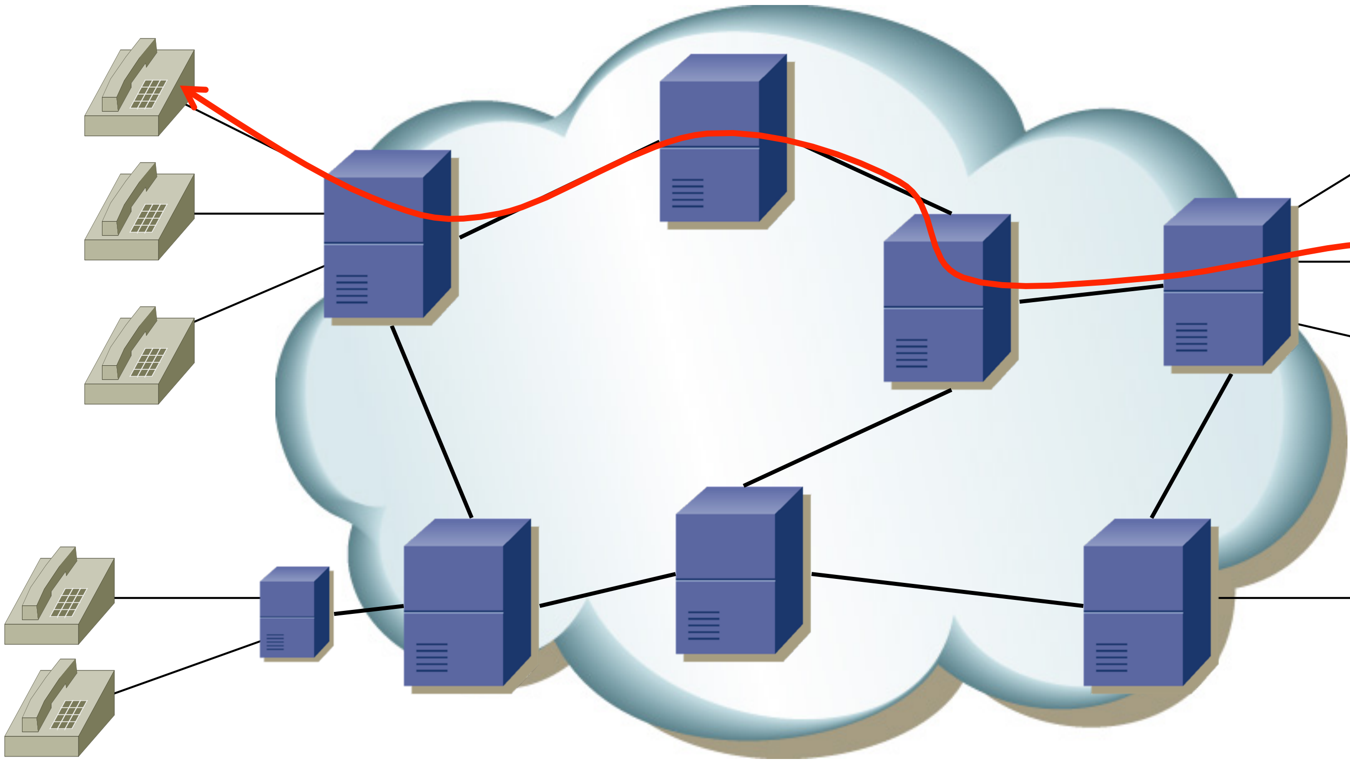
Ingeniería de Teletráfico

Objetivos del tema

- Introducción a la problemática
- Caso de dimensionamiento de redes con bloqueo
- Escenarios donde llegan solicitudes de servicio
- Si no pueden ser atendidas son rechazadas

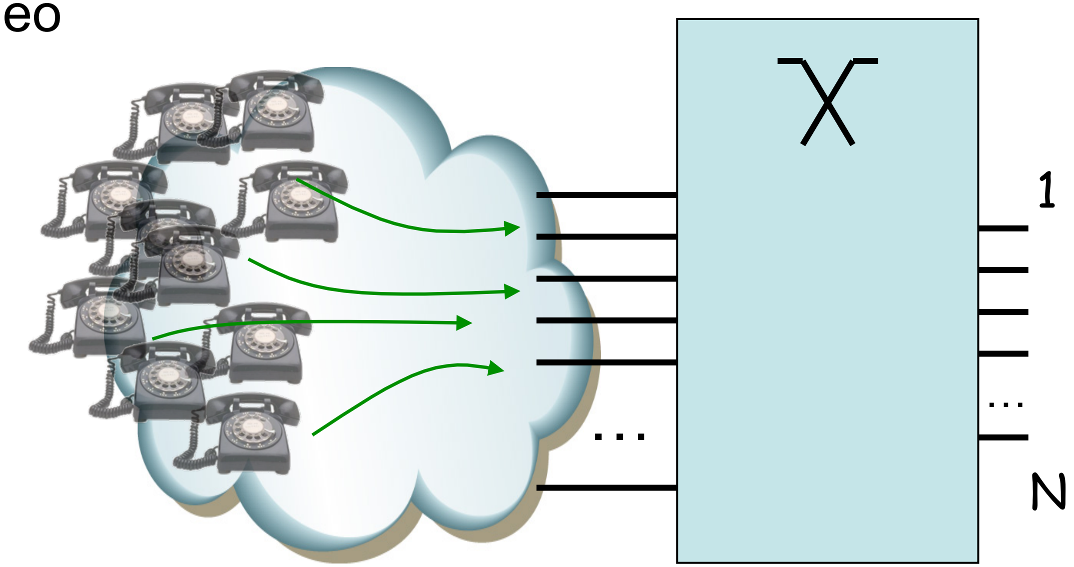
Problema en Telefonía

- Cuántos canales poner en los enlaces troncales para que “casi siempre” se puedan establecer las llamadas
- Caso peor es sobredimensionamiento

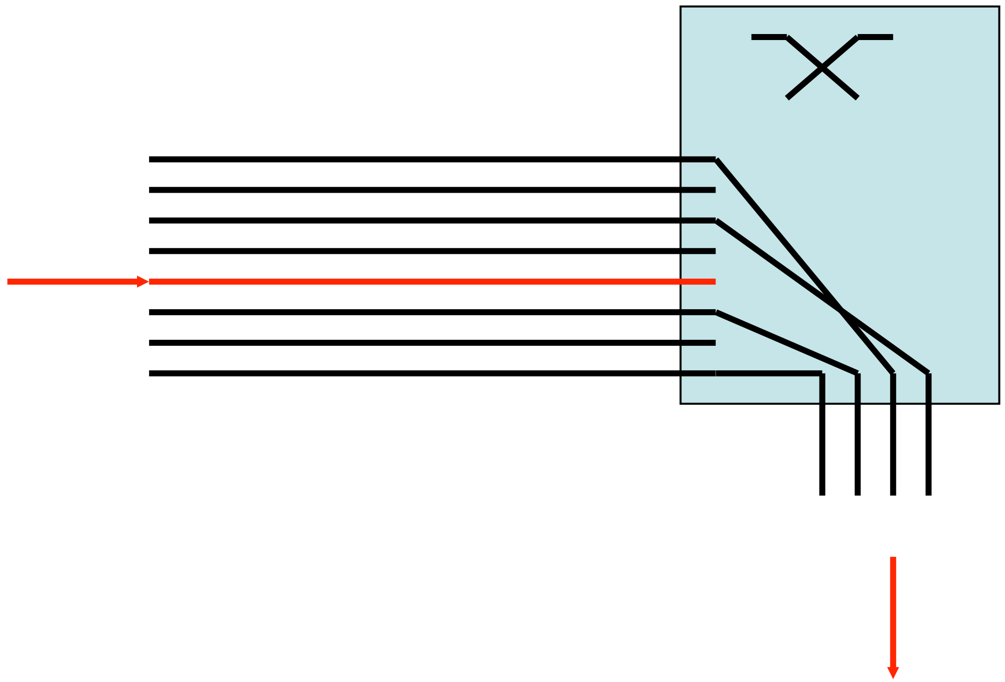


Problema tipo a resolver

- Solo vamos a ver un caso muy simple, donde no hay cola
- Conmutador con líneas de entrada y de salida
- Entradas usuarios finales o troncales: lo que nos importará es la llegada de llamadas al conmutador
- Salidas troncales (máximo N llamadas simultáneas)
- Llamadas entrantes o salientes
- Decidir N para poder cursar las llamadas con una probabilidad de ***bloqueo*** máxima objetivo
- o decidir la cantidad de usuarios con un N y ese máximo bloqueo

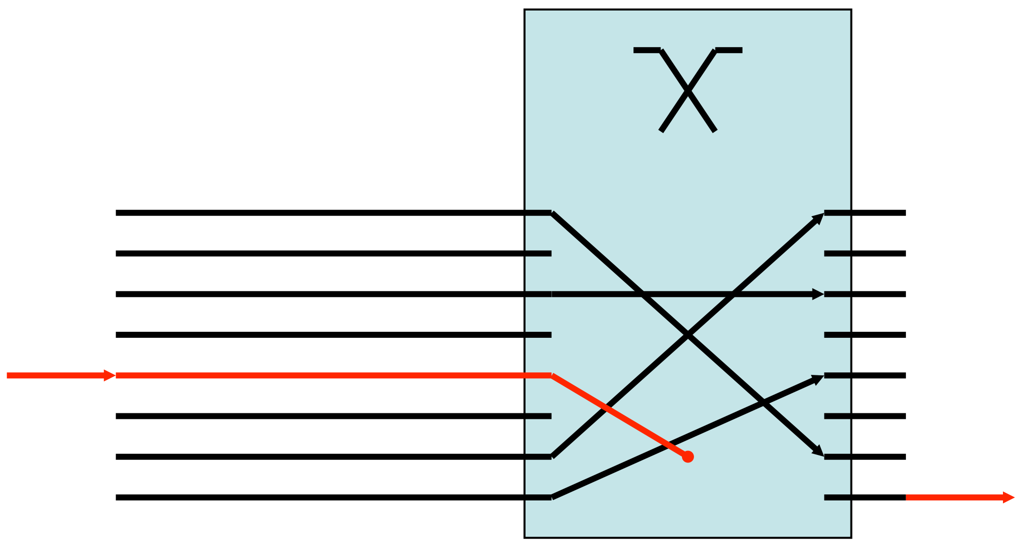


- **Bloqueo externo**
  - No se puede interconectar dos estaciones aunque estén libres
  - El conmutador no tiene suficientes recursos de salida para cursar una nueva llamada



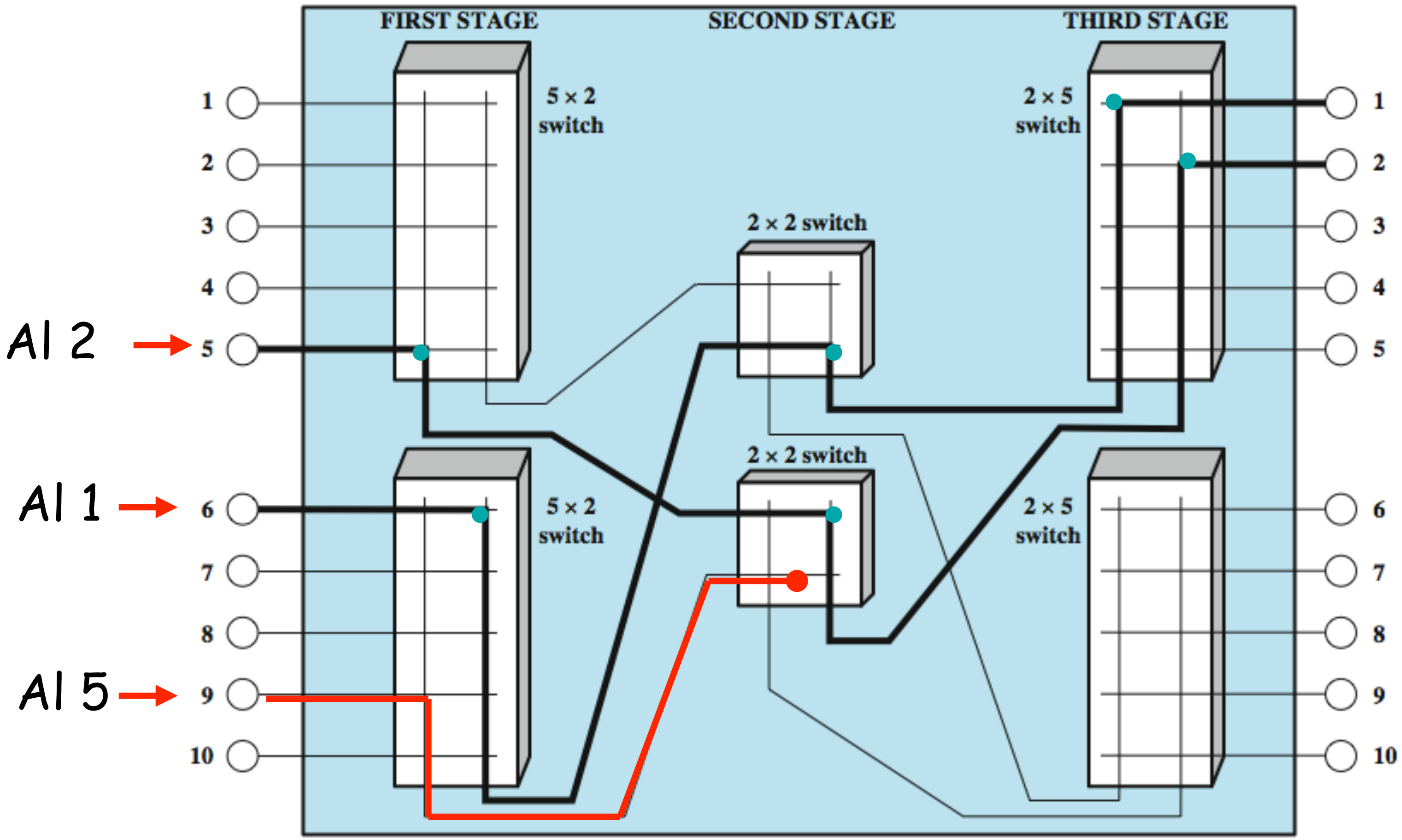
Bloqueo

- Bloqueo interno
  - El conmutador no tiene recursos internos para interconectar una entrada con una salida (…)



Bloqueo

- Bloqueo interno
  - El conmutador no tiene recursos internos para interconectar una entrada con una salida. Ejemplo (…)



Definiciones

Capacidad

- Recursos de un sistema para dar un servicio, número de líneas de salida…
- Ej: nuestra centralita tiene 5 líneas para llamadas salientes

Carga (Intensidad de trafico)

- Cantidad de servicio demandada al sistema, medida como cantidad de recursos necesarios en un determinado momento
- Ej: nuestra centralita tiene en media 3.2 llamadas con el exterior

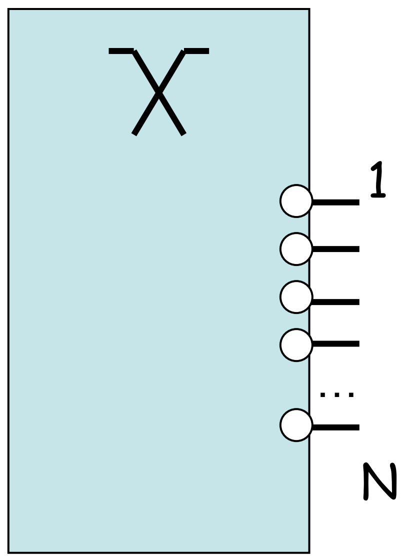
Calidad de servicio

- Medida del servicio obtenido del sistema
- Ej: nuestra centralita con las líneas de entrada que tenemos y la carga típica que soporta pierde menos del 0.1% de las llamadas

A continuación en más detalle…

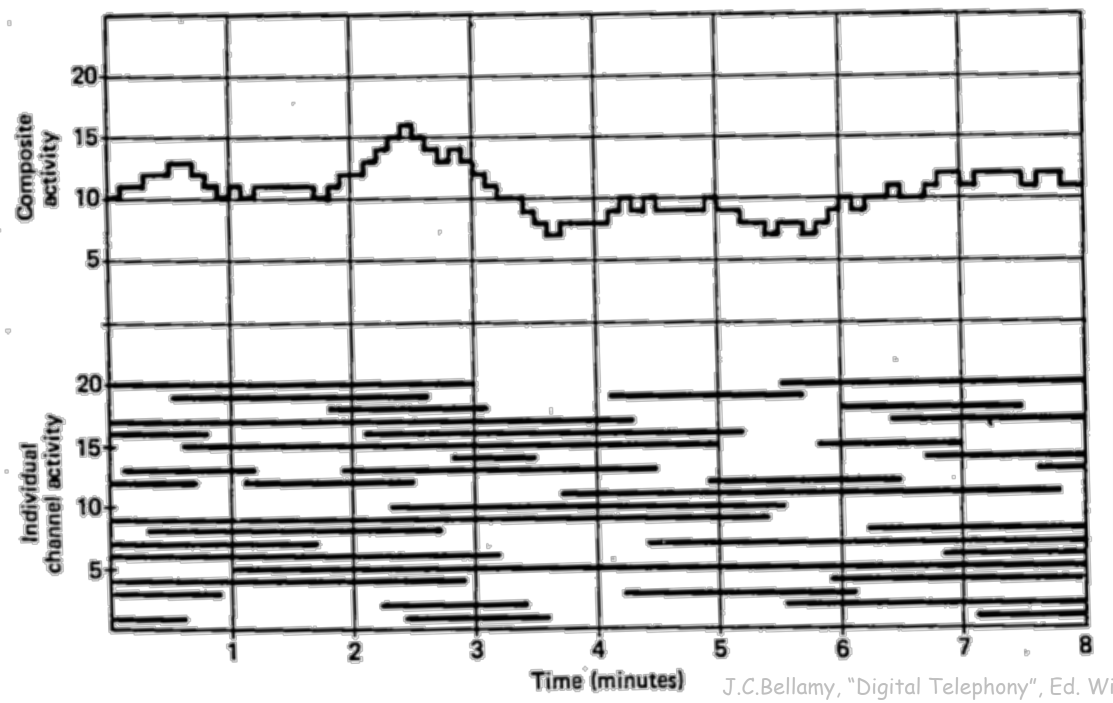
Capacidad

- Medida de la habilidad del sistema para proporcionar servicio
- Típicamente se mide como el número de servidores (líneas de salida, puertos de un conmutador…)
- Variable de diseño del sistema
- Coste proporcional
  - Más capacidad = más coste y más calidad de servicio

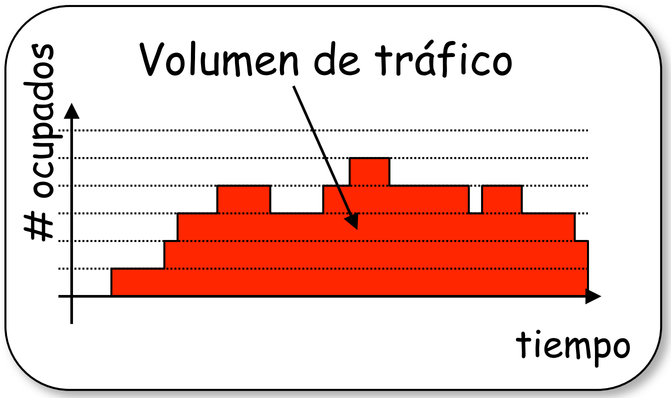


Carga o Tráfico

- Medida de la demanda de servicio al sistema
- Agregación de todas las peticiones de servicio de los usuarios
- = recursos en uso del sistema bajo condiciones de servicio ideales
- **Aleatorio**
  - Peticiones de servicio llegan de forma aleatoria
  - Solicitan servicio durante una cantidad de tiempo no predecible
- Volumen de tráfico: suma de las duraciones de los servicios



J.C.Bellamy, "Digital Telephony", Ed. Wiley Interscience



Carga o Tráfico

- Depende de
  - Número de usuarios (n)
  - Tasa a la que generan llamadas (λ_i)
  - Duración de las llamadas (s)
- El servidor no distingue el efecto de *n* o de λ_i
  - Ej: 600 usuarios, cada uno con una petición por hora, es equivalente a 10 usuarios con una petición por minuto cada uno
- Se reduce a:
  - Tasa de generación de llamadas de todos los usuarios (λ)
  - Duración de las llamadas (s)
- El primer paso del análisis de tráfico es la caracterización de las llegadas de peticiones y la duración de las mismas

Medida del Tráfico

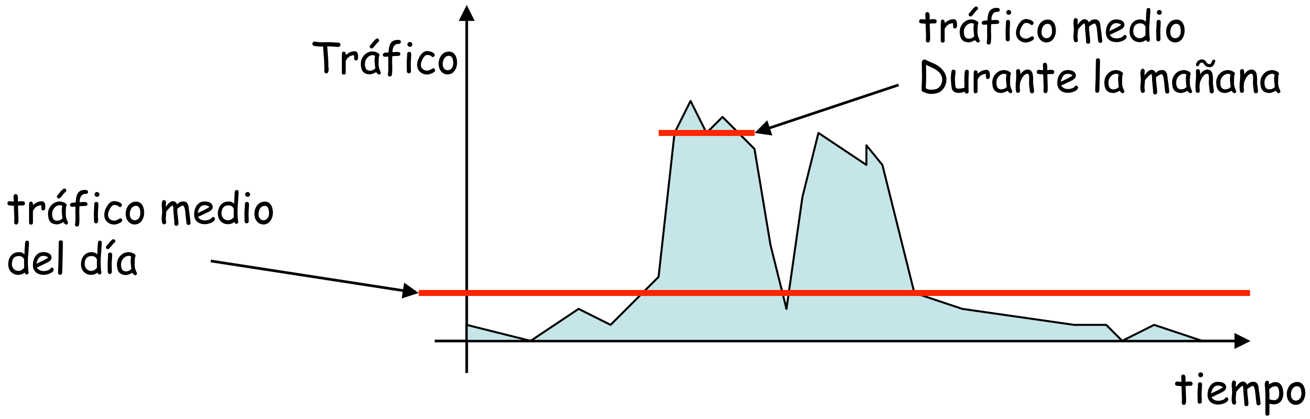
- Intensidad de tráfico

$$I = \frac{\text{Volumen de tráfico}}{\text{Tiempo de observación}} = \frac{\text{Tiempo acumulado de ocupación}}{\text{Tiempo de observación}}$$

- Sin unidades físicas. Se mide en *Erlangs (E)*
- **1 Erlang** = el tráfico que mantiene ocupada completamente una línea durante el tiempo de observación
- Ejemplo:
  - 600 usuarios, cada uno en media hace 1 llamada por hora
  - El tiempo medio de duración de las llamadas es de 3 minutos
  - ¿Cuanto tráfico representan?
  - Tiempo observación: 60 minutos
  - T. acumulado de ocupación : 600llamadas x 3minutos/llamada = 1800min
  - 1800/60 = 30 Erlangs
  - ¿Significa esto que necesitamos 30 líneas?

Medida del Tráfico

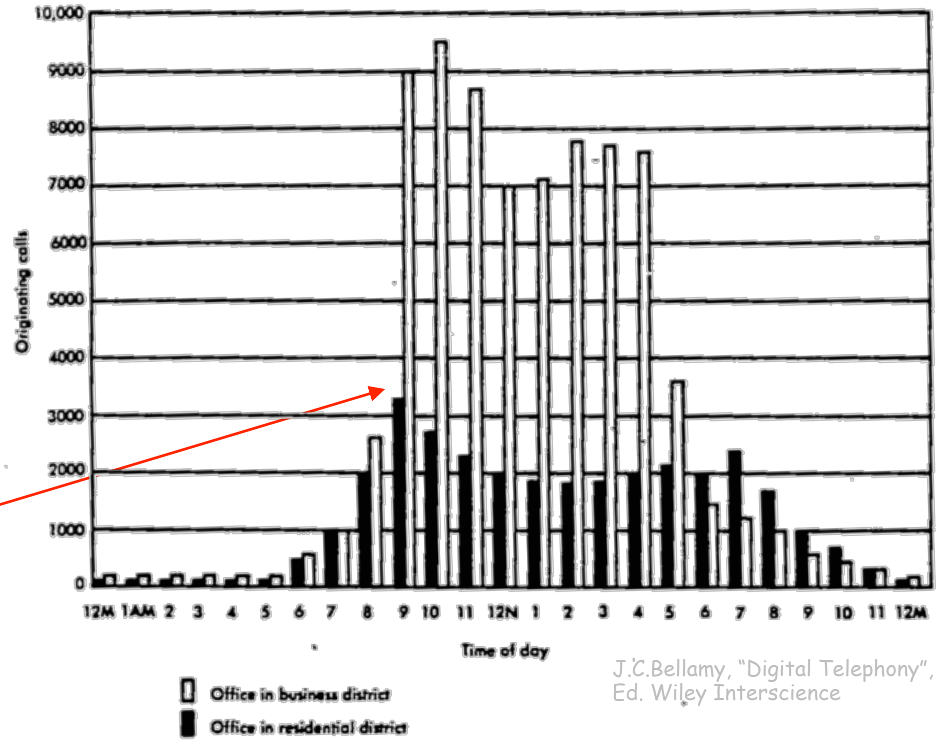
- Normalmente la intensidad del tráfico varía con el tiempo (no es un proceso estocástico estacionario) pero se puede considerar estable en un tiempo limitado



- En telefonía se caracteriza por horas
- Varía entre meses, entre días y entre horas del mismo día (y dentro de la hora)
- Suele haber patrones semanales
- Días de fiesta, el clima, etc. afectan al patrón

Hora cargada ("busy hour")

- Periodo de 60 minutos consecutivos durante los cuales el volumen de tráfico es máximo
- Los análisis para dimensionamiento de equipos se efectúan siempre sobre la **hora cargada**
- Para determinarla se suelen toman medidas en **intervalos de 15min** y entonces es el periodo de tiempo de 4 intervalos consecutivos con mayor volumen de tráfico
- Se calcula la hora cargada en un periodo largo (unas semanas) en la época del año de mayor tráfico
- Diferentes patrones usuarios residenciales y empresariales
- No es el volumen de tráfico mayor del año (nochevieja, día de la madre,...) pues llevaría a un sobredimensionamiento para la mayor parte del tiempo
- 1 teléfono en hora cargada approx. 0.05-0.1 E y 3-4min duración



J.C.Bellamy, "Digital Telephony", Ed. Wiley Interscience

“Calidad” de servicio

- Medida de la bondad del servicio proporcionado
- En telefonía:
  - Probabilidad de bloqueo = probabilidad de que el sistema no pueda aceptar una llamada entrante.

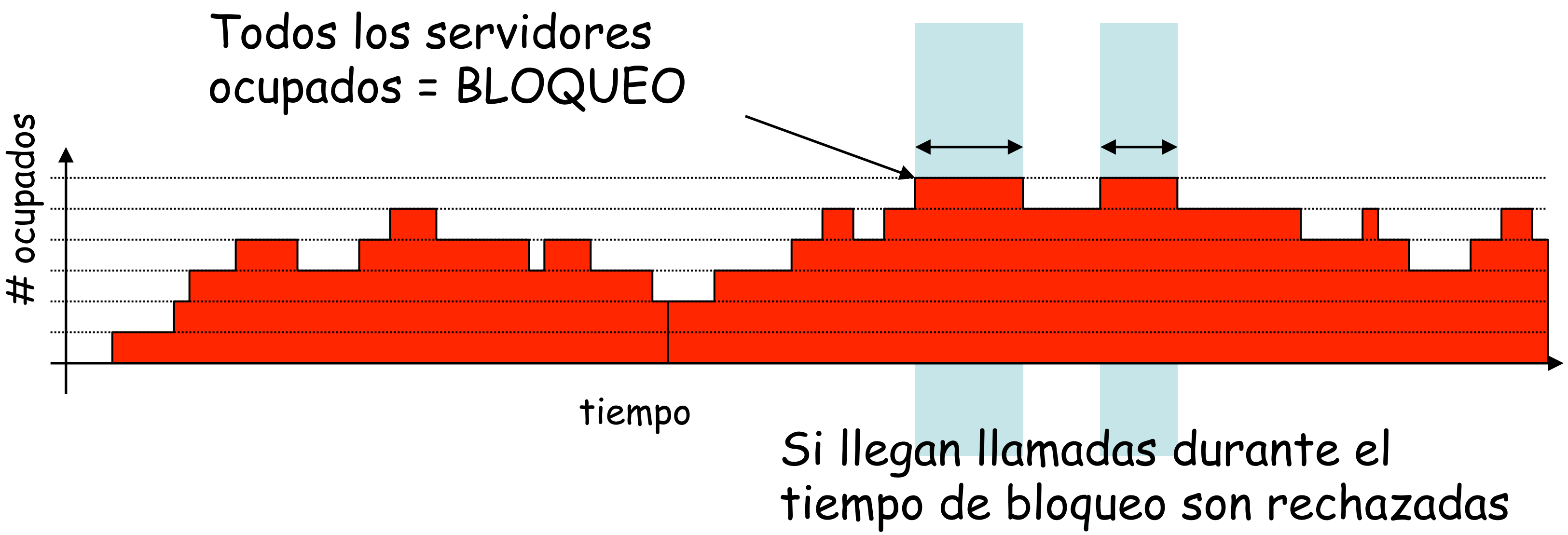
  En ese caso:
  - Se descarta: La llamada es rechazada y el usuario a veces no puede hacer una llamada → Menos calidad de servicio (*congestion theory*)
  - Se hace esperar la llamada hasta que se libere un servidor: El usuario a veces ve que sus llamadas tardan más en establecerse → Menos calidad de servicio (*queueing theory*)
- Requisito de diseño del sistema: probabilidad de bloqueo objetivo y dimensionar la capacidad para conseguirla
- Se suele distinguir:
  - Sistema en **situación de Bloqueo**

    Todos los recursos están ocupados y una llamada nueva que llegue será rechazada
  - Sistema en **situación de Congestión**

    Se han empezado a rechazar llamadas

Tráfico ofrecido vs cursado

- Tráfico ofrecido: el tráfico total que sería cursado por una red que pudiera dar servicio a todas las peticiones
- Diseño (por economía) hace que en ciertas situaciones no se pueda cursar todo el tráfico (llamadas bloqueadas)
- Asumiremos que las llamadas bloqueadas se “pierden” (no hay reintento)
- El tráfico cursado es siempre menor o igual al ofrecido



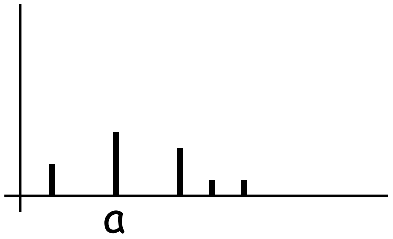
Modelando la carga

Variables aleatorias (V)

- No tiene un valor sino que describe el resultado de un experimento aleatorio
- Se caracteriza por la descripción de los posibles resultados que puede tomar en términos de probabilidad
- **Función de distribución / densidad de probabilidad**

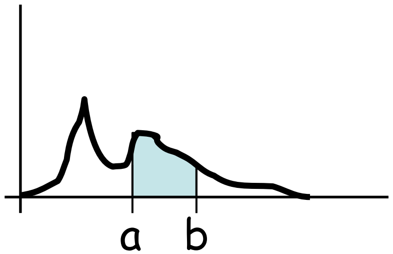
Variable discreta

$$p(a) = P[V = a]$$



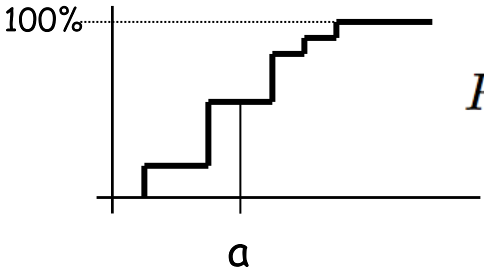
Variable continua

$$P[a < V < b] = \int_a^b p(t)dt$$



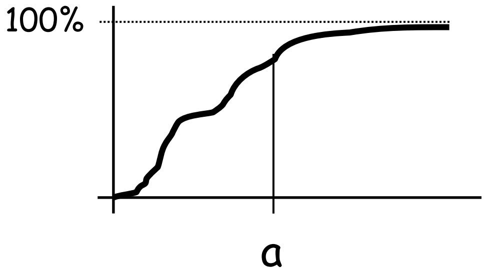
- **Función acumulada de probabilidad / distribución**

Variable discreta



$$P[V \leq a] = F(a)$$

Variable continua



Modelando la carga

Procesos estocásticos (V)

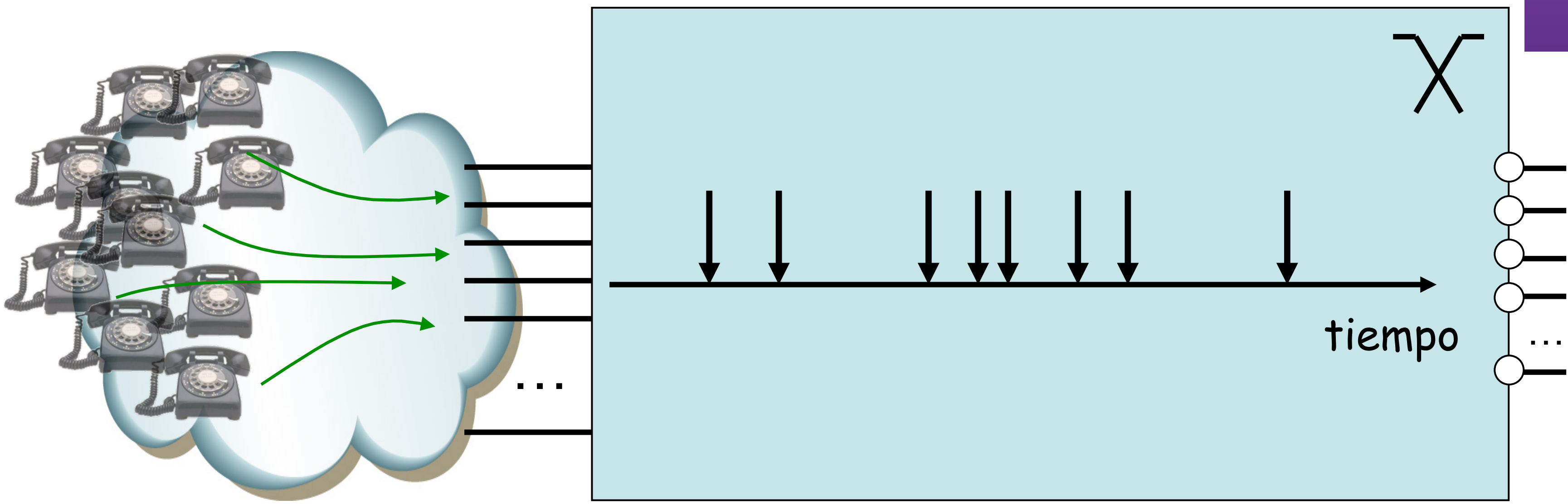
- Una familia de variables aleatorias

$$\{X_t : t \in T\}$$

- Hablaremos de
  - “Tiempo continuo” cuando T es real, por ejemplo T = [0,∞]
  - “Tiempo discreto” cuando T es numerable, por ejemplo T = {0,1,2…}

Proceso de llegadas

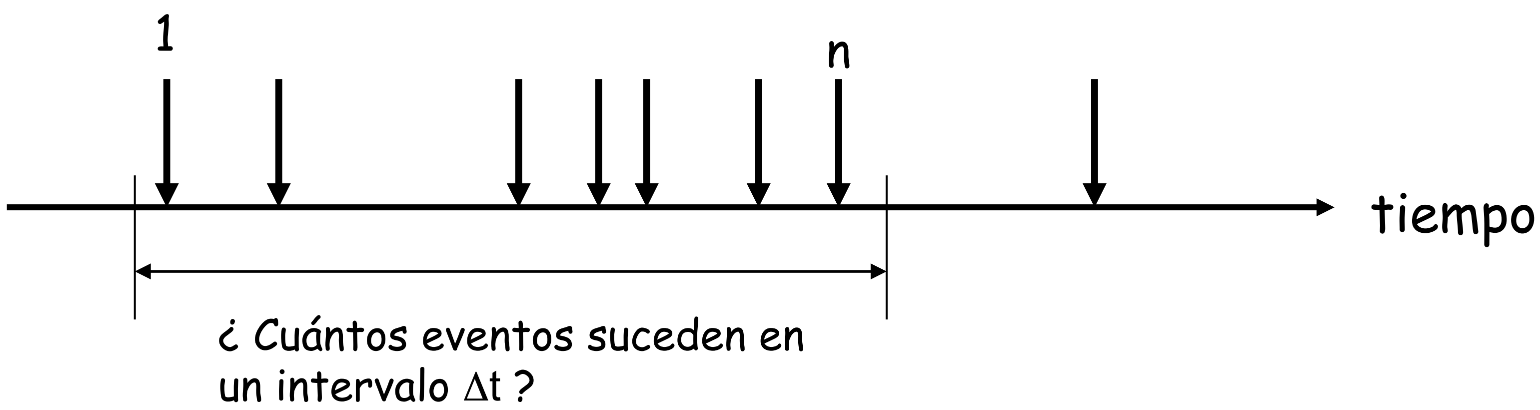
- Hipótesis fundamental en teoría clásica: llegadas independientes
- Tasa media de llegadas de llamadas de una gran población de fuentes (usuarios) independientes: λ



Número de llegadas

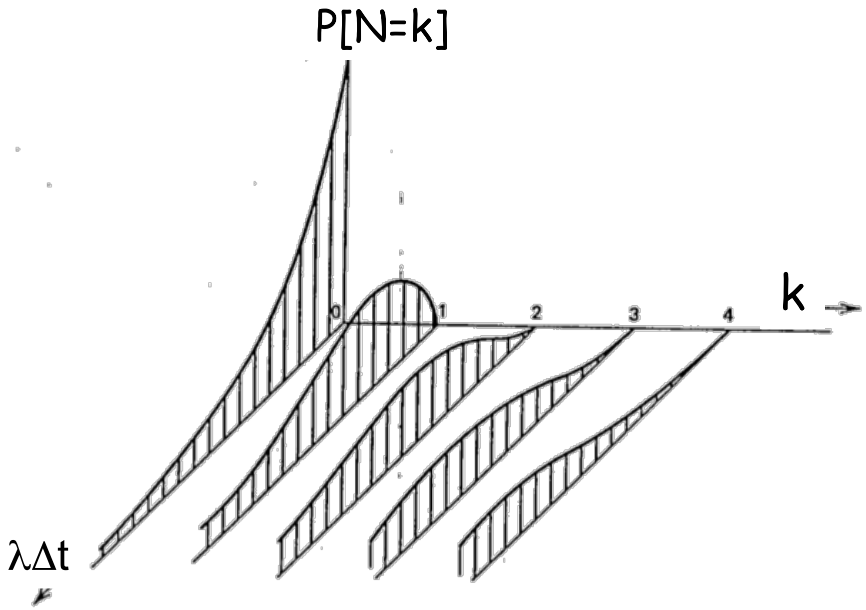
- Hipótesis:
  - En un intervalo suficientemente pequeño solo puede producirse una llegada
  - La probabilidad de una llegada en un intervalo suficientemente pequeño es directamente proporcional a la longitud del mismo (probabilidad $\lambda\Delta t$)
  - La probabilidad de una llegada en un intervalo 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:

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

- Su valor medio es λΔt :

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

Tiempos entre llegadas

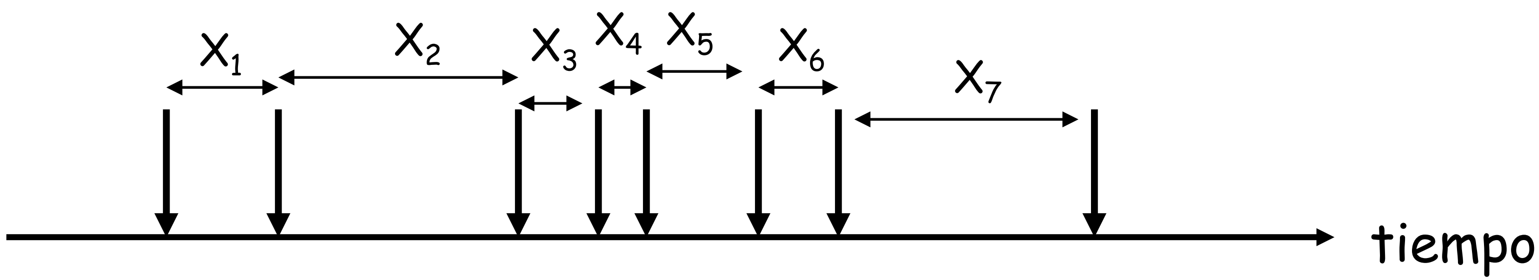
- Se demuestra que: si el número de eventos que ocurren en un intervalo sigue una distribución de Poisson los tiempos entre llegadas de eventos siguen una distribución exponencial
- El tiempo entre llegadas sigue una v.a. exponencial de parámetro λ
- X_i variables aleatorias independientes e idénticamente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

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

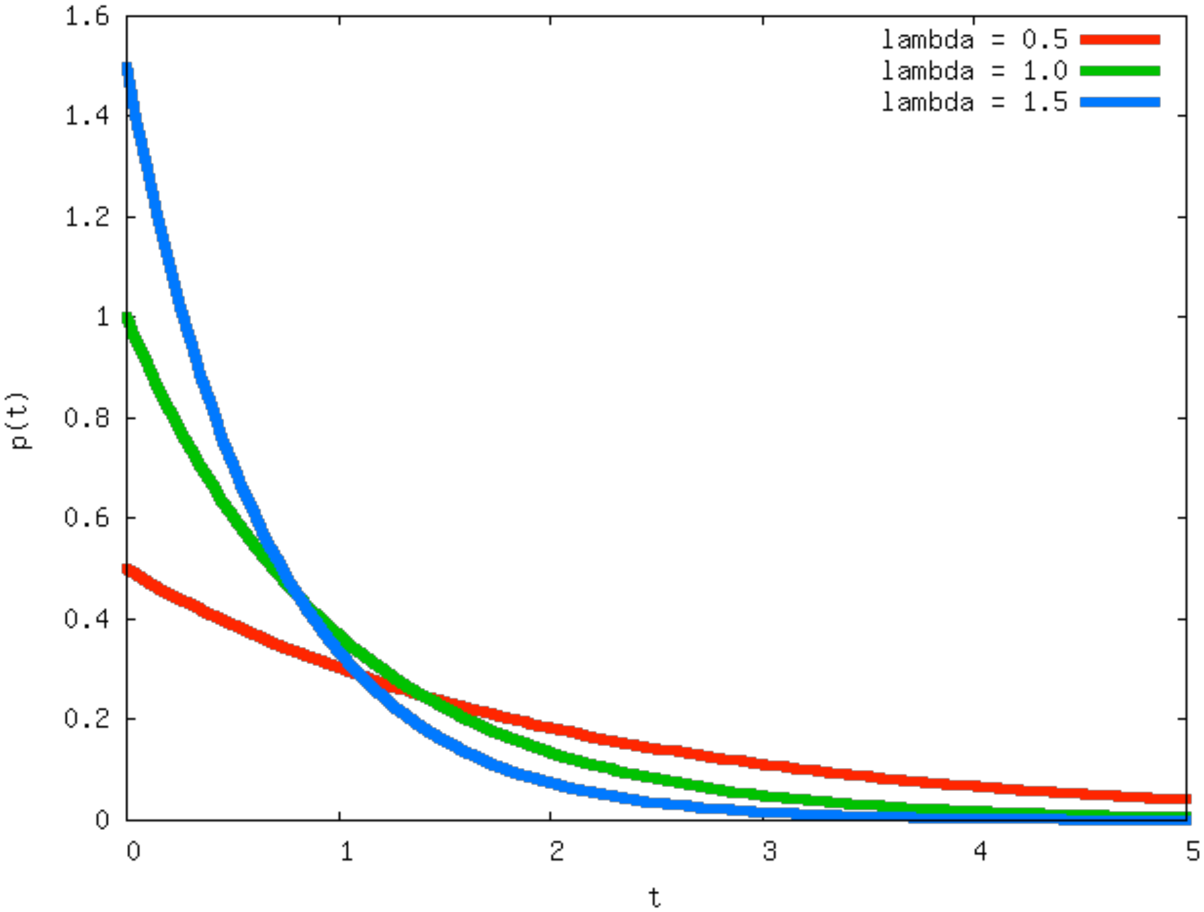
$$P[X < t] = 1 - e^{-\lambda t}$$

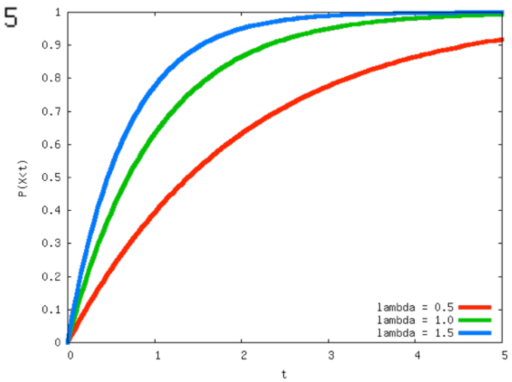
- Media: $E[X] = \int_0^\infty t\lambda e^{-\lambda t} = 1/\lambda$

- Tiempo medio entre llegadas 1/λ ⇒ en media λ llegadas por segundo

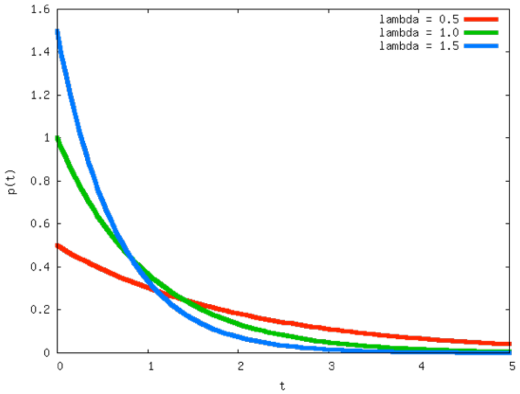


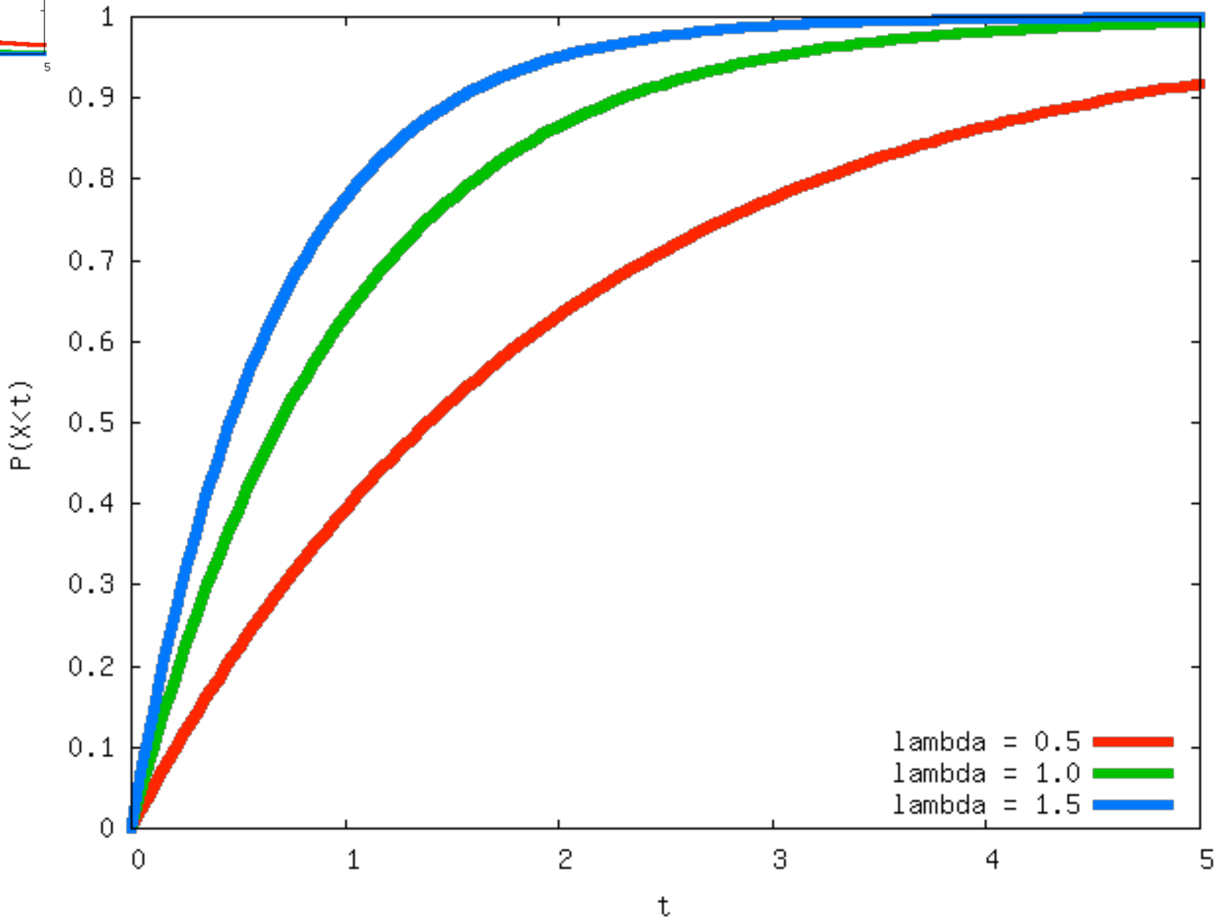
Variable aleatoria exponencial





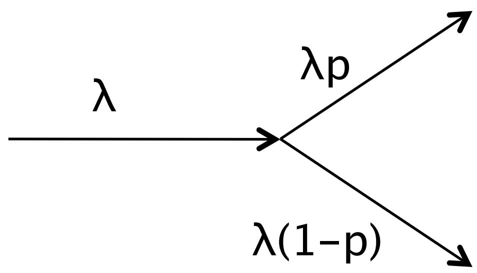
Variable aleatoria exponencial





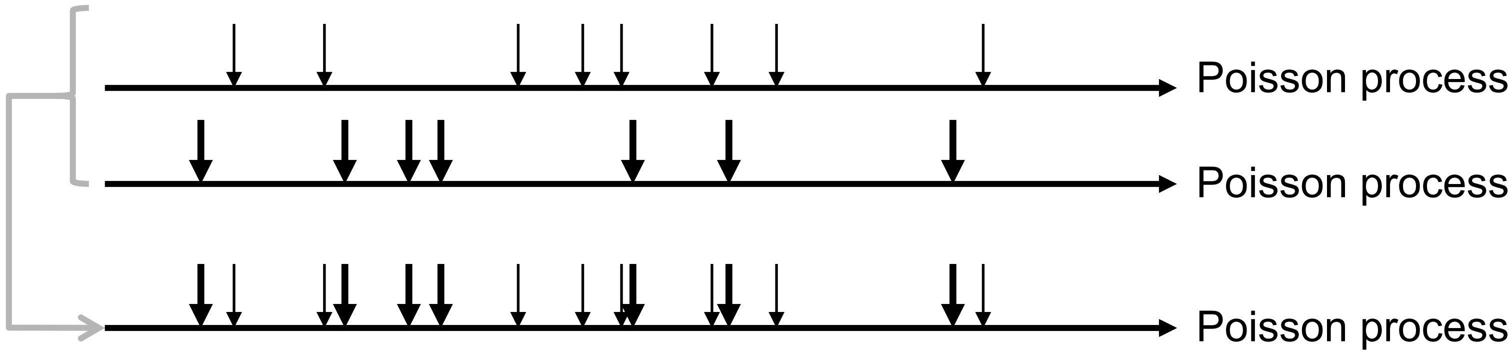
Random splitting

- Proceso de Poisson con tasa λ
- Repartidas las llegadas en dos grupos mediante Bernoulli de parámetro p
- Los procesos resultantes son procesos de Poisson de tasas λp y λ(1-p)

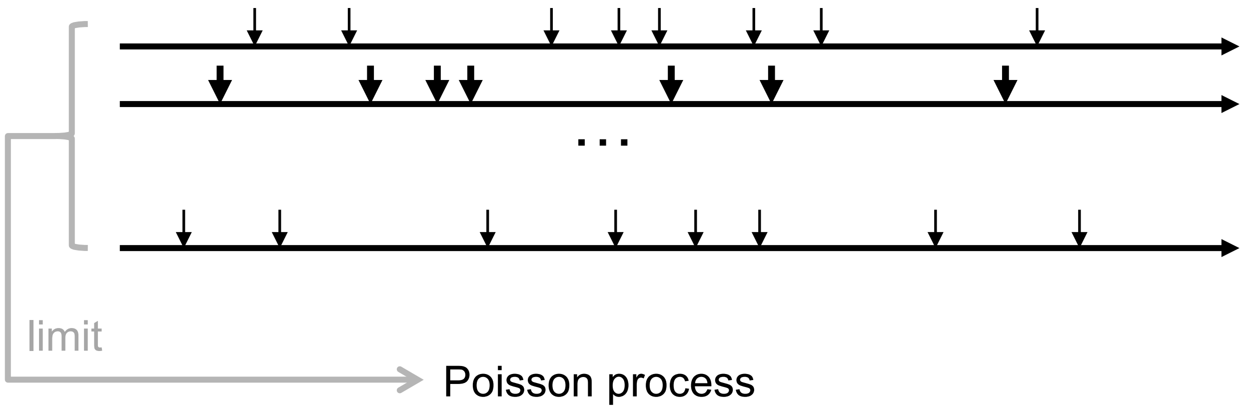


Superposición

- La superposición de dos procesos de Poisson es un proceso de Poisson de tasa la suma de las dos



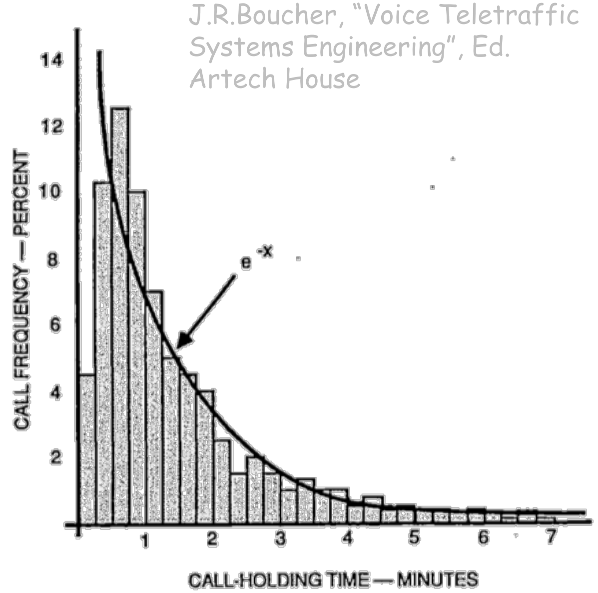
- Para ciertos procesos muy comunes (independientes), la superposición de un grán número de ellos tiende a un proceso de Poisson



Tiempo de ocupación

- Duración de las llamadas
- Lo más simple: tiempo constante
  - Poco realista para llamadas
  - Actividades automáticas: reproducción de mensajes, procesado de señalización, etc.
- Tiempo exponencial
  - Variables aleatorias (continuas) 's_i'
  - i.i.d. ('s')
  - Tiempos menores de la media muy comunes
  - Cada vez menos comunes tiempos mayores que la media
  - Propiedad: el tiempo restante de una llamada es independiente de lo que haya durado hasta ahora
- Duración exponencial: 's' caracterizada por su función de densidad

J.R.Boucher, "Voice Teletraffic Systems Engineering", Ed. Artech House



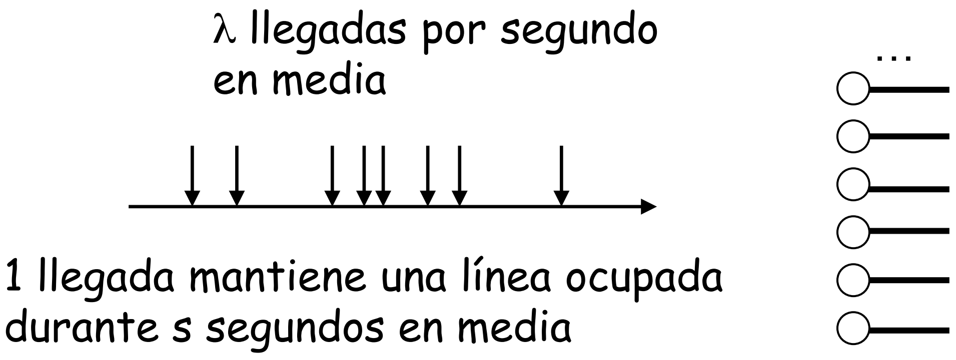
$$p_s(t) = \mu e^{-\mu t} \qquad \text{(t>0)}$$

$$\int_0^\infty \mu e^{-\mu t} = 1 \qquad \text{es una fdp}$$

$$\bar{s} = E[s] = 1/\mu$$

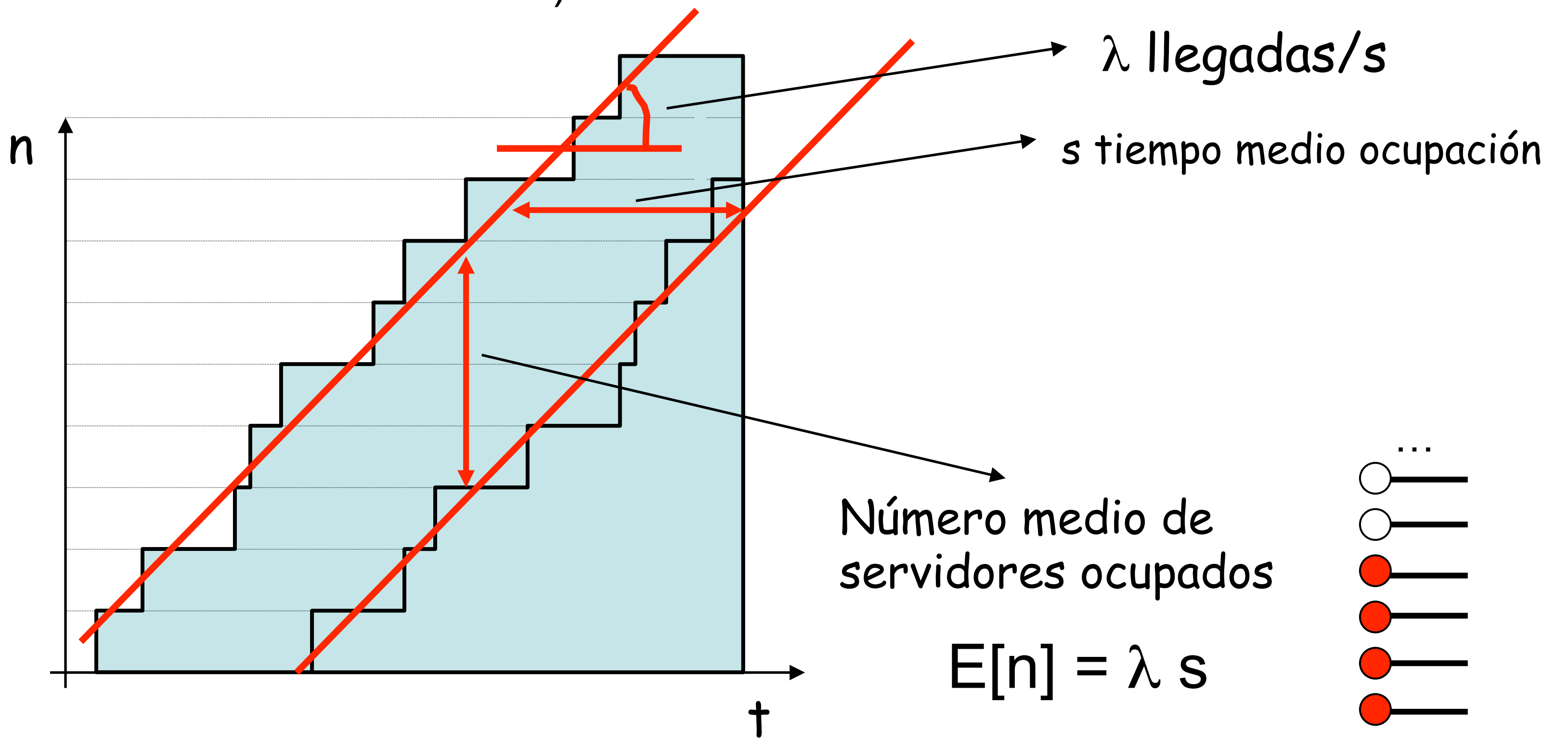
Intensidad de trafico

- Infinitas líneas
- Llamadas que se generan con una tasa media λ
- Tiempo medio de duración s
- ¿ Intensidad de tráfico que representan ?



Intensidad de trafico

- $E[n] = \lambda\ s$
- Esto es conocido como la Fórmula de Little
- $\lambda\ s$
  - Es el tráfico medido en Erlangs
  - Equivalente al número de recursos que se ocuparían en el sistema con esa carga si el sistema tuviera infinitos recursos (condiciones de servicio ideales)



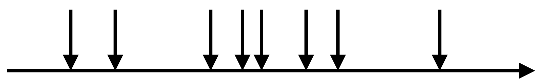
Número de líneas ocupadas

- Hipótesis:
  - Llamadas proceso de **Poisson** con tasa λ
  - Solicitudes de servicio de duración **constante** ‘s’
- ¿ Número de líneas ocupadas en un instante cualquiera ?
  - Es una variable aleatoria
  - La probabilidad de que ‘j’ líneas estén ocupadas en un instante es la probabilidad de ‘j’ llegadas en el intervalo previo de duración ‘s’
  - Depende solo de la intensidad de tráfico λs, que es la media de esta variable (A = λs)
  - Resulta ser válido independiente de la distribución de ‘s’ (sin demostración)

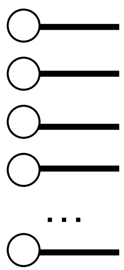
Intensidad de tráfico

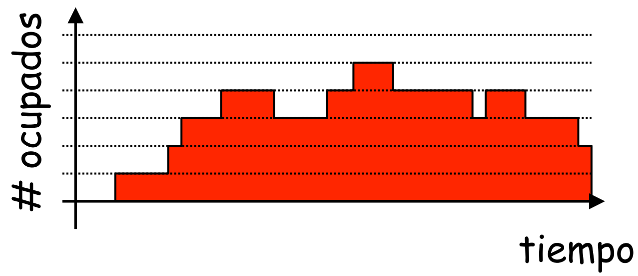
$$P_{\lambda s}[N=j] = \frac{(\lambda s)^j}{k!} e^{-\lambda s}$$

λ Llegadas
por segundo



1 llegada mantiene una línea ocupada
durante s segundos



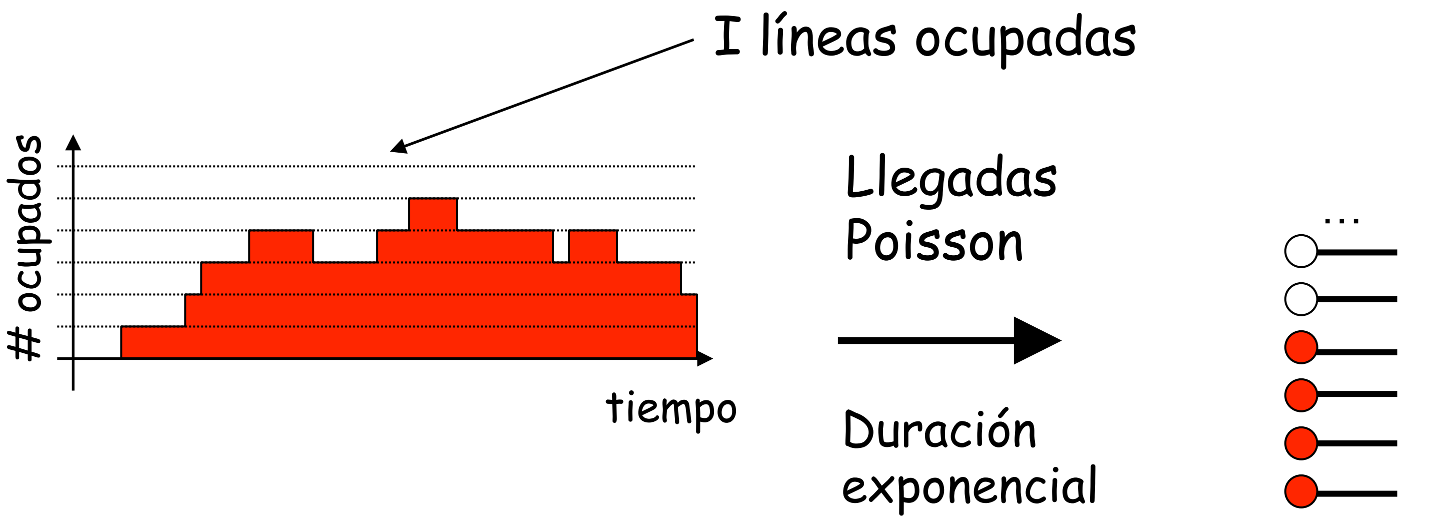


Recursos finitos

- Normalmente dispondremos de recursos finitos (capacidad)
- Problemas de interés
  - ¿ Cuál es la probabilidad de que una llamada encuentre el sistema ocupado ?
  - ¿ Cuál es el número de líneas necesarias para una probabilidad objetivo ?
  - ¿ Cuál es el tráfico que atraviesa ese sistema y forma la carga del siguiente sistema ?

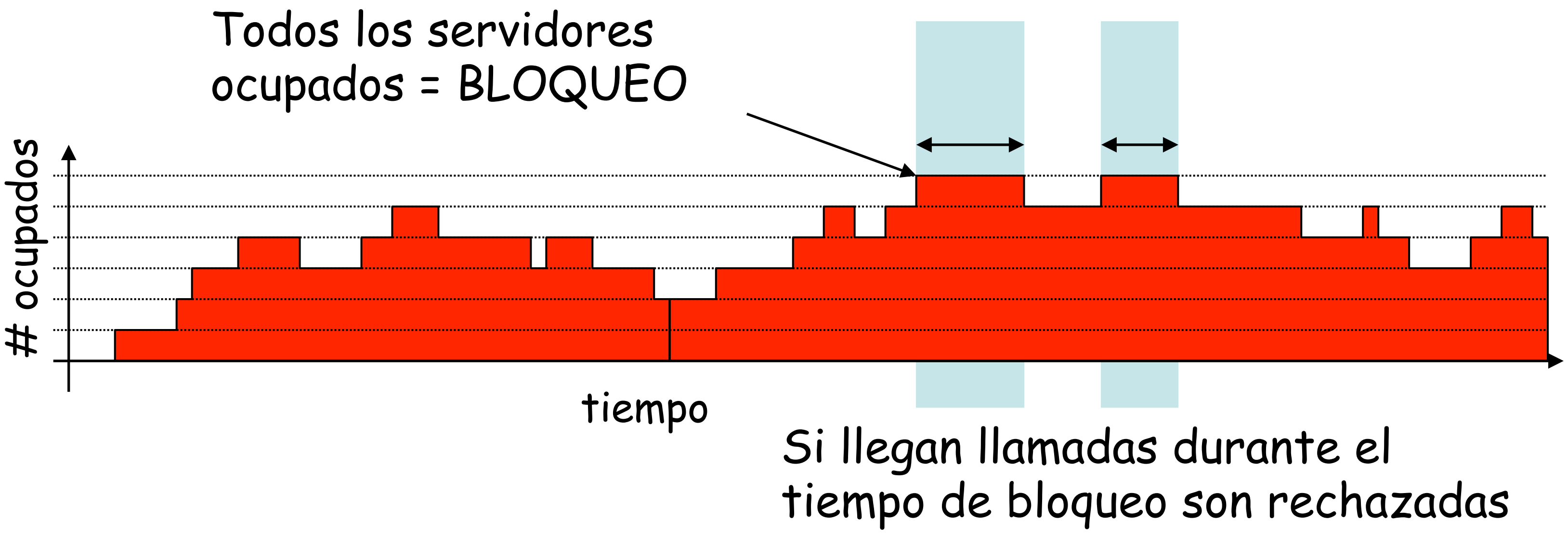
Probabilidad de bloqueo

- Llegadas según proceso de Poisson de tasa λ
- Duración exponencial de media s
- Variable aleatoria (o más bien proceso aleatorio)
  - I número de servidores ocupados en cada instante de tiempo
  - La intensidad de tráfico es $E[I] = A = \lambda s$



Probabilidad de bloqueo

- Cuando la variable `I` toma valor = número de servidores el sistema está en BLOQUEO
- ¿ Cuál es la probabilidad de que el sistema esté en situación de bloqueo ?



Probabilidad de bloqueo

- En un sistema con
  - Llegadas Poisson(λ)
  - Duraciones Exp(1/s)
  - Tráfico de entrada A = λs
  - k servidores
  - Las llamadas que llegan al sistema bloqueado se pierden
  - Probabilidad de bloqueo: ¿Cuál es P[I=n]? (…)
- P[I=n] = B(a,k)
- B(a,k) es conocida como función B de Erlang (o ErlangB)

B de Erlang

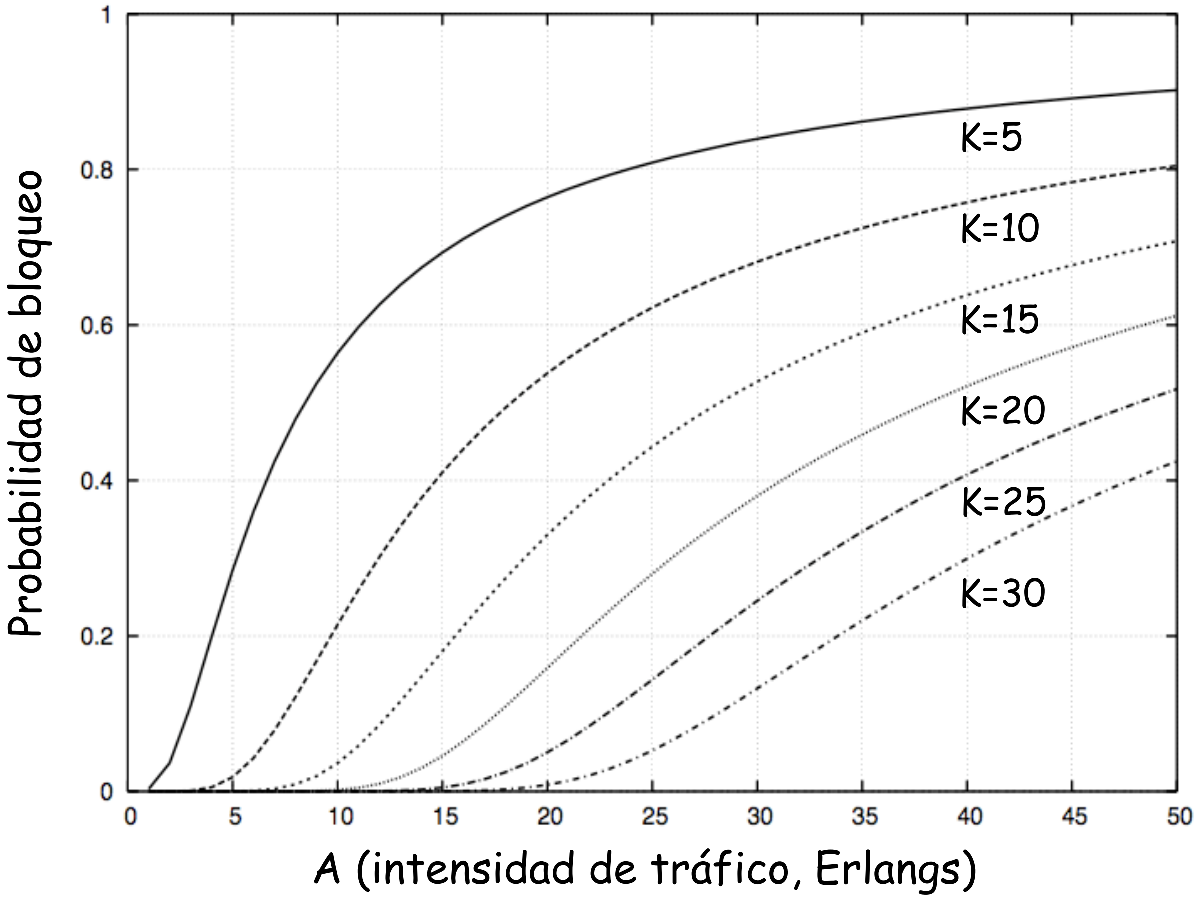
- Fórmula:

$$B(A,k) = \frac{\frac{A^k}{k!}}{\sum_{i=0}^{k} A^i/i!}$$

- Cálculo recursivo:

$$B(A,0) = 1$$

$$B(A,j) = \frac{A \cdot B(A,j-1)}{A \cdot B(A,j-1) + j}$$



Tráfico cursado

- Si un conjunto k de líneas tiene un tráfico ofrecido de I Erlangs y una probabilidad de bloqueo, ¿cuánto tráfico atraviesa las líneas?

  Esto será el **tráfico cursado** y será a su vez el tráfico ofrecido al siguiente sistema al que lleguen las líneas

  $I_c = I_{in}\,(1 - P_b) = I_{in}\,(1 - B(I_{in}, k))$

  I_c : tráfico cursado

  I_{in} : tráfico ofrecido o de entrada

Tráfico de desbordamiento

No puede ser cursado por el camino principal (por bloqueo)

Se “desborda” (overflow) a una ruta secundaria

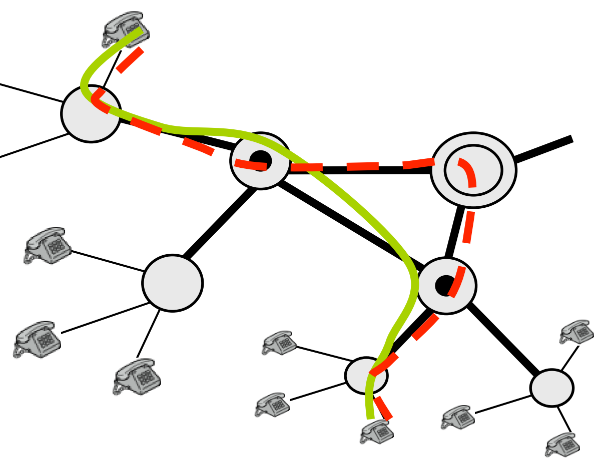
Un proceso de Poisson del que se eliminan aleatoriamente (iid) muestras con probabilidad p sigue siendo un proceso de Poisson, pero con menor tasa ($p\lambda$)

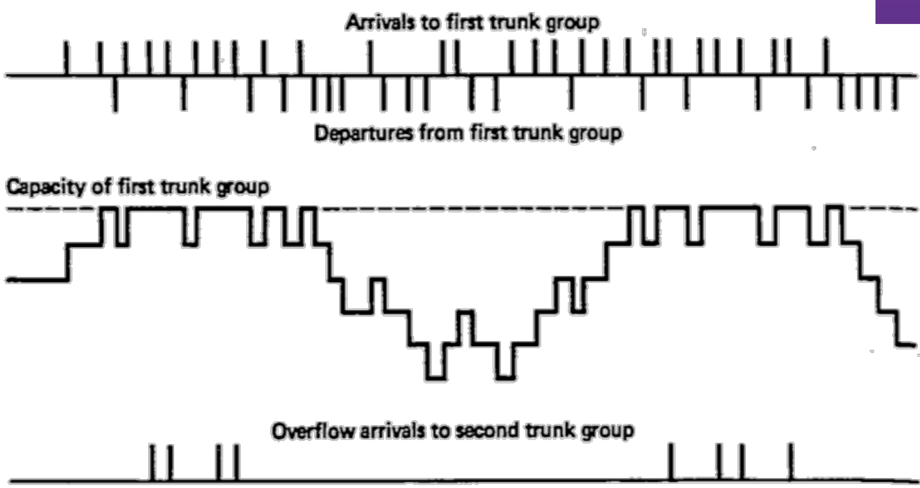
En nuestro caso las llamadas desbordadas suelen ir en bloques

Eso da mayores probabilidades de bloqueo que con un proceso de Poisson de igual media

Se aproxima con un proceso de Poisson de mayor tasa

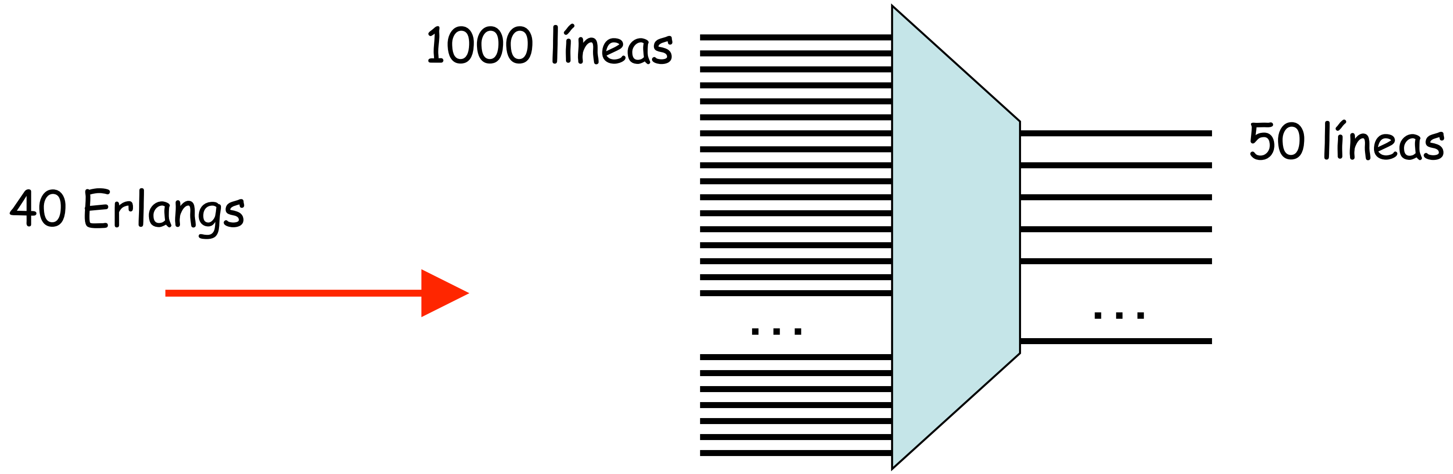
(En los problemas en caso de no disponer de las tablas emplearemos Poisson de igual tasa, aunque esto es subdimensionar)





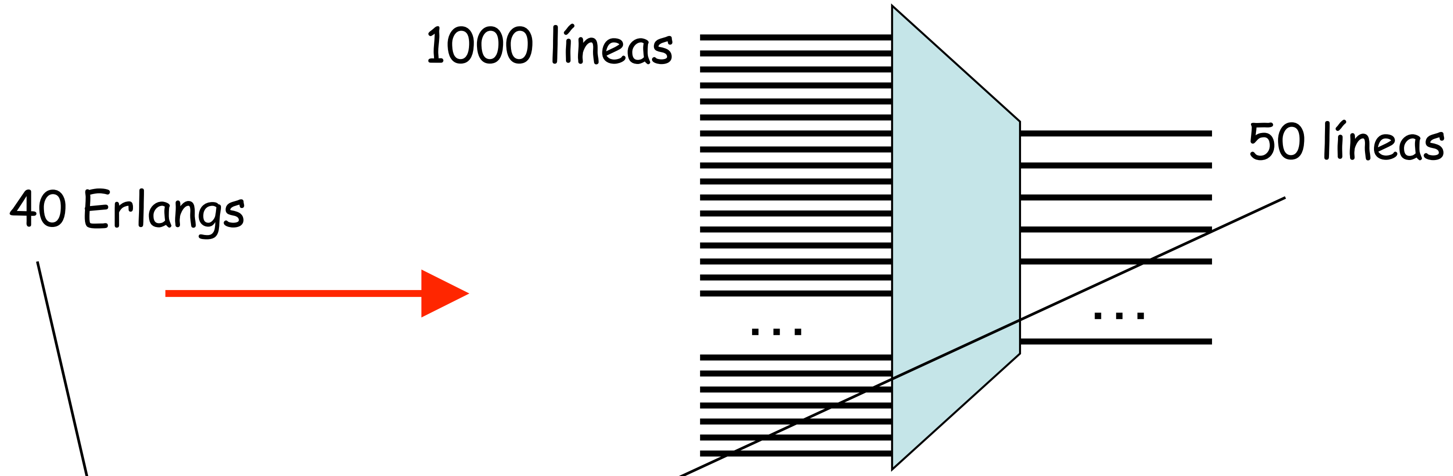
Ejemplos (1)

- 1000 líneas llegan a un concentrador que selecciona 50 para entrar a una centralita
- Los usuarios generan un tráfico de 40 Erlangs
- ¿ Cuál es la probabilidad de bloqueo ?



Ejemplos (1)

- 1000 líneas llegan a un concentrador que selecciona 50 para entrar a una centralita
- Los usuarios generan un tráfico de 40 Erlangs
- ¿ Cuál es la probabilidad de bloqueo ?

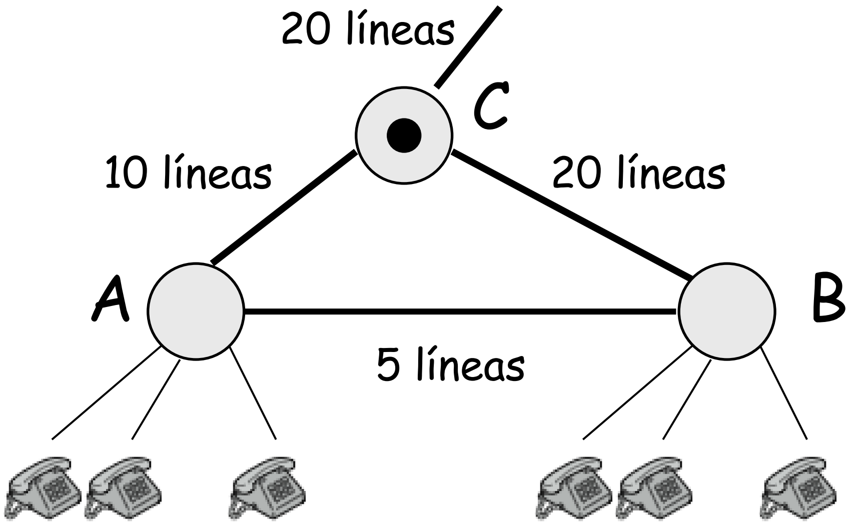


- La probabilidad de bloqueo es

$P_b = B(40, 50) = 0.0187$ casi un 2%

Ejemplos (2)

- En la centralita A de la figura las llamadas con destino a B se encaminan si es posible por el enlace directo a B y en caso de estar ocupado a través de la central primaria
- ¿ Cuál es el tráfico que cursa el enlace A-C y cuál es la probabilidad de bloqueo de una llamada de un abonado de A a uno de B ?

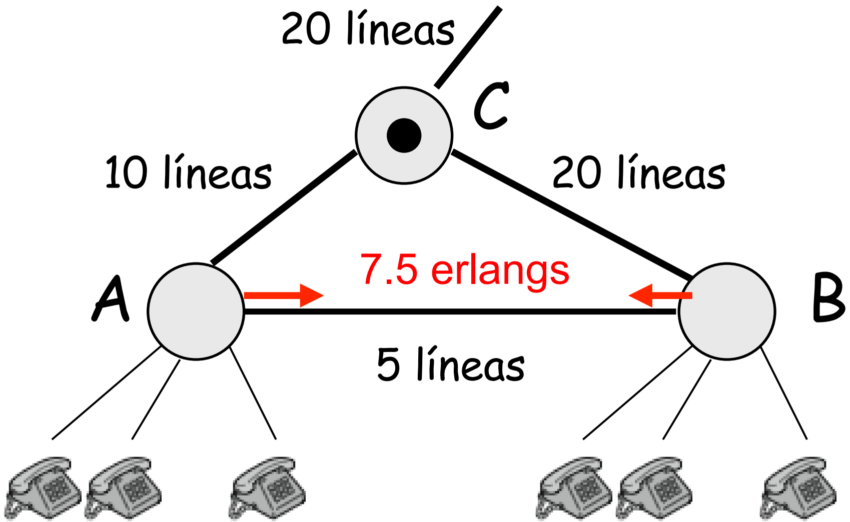


Demanda en Erlangs

Origen	a A	a B	Al exterior
De A	2	4.5	4.5
De B	3	3.2	5
Exterior	2	2	-

Ejemplos (2)

- Las 5 líneas entre A-B soportan un tráfico de 3+4.5=7.5 Erlangs
- Al ser 5 líneas la probabilidad de bloqueo es $p_1 = B(7.5,5) \approx 0.45$
  - Casi el la mitad de las llamadas no puede ir por la sección directa
  - Eso genera que un 45% del trafico que iba por ahí acabe yendo por C
  - Definimos: $q_1 = 1\text{-}p_1 = 0.55$

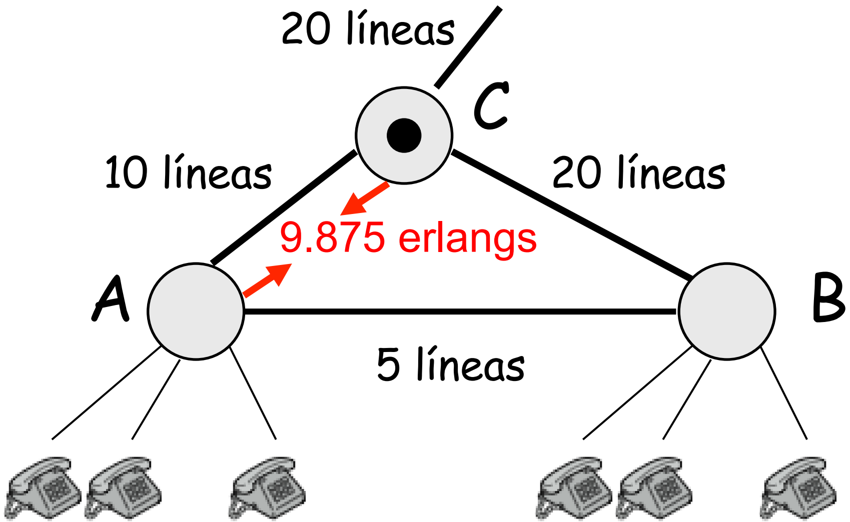


Demanda en Erlangs

Origen	a A	a B	Al exterior
De A	2	4.5	4.5
De B	3	3.2	5
Exterior	2	2	-

Ejemplos (2)

- El enlace entre A-C soporta un tráfico de:
  - Llamadas entre A y el exterior: 4.5 + 2 = 6.5 Erlangs
  - Llamadas entre A y B que no pueden ir directamente: 7.5 x 0.45 = 3.375 E
  - Total 9.875 Erlangs (aproximamos que es Poisson)
- 10 líneas con 9.875 Erlangs de demanda tienen una probabilidad de bloqueo de $p_2 = B(9.875,10) \approx 0.21$ (21%) ($q_2=1-p_2=0.79$)
- El enlace A-C tiene una probabilidad de bloqueo en torno al 21%

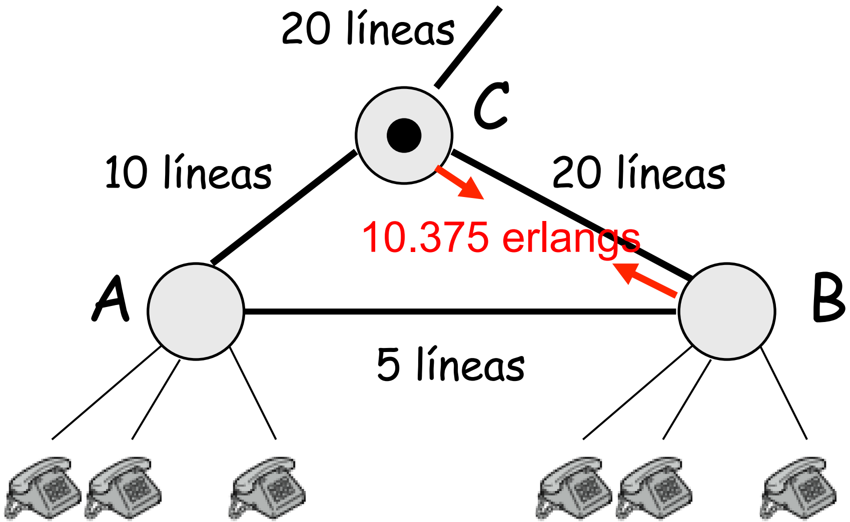


Demanda en Erlangs

Origen	a A	a B	Al exterior
De A	2	4.5	4.5
De B	3	3.2	5
Exterior	2	2	-

Ejemplos (2)

- El enlace B-C soporta un tráfico de:
  - Llamadas entre B y el exterior: 5 + 2 = 7 Erlangs
  - Llamadas entre A y B que no pueden ir directamente: 7.5 x 0.45 = 3.375 E
  - Total 10.375 Erlangs (aproximamos que es Poisson)
- 20 líneas con 10.375 Erlangs de demanda tienen una probabilidad de bloqueo de p_3 = B(10.375,20) ≈ 0.0027 (0.27%)
- Prácticamente despreciable ($q_3 = 1-p_3 \approx 1$ comparado con el resto)



Demanda en Erlangs

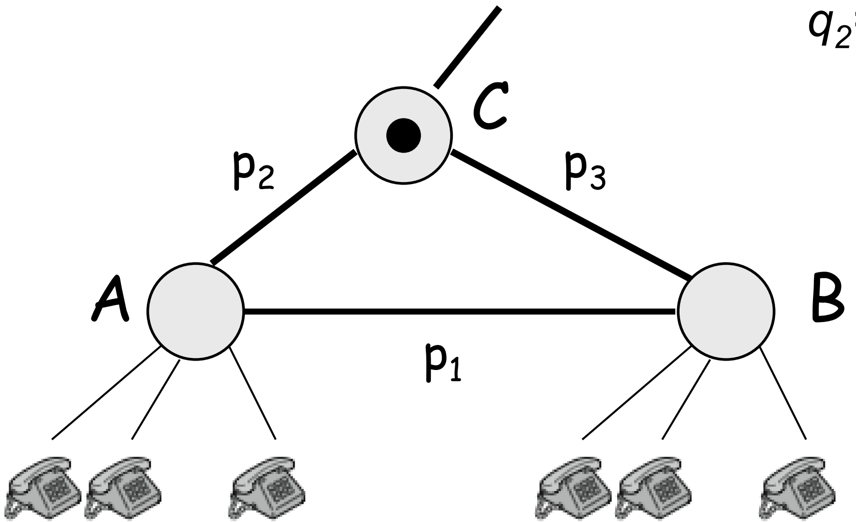
Origen	a A	a B	Al exterior
De A	2	4.5	4.5
De B	3	3.2	5
Exterior	2	2	-

Ejemplos (2)

- Probabilidades de bloqueo en cada enlace: p_1, p_2 y p_3
- Asumimos independencia
- Probabilidad de bloqueo de llamadas entre A y B: que ambos caminos se bloqueen (A-B y A-C-B)
- Probabilidad de que se bloquee el camino A-C-B = probabilidad de que se bloquee al menos uno de los dos (A-C y/o A-C-B) = 1 – probabilidad de que ninguno de los dos se bloquee

$$Pbloq_{A-B} = p_1(1-(1-p_2)(1-p_3)) = p_1(1-q_2q_3) \approx p_1p_2$$

$q_2=1-p_2$, $q_3=1-p_3\approx 1$

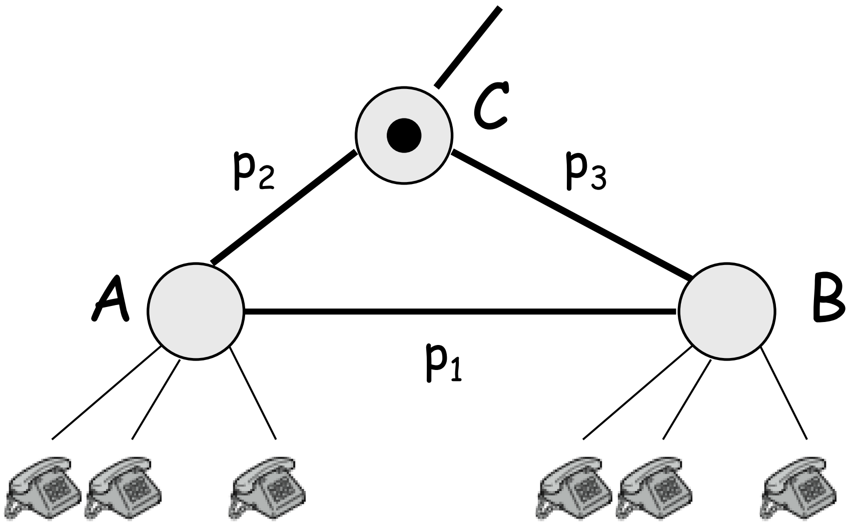


Demanda en Erlangs

Origen	a A	a B	Al exterior
De A	2	4.5	4.5
De B	3	3.2	5
Exterior	2	2	-

Ejemplos (2)

- Tráfico cursado por el enlace A-C:
  - Ofrecido a A-C-B (el desbordado de A-B) que es cursado: 3.375 x $q_2 q_3$
  - + tráfico de A con el exterior que es cursado: 6.5 x q_2
  - = 3.375 x (1-0.21)(1-0.0027) + 6.5 x (1-0.21) = 7.794 Erlangs

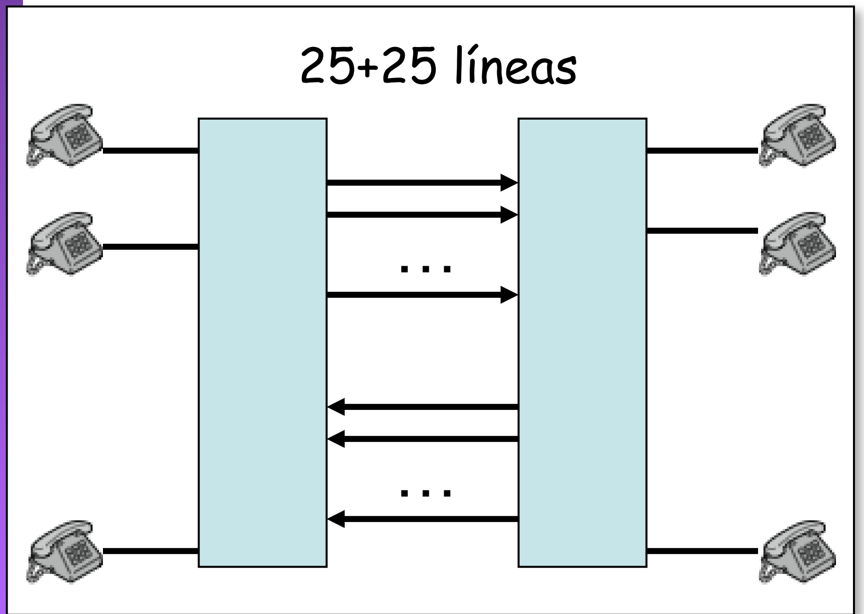


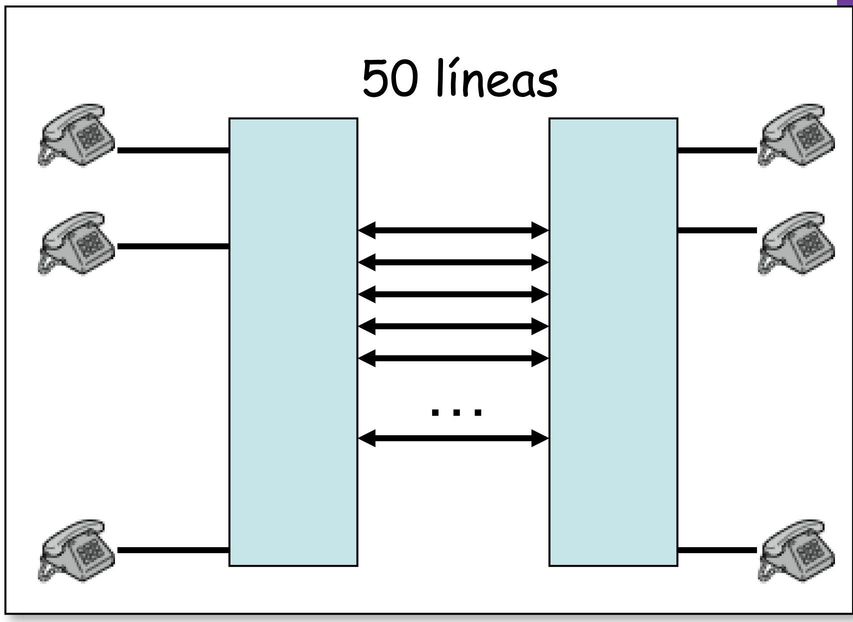
Demanda en Erlangs

Origen	a A	a B	Al exterior
De A	2	4.5	4.5
De B	3	3.2	5
Exterior	2	2	-

Ejemplos (3)

- Entre dos centralitas tenemos la posibilidad de:
  - asignar 25 troncales para llamadas salientes de A y 25 troncales para llamadas entrantes a A
  - O bien asignar las 50 troncales para que se puedan usar indistintamente en llamadas en cualquier dirección
- ¿ Qué es mejor ?





Ejemplos (3)

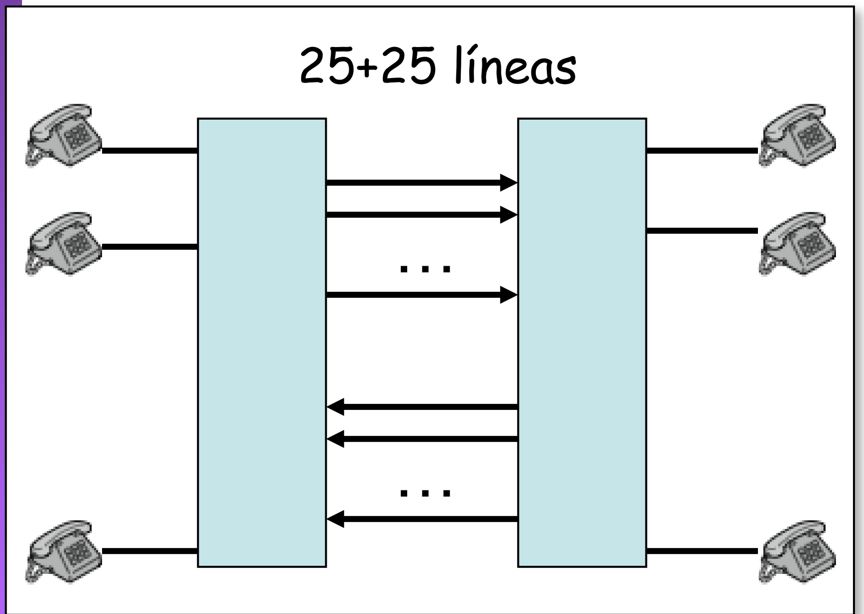
- Suponiendo que el tráfico que intenta ir de B a A es el mismo que el de A a B llamémosle I (pongamos 15 erlangs)
- Probabilidad de bloqueo en el caso 1:

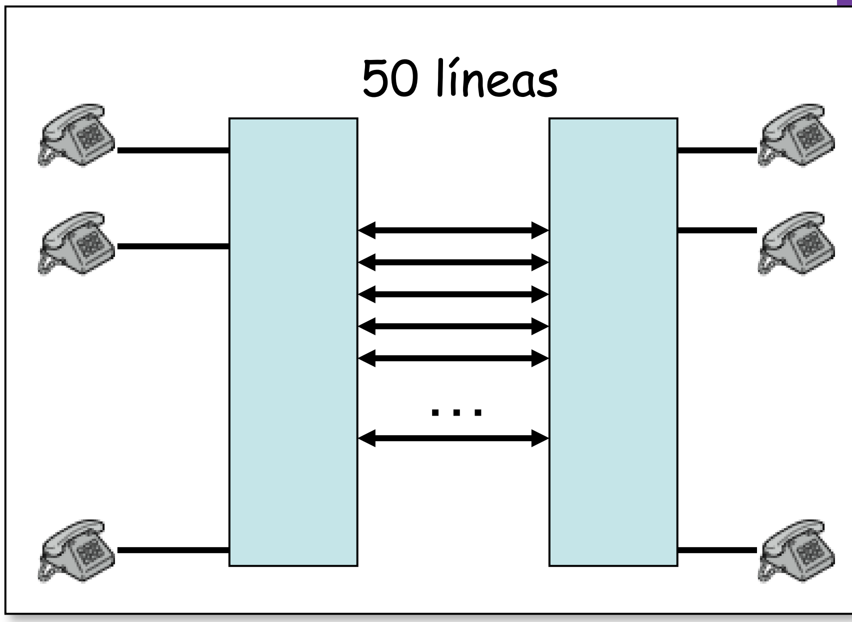
  $P_b(A\text{->}B)=B(I,25)$ $P_b(B\text{->}A)=B(I,25)$

  B(15,25)=0.005 0.5%
- Probabilidad de bloqueo en el caso 2:

  $P_b(\text{cualquier dirección})=B(I+I,50)$

  B(30,50)=0.0002 0.02% 20 veces menos !!!





Mayor complejidad

- *¿ Qué ocurre si las llamadas se retienen hasta que sean atendidas ?*

Teoría de colas (función C de Erlang)

- *¿ Qué ocurre si tenemos en cuenta que hay un número finito (y conocido) de usuarios ?*

Fórmula de Engset

Conclusiones

- El tráfico telefónico se modela mediante procesos de llegadas de Poisson y duraciones exponenciales
- La probabilidad de bloqueo se calcula mediante la B de Erlang
- Aproximaciones con tráfico de desbordamiento

Referencias

- Richard A.Thompson, “Telephone switching systems”, Ed. Artech House, capítulo 5
- John C. Bellamy, “Digital Telephony”, Ed. Wiley Interscience, último capítulo

Ingeniería de Teletráfico