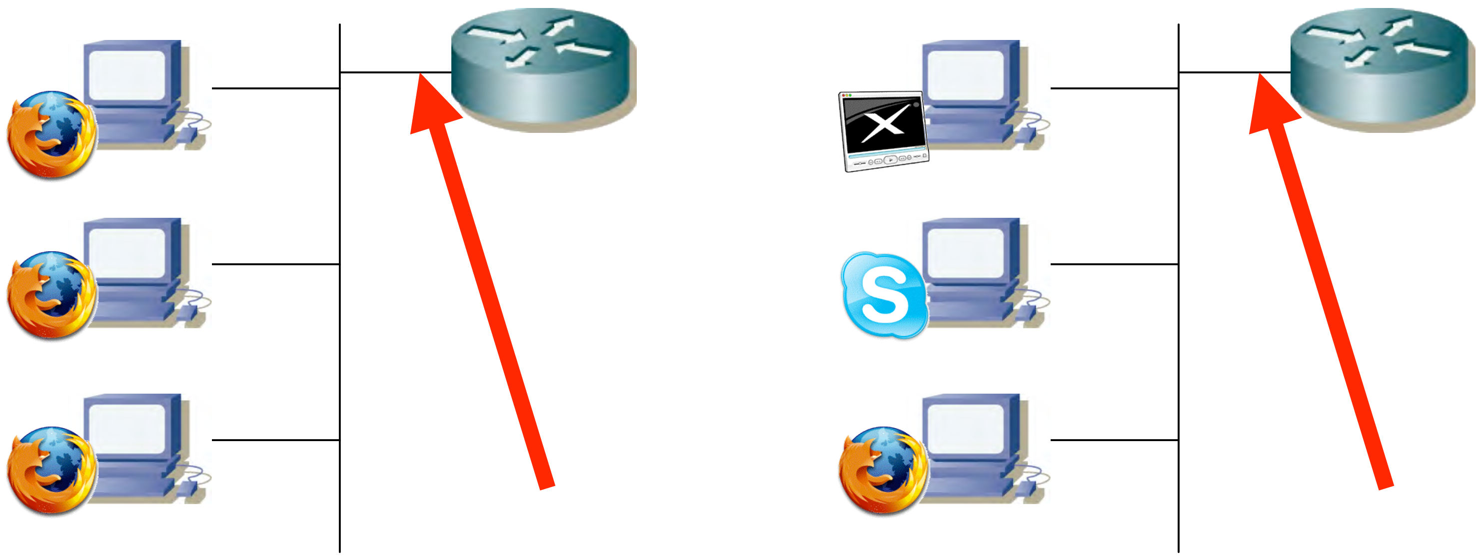




Punto de medida

Tráfico agregado de aplicaciones

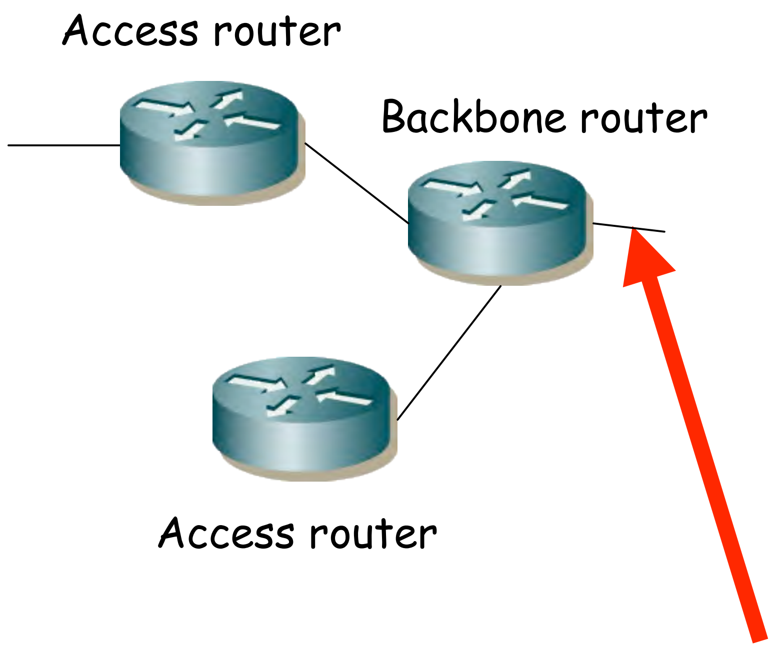
- Escenario homogéneo en aplicaciones
- O heterogéneo



Punto de medida

Tráfico agregado de red

- Agregación de gran número de flujos
- No es simplemente la superposición de los flujos individuales



Internet Traffic

Algunas características

- La mayor parte del tráfico es TCP
  - Se observó ya en los 90
  - Se esperaba que creciera el uso de UDP
  - Hoy en día muchos contenidos multimedia sobre TCP

Algunas características

- La mayor parte del tráfico es TCP
- Los flujos son bidireccionales pero asimétricos
  - ftp, http, ssh, email, etc. flujos bidireccionales
  - Request-response genera asimetría
  - P2P pero accesos asimétricos

Algunas características

- La mayor parte del tráfico es TCP
- Los flujos son bidireccionales pero asimétricos
- La mayoría de las conexiones TCP son cortas
  - WWW contenidos pequeños
  - HTTP 1.0 una conexión por objeto
  - HTTP 1.1 una conexión por página
  - Distribución de cola pesada

Algunas características

- La mayor parte del tráfico es TCP
- Los flujos son bidireccionales pero asimétricos
- La mayoría de las conexiones TCP son cortas
- Las llegadas de paquetes no forman un proceso de Poisson
  - Muchos estudios asumen tráfico agregado Poisson homogéneo
  - No parece ser así
  - Interarrival times correlados
  - Y puede que ni sea estacionario

Algunas características

- La mayor parte del tráfico es TCP
- Los flujos son bidireccionales pero asimétricos
- La mayoría de las conexiones TCP son cortas
- Las llegadas de paquetes no forman un proceso de Poisson
- Las conexiones se aproximan más
  - Muchos usuarios
  - Los usuarios suelen ser independientes
  - Similar al caso telefónico

Algunas características

- La mayor parte del tráfico es TCP
- Los flujos son bidireccionales pero asimétricos
- La mayoría de las conexiones TCP son cortas
- Las llegadas de paquetes no forman un proceso de Poisson
- Las conexiones se aproximan más
- Los tamaños de paquetes suelen seguir una distribución bimodal
  - En torno al 50% de la MTU de Ethernet
  - En torno al 40% lo más pequeños posible (ACKs)
  - El resto distribuido entre ellos
  - Algunos picos en tamaños de fragmentación (otras MTUs)

Algunas características

- La mayor parte del tráfico es TCP
- Los flujos son bidireccionales pero asimétricos
- La mayoría de las conexiones TCP son cortas
- Las llegadas de paquetes no forman un proceso de Poisson
- Las conexiones se aproximan más
- Los tamaños de paquetes suelen seguir una distribución bimodal
- Reparto no uniforme del tráfico
  - El 90% del tráfico entre el 10% de los nodos

Algunas características

- La mayor parte del tráfico es TCP
- Los flujos son bidireccionales pero asimétricos
- La mayoría de las conexiones TCP son cortas
- Las llegadas de paquetes no forman un proceso de Poisson
- Las conexiones se aproximan más
- Los tamaños de paquetes suelen seguir una distribución bimodal
- Reparto no uniforme del tráfico
- El tráfico agregado parece tener naturaleza multi-fractal
  - En los 80 se creía Poisson
  - En los 90 aparece la hipótesis de autosimilaridad
  - Hoy en día, en escalas medias-altas parece no estacionario

Algunas características

- La mayor parte del tráfico es TCP
- Los flujos son bidireccionales pero asimétricos
- La mayoría de las conexiones TCP son cortas
- Las llegadas de paquetes no forman un proceso de Poisson
- Las conexiones se aproximan más
- Los tamaños de paquetes suelen seguir una distribución bimodal
- Reparto no uniforme del tráfico
- El tráfico agregado parece tener naturaleza multi-fractal
- ¡ Sigue cambiando !
  - 80-95 : e-mail, acceso remoto
  - 95-00 : www, grandes transferencias
  - 00-futuro: p2p, www, …

Descripción del tráfico

Descripción de proceso estocástico

- Es una familia de variables aleatorias indexadas por un parámetro

$$\{S(t),\, t \in T\}$$

- Descripción completa requiere la *joint probability density function*

$$S(t_1,s_1,t_2,s_2,\ldots) = P\left(S(t_1) \le s_1, S(t_2) \le s_2, \ldots\right)$$

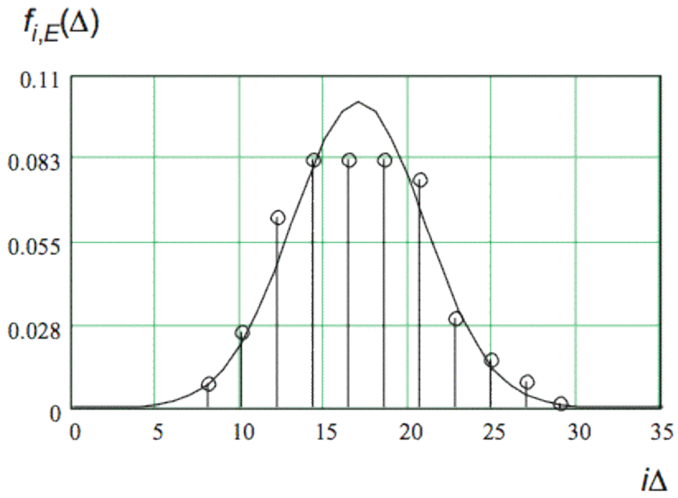
- Inmanejable, nunca se considera esa opción
- Si el proceso es estacionario en el sentido amplio
  - La media de S(t) no depende del tiempo
  - La autocorrelación depende solo de la diferencia de índice $R_S(\Delta t)$
- Normalmente nos interesará marginal y autocorrelación
- Nota: Autocorrelación = $R_S(t_1,t_2) = E\left[S(t_1)S(t_2)\right]$

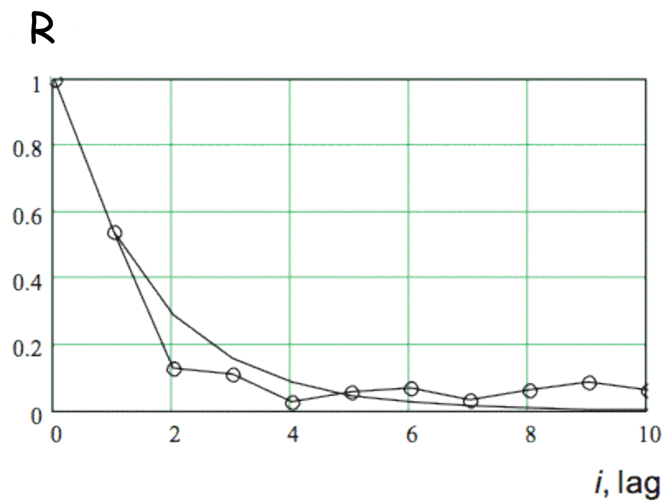
  Autocovarianza = $C_S(t_1,t_2) = E\left[\left(S(t_1) - \mu(t_1)\right)\left(S(t_2) - \mu(t_2)\right)\right]$

  $\mu(t) = E[S(t)]$

Descripción

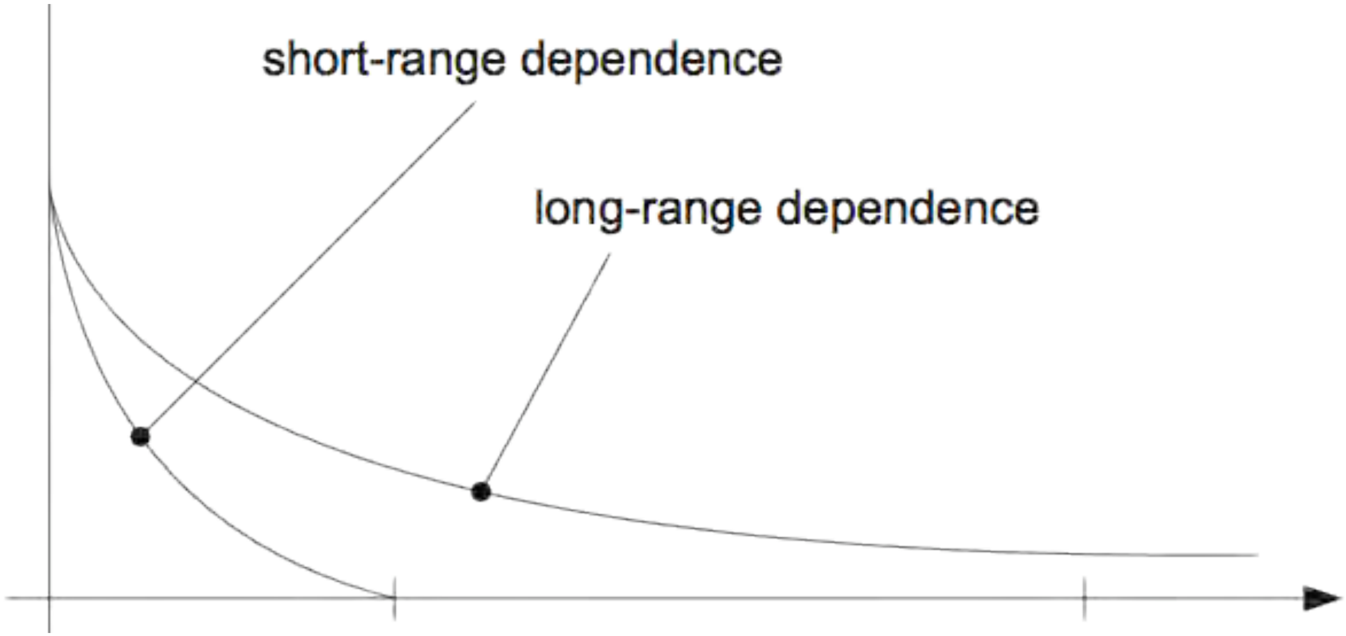
- Normalmente tenemos muchos datos
- Capturas de días o de enlaces de muy alta velocidad
- Millones de paquetes
- Se reduce si queremos un intervalo estacionario (1h?)





Efectos de la autocorrelación

- Importante en análisis de prestaciones
- Alta correlación en los tiempos entre llegadas de paquetes suele implicar mayores pérdidas
- Dos manifestaciones importantes
  - Dependencia a corto plazo (caída exponencial/geométrica)
  - Dependencia a largo plazo (*power decay*) (no confundirla con no-estacionariedad)



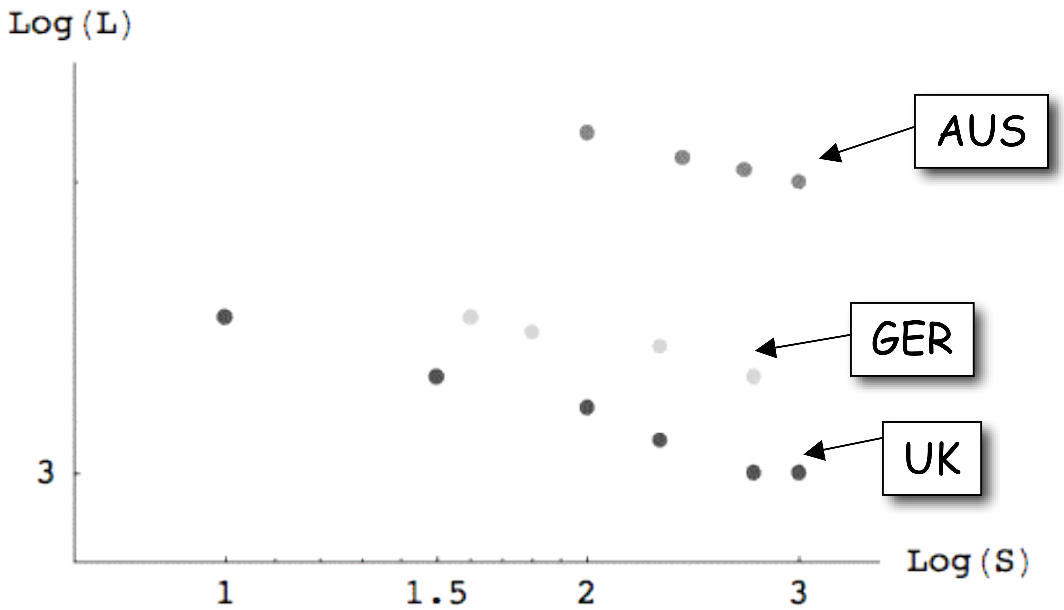
Escalas

- S = separación del compás
- L = longitud de la costa

$$\log(L) = -D\log(s)$$

$$L = S^{-D}$$

- *Power Law scaling*



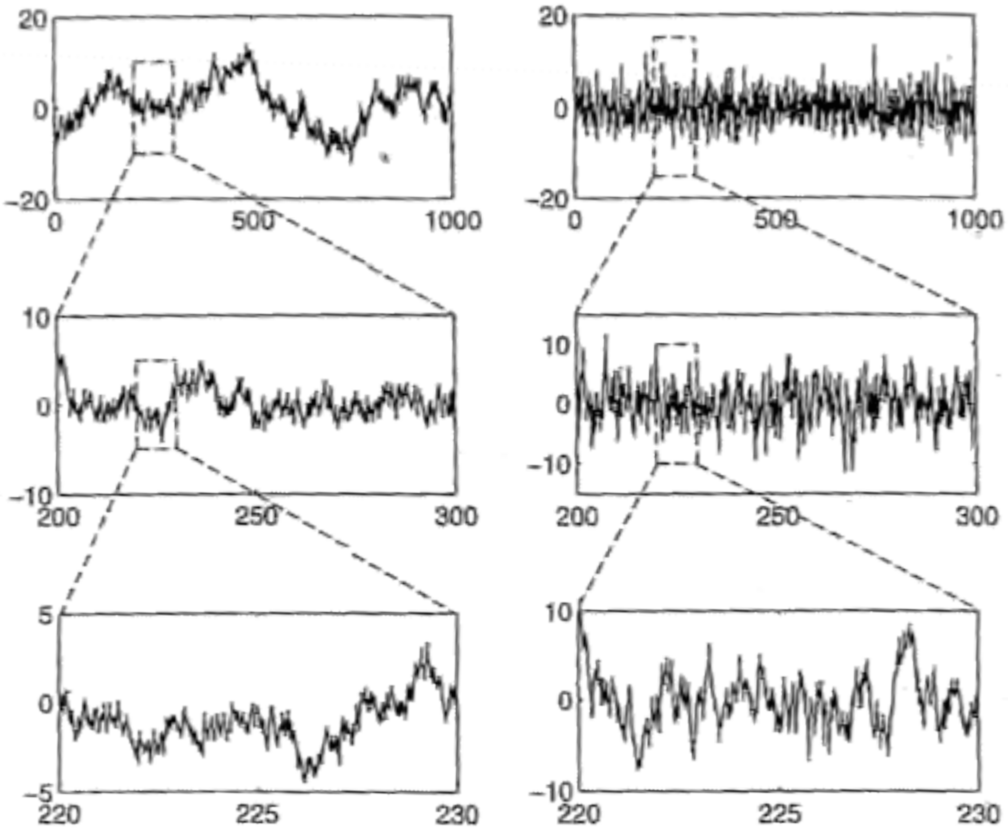
Escalas en el tráfico

Autosimilar/Fractal

- Al aumentar la escala tiene un comportamiento similar

Poisson y SRD

- Variabilidad se reduce rápidamente con la escala



(a) Self-similar process (b) Non-self-similar process

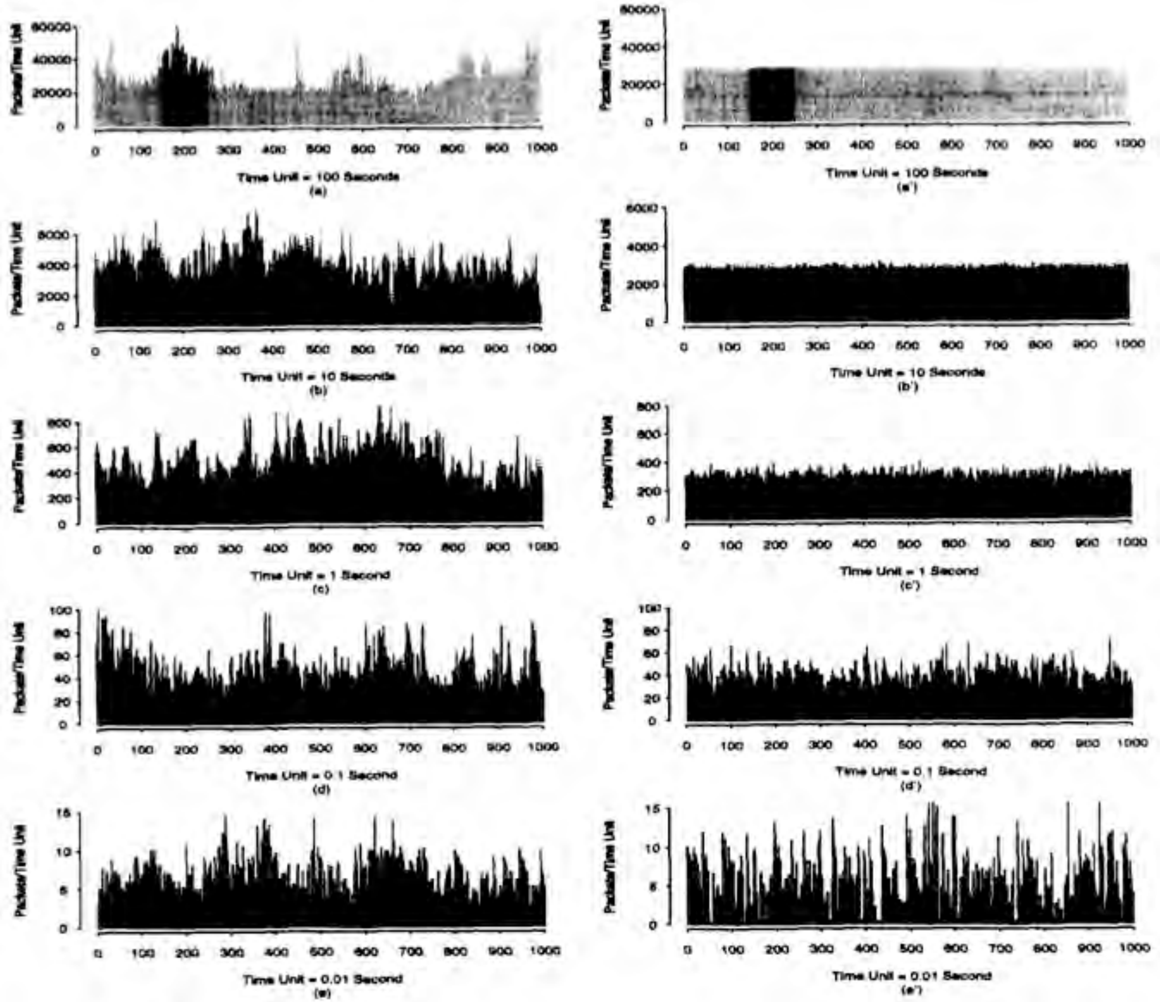
Escalas en el tráfico

Autosimilar/Fractal

- Al aumentar la escala tiene un comportamiento similar

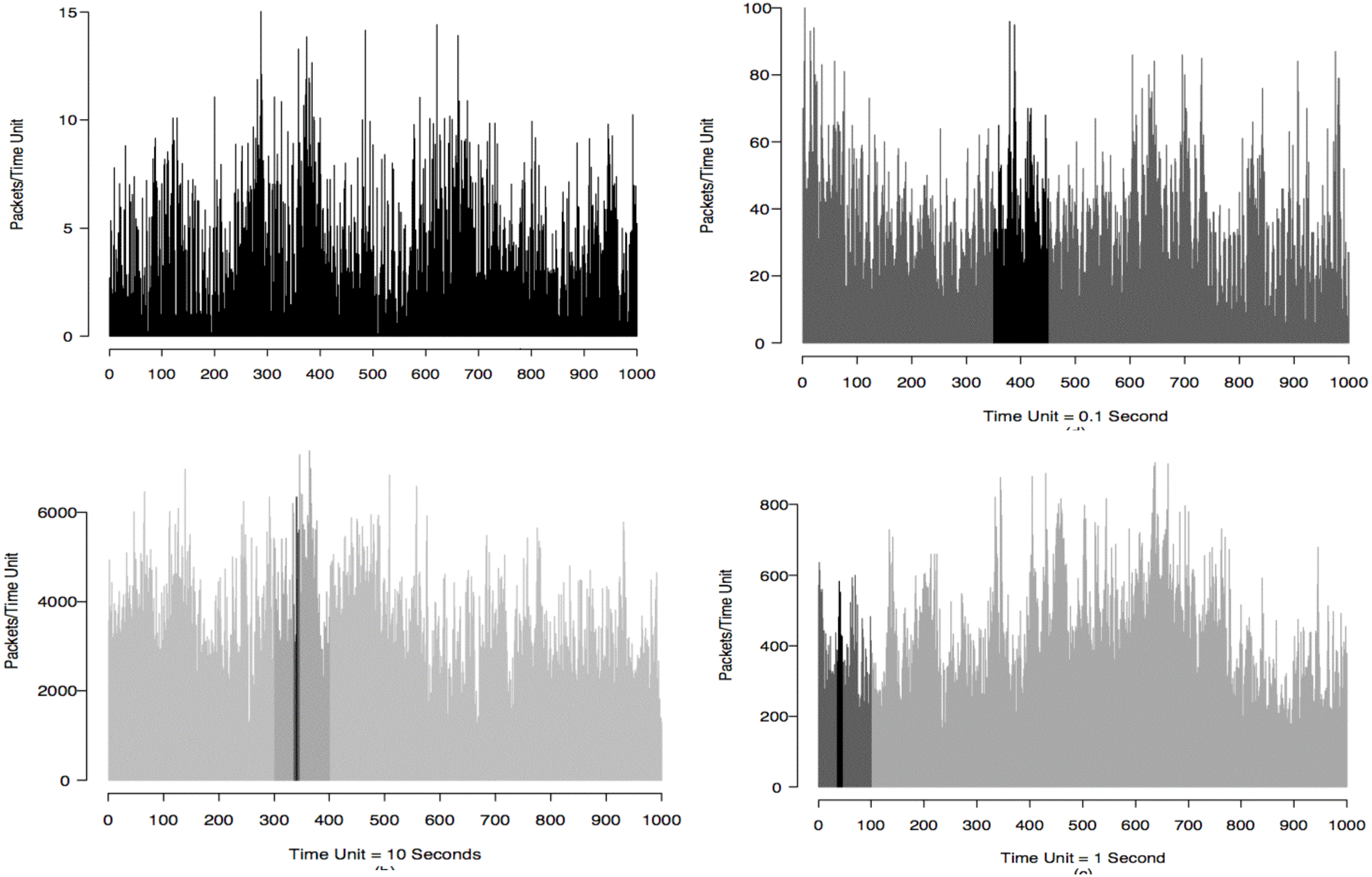
Poisson y SRD

- Variabilidad se reduce rápidamente con la escala



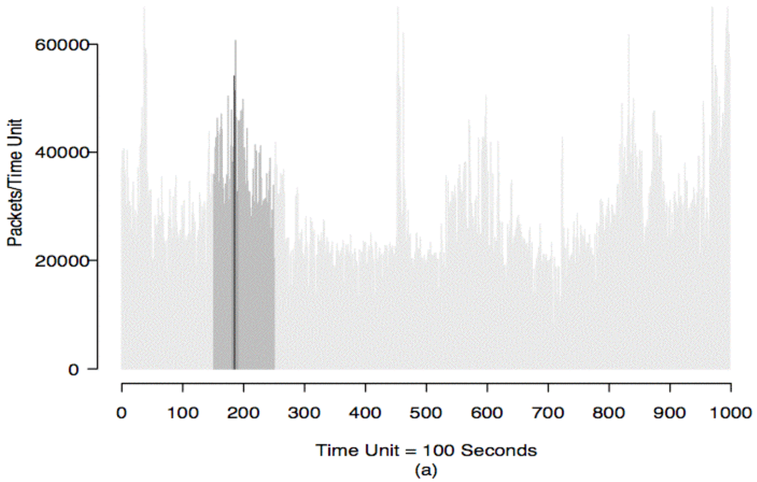
Trazas Bellcore 80-90

- Mayoritariamente IP
- Login remoto, e-mail, NFS, radio



Trazas Bellcore 80-90

- Mayoritariamente IP
- Login remoto, e-mail, NFS, radio



$$R(d)/S(d) = \frac{\max(0,W_1,W_2,\ldots,W_d) - \min(0,W_1,W_2,\ldots,W_d)}{S(d)}$$

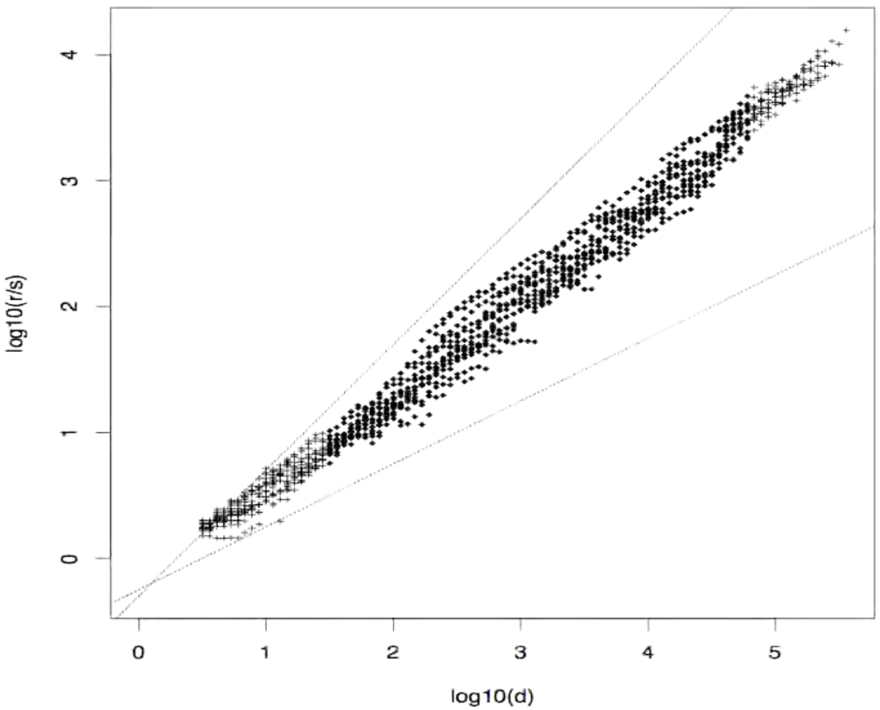
$$W_k = (X_1 + X_2 + \ldots + X_k) - k\overline{X}(d)$$

$\overline{X}(d)$ media muestral

$S^2(d)$ varianza muestral

- "Rango reescalado" ("*rescaled adjusted range statistic*")
- Power law scaling

$$E[R(d)/S(d)] \underset{d\to\infty}{\sim} d^H$$



Statistically Self-Similar

- x(t) es estadísticamente auto-similar con parámetro H si

$$x(t) \underset{d}{=} a^{-H} x(at) \qquad 0.5 \le H \le 1$$

donde la igualdad es en sus propiedades estadísticas (media, varianza, autocorrelación)

- Ejemplo: Fractional Brownian Motion (FBM)

Tiempo discreto

- Para un proceso estacionario de tiempo discreto x se define la versión agregada y reescalada:

$$x_k^{(m)} = \frac{1}{m} \sum_{i=km-(m-1)}^{km} x_i$$

- El proceso es ***exactly second-order self-similar*** con parámetro β si para todo m=1,2,…

$$R_{x^{(m)}}(k) = R_x(k)$$

$$Var\left(x^{(m)}\right) = \frac{Var(x)}{m^{\beta}}$$

- $H = 1 - \beta/2$
- Se dice que es ***asymptotically second-order self-similar*** si lo anterior se cumple solo para m→∞

Consecuencias de SS

- Para un proceso no auto-similar lo normal es que β=1, es decir, la varianza caiga con 1/m:

$$Var\left(x^{(m)}\right) = \frac{Var(x)}{m}$$

- y la autocorrelación tienda a cero

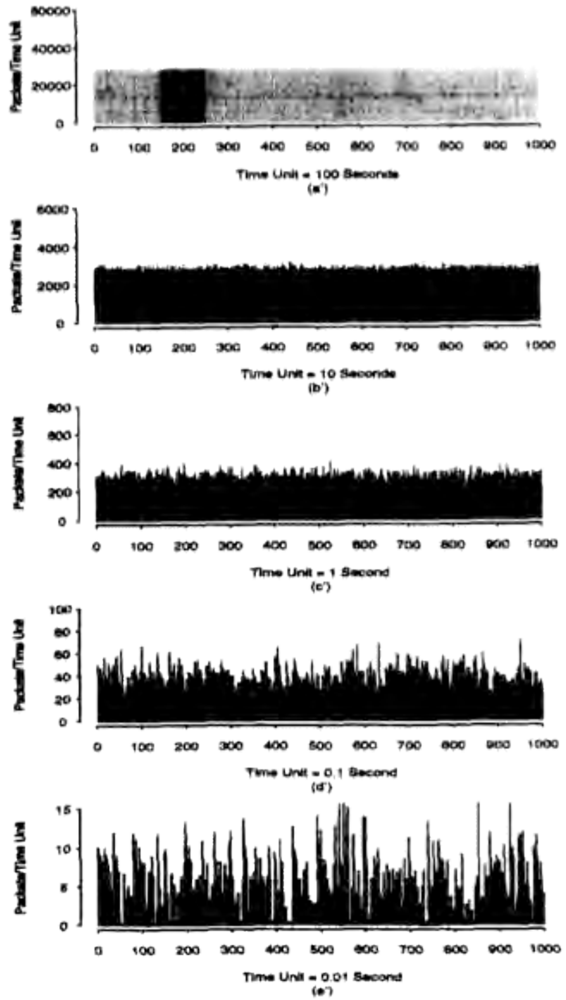
$$R_{x^{(m)}}(k) \underset{m\to\infty}{\longrightarrow} 0$$

- con lo que cada vez se parece más a un ruido blanco
- Modelos para decidir el BW o buffer que se basen en esto darán valores optimistas ante tráfico SS
- La autocovarianza (0.5 < H < 1):

$$C(k) \sim |k|^{-\beta}, \quad \text{as } |k| \to \infty, \quad 0 < \beta < 1$$

$$\sum_k C(k) = \infty$$

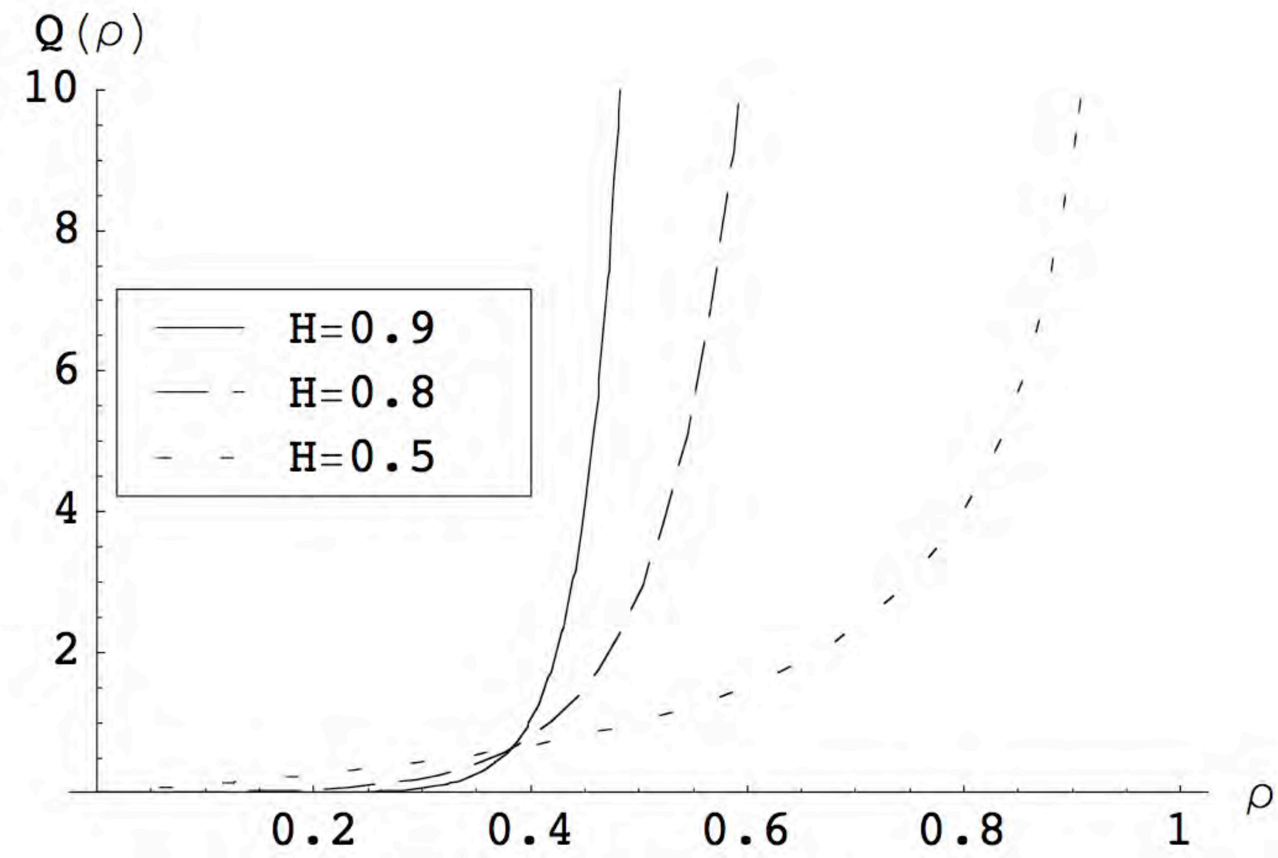
¡ LRD !



Consecuencias de SS

Ejemplo

- Longitud media de cola
- H = 0.5 (M/M/1)
- Para H > 0.5 las ráfagas hacen crecer la cola



Artículos

Objetivos de la presentación

- Explicar lo que has/han hecho y que se entienda
- Convencer de que
  - El tema es interesante
  - Es importante
  - Has aportado algo
  - Lo has hecho correctamente
- ¡ Pero solo si es así !
- Vender el trabajo (marketing)
- Pero sin engañar (ciencia)
- Si lo necesitas para entender el paper consulta sus referencias

Cómo presentarlos

- Di qué artículo vas a presentar y dónde está publicado
- Quiénes son los autores, de dónde son
- Índice de la presentación
- Introducción: tema general, antecedentes
- Escenario: tema concreto, escenario concreto
- Qué hacen
- Qué resultados obtienen
- Qué conclusiones se extraen
- Qué te ha parecido útil
- Qué te parece inútil, incorrecto o mejorable
- Básico:
  - Entérate tú
  - Consigue que se enteren los demás
- 30 minutos

Discussant

- Leer el paper
- Preparar crítica al mismo
  - Lo que no se entiende
  - Lo que no se explica bien
  - Lo que no parece correcto
  - Lo que se podría mejorar y cómo
- 5-10 min
- A partir de ahí discusión general

Traffic Analysis

- Introducción -

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

Programa de Tecnologías para la gestión distribuida
de la información

Ejercicios de simulación

1. Bicing
2. Cafetería del Sario
3. Servidor de streaming
4. P2P
5. Cache
6. Flujos