

Introducción a la simulación de redes de comunicaciones

Carlos Ortega Marchamalo & Pablo Collado Soto

Conmutación

Universidad de Alcalá

Marzo
2020



Contenidos

1	Introducción	2
2	Descripción del <i>backbone</i> de datos	2
3	Retardo en el enlace Londres <-> Dublín	3
4	Retardo del trayecto Dublín -> Londres	4
5	Variando el valor de α	4
6	Aumentando el número de iteraciones	5
7	Bloqueo de llamadas	6
8	Ocupación del enlace Dublín --> Londres	7
9	Aumentando la velocidad de los enlaces	8
10	Conclusión	9

1 Introducción

A lo largo de la carrera nos hemos acostumbrado a analizar gran cantidad de escenarios de manera teórica aproximando sistemas reales por modelos matemáticos que si bien trataban de capturar la esencia de la realidad no pueden plasmar todos los detalles que la caracterizan. Con esta introducción en el mundo de la simulación pretendemos descubrir las sutilezas de esta herramienta y aprender a sacarle todo el provecho que podamos.

Para lograr los objetivos expuestos desarrollaremos una serie de experimentos para posteriormente analizar la validez de los resultados obtenidos comparándolos con los valores esperados de acuerdo a modelos teóricos. Observaremos que si bien existe un parecido también habrá diferencias que debemos ser capaces de explicar y aprender a tolerar ya que si bien el análisis basado en modelos matemáticos es rápido también puede pecar de una relativa simpleza a la hora de representar complejidades de sistemas existentes. Nos daremos cuenta de cómo la simulación puede ayudarnos a llenar este "vacío" y hasta qué punto podemos exigir unos resultados arbitrariamente fidedignos.

Nos aventuraremos también en el mundo de la estadística para explicar el por qué de algún que otro suceso en un intento de conferir una mayor rigurosidad a nuestro trabajo.

Comenzaremos por describir la red de datos de nuestro escenario para pasar a analizar valores característicos de la misma. Estudiaremos la influencia de varios parámetros en estas magnitudes y descubriremos cómo ser capaces de obtener valores de una mayor calidad a través de un incremento del número de iteraciones o del tiempo de ejecución de las distintas pruebas. Acompañamos el texto relativo a cada uno de estos puntos de una serie de conclusiones que nos permitirán esclarecer las razones detrás de los hechos observados.

Esperamos haber sabido condensar todo lo aprendido en las líneas necesarias para facilitar la comprensión del lector a la vez que intentamos evitar excedernos en nuestras explicaciones.

2 Descripción del *backbone* de datos

La mayoría de los ejercicios que debemos realizar se relacionan con enlaces de datos que conectan las centrales de **Dublín** y **Londres**. En nuestro escenario observamos cómo Dublín generará en un principio mensajes siguiendo un proceso de *Poisson* con una tasa $\lambda_D = \frac{1}{5} \frac{msg}{s}$ y dirigidos a **Londres**. Más adelante veremos como la tasa de mensajes generados en **Londres** (λ_L) resulta ser exactamente la misma teniendo todos ellos la capital irlandesa como destino. Ahora, el tamaño de estos mensajes **NO** es similar. Los mensajes generados en la central irlandesa tendrán un tamaño distribuido exponencialmente con una media de $1000 B$, esto es, $L_D \exp(-\frac{1}{1000 B}) \rightarrow E[L_D] = 1000 B$. En cambio las respuestas generadas por la central inglesa tendrán un tamaño fijo de $200 B$ lo que implica que $L_L \delta(x - 200 B)$. Veremos cómo esta diferencia tendrá implicaciones importantes en los resultados obtenidos.

Llamamos también la atención a que el único tráfico de la red es el anteriormente descrito con lo que todos estos paquetes transitarán por este enlace. Asimismo observamos que el enlace de interés (*PPL*) tiene una tasa binaria $R_{PPL} = 128 kbps$ y que las centrales se conectan a los routers en los extremos del mismo a través de los enlaces A y A' con una tasa binaria $R_A = R_{A'} = 64 kbps$ en ambos casos. Señalamos también que estos enlaces son *full-duplex* en el sentido de que podemos transmitir en ambos sentidos de manera simultánea. Esto nos permitirá modelar sendas direcciones de manera independiente lo que facilitará mucho el desarrollo y la aproximación teórica.

Tal y como adelantábamos en el primer párrafo el hecho de que no haya más tráfico que el que hemos recogido arriba supone que las tasas de envío de mensajes serán las mismas en **Dublín** y **Londres**, esto es, $\lambda_D = \lambda_L = \lambda = 0,2 \frac{msg}{s}$. Dado que solo existe una ruta por la que los enlaces pueden fluir nos aseguramos de que todos los paquetes generados en Irlanda llegarán a Reino Unido y viceversa. Dado que la central londinense se limita tan solo a responder los mensajes que le llegan (no "genera" tráfico de manera independiente) se puede comprobar que la igualdad anterior en efecto se mantiene.

Al ser el retardo que sufren los paquetes en el enlace una magnitud a estudiar veremos que esta comprende dos componentes. Por un lado debemos tener en cuenta el tiempo de transmisión del enlace que refleja la cantidad de datos que puede manejar éste de manera concurrente y es definido como $t_{tx} = \frac{L}{R} s$ para un paquete de $L bits$ y un enlace con una tasa de $R bps$. La segunda fuente de latencia viene dada por el propio tiempo que tarda en propagarse la información. Al fin y al cabo en el enlace tenemos señales de naturaleza eléctrica u óptica y tal como demostró *Einstein* nada viaja más rápido que la luz con lo que existirá lo que denominamos un retardo de propagación t_{prop} que dependerá principalmente de la distancia física del enlace. Nosotros hemos ignorado este parámetro por no aportar suficiente información como para asumir las implicaciones que supone tenerlo en cuenta pero siendo estrictos el retardo en estos enlaces vendría dado por $t_t = t_{tx} + t_{prop}$. Nos gustaría resaltar que a pesar de lo que parece implicar la expresión de t_{tx} este retardo no tiene por qué ser constante si alguno de

sus parámetros no lo es y tal y como comentábamos el tamaño de los paquetes de Dublín se modelan como una variable aleatoria que es intrínsecamente estadística...

Concluimos pues comentando que a la luz de los parámetros que definen nuestro escenario modelaremos el enlace en el sentido **Dublín -> Londres** a través de un sistema de colas $M/M/1$ y la conexión en el sentido opuesto como un $M/D/1$. Atendiendo a la notación de *Kendall* vemos como en ambos el tiempo entre llegadas se distribuye exponencialmente, solo se envía un paquete de manera concurrente y se asume una cola infinita en los routers implicados. La diferencia viene por la distribución de los tiempos de envío de estos paquetes. En el primer caso un tamaño exponencial supondrá un tiempo de servicio igualmente distribuido mientras que en el segundo un tamaño constante implicará un tiempo de servicio determinista. Somos conscientes de las limitaciones de estas premisas (los routers no cuentan con memoria infinita por ejemplo) pero dada la situación creemos que son buenas aproximaciones. A lo largo del desarrollo de los diversos experimentos veremos como ésto es efectivamente así. Además analizaremos por qué modelar el enlace de vuelta como una cola $M/D/1$ es incorrecto a pesar de lo que podríamos pensar.

3 Retardo en el enlace Londres <-> Dublín

Con el escenario claro nos disponemos a recoger los primeros datos. La simulación nos informa de que el retardo medio de tránsito en el sentido **Londres -> Dublín** es de 12,5 ms y, para nuestra sorpresa, el intervalo de confianza es 0, esto es, el retardo es **exactamente** el proporcionado. Si lo comparamos con el resultado de aplicar un modelo $M/D/1$:

$$E[T] = \frac{\frac{1}{\mu}}{1 - \rho} \cdot (1 - \frac{\rho}{2}); \quad \rho = \frac{\lambda}{\mu} = \frac{\lambda}{\frac{1}{T_s}} = \frac{0,2 \frac{msg}{s}}{\frac{128 \text{ kbps}}{200 \cdot 8 \text{ b}}} = 0,0025$$

Conociendo ya todos los datos necesarios llegamos a que $E[T] = 12,516 \text{ ms}$. No obstante los resultados teóricos difieren de los obtenidos lo que nos motiva a encontrar la razón de la discrepancia. Observando de nuevo la red nos percatamos de que los mensajes generados por **Londres** atraviesan un enlace de tasa $R_A = 64 \text{ kbps}$ para luego llegar al enlace estudiado de tasa $R_{PPL} = 128 \text{ kbps}$. Este hecho unido al tamaño determinista de los paquetes provoca una situación inusual. Si nos ponemos en la piel de uno de estos mensajes veremos que tardaremos $\frac{200 \cdot 8}{64 \text{ kbps}} = 25 \text{ ms}$ en atravesar el enlace entre routers. En cambio, en la segunda conexión tenemos una tasa $R_{PPL} = 128 \text{ kbps}$ lo que implica un retardo de transmisión de 12,5 ms. Observamos pues que **NO** se nos generará una cola en el enlace analizado pues nos da tiempo a transmitir los paquetes por éste antes de que nos llegue uno nuevo. En ausencia de cola no es correcto modelar la situación a través de un $M/D/1$, de ahí el error obtenido anteriormente.

Tal y como comentábamos el segundo enlace puede caracterizarse como un $M/M/1$ pero la situación anterior nos ha llevado a cerciorarnos de que no se nos escapa nada. Si bien es cierto que los mensajes siguen saliendo de un enlace lento a uno rápido ahora su tamaño no se distribuye de manera determinista sino exponencial. Ésto impide asumir la desaparición de la cola pues podemos tener ráfagas de paquetes muy grandes que bloqueen a paquetes de tamaño mucho más reducido. En cuanto entra en juego la estadística no podemos seguir haciendo predicciones y nos debemos poner en manos de modelos y simulaciones. Los valores teóricos obtenidos son:

$$E[T] = \frac{1}{\mu - \lambda}; \quad \mu = \frac{R}{E[L]} = \frac{128 \text{ kbps}}{1000 \cdot 8 \text{ bit}} = 16 \frac{pkt}{s} \rightarrow E[T] = 63,29 \text{ ms}$$

Mostrando los datos recopilados en una tabla junto a los teóricos observamos cómo las diferencias son razonables ya que existe una clara similitud:

[ms]	Media	L. Inferior	L. Superior	Int. Confianza	Media Teórica
Retardo <i>PPL</i> D->L	64,38	55,12	73,65	9,26	63,29
Retardo <i>PPL</i> L->D	12,5	12,5	12,5	0	12,5

Nótese que dado lo determinista del retardo en el enlace en el sentido **L -> D** no lo incluiremos en tablas posteriores pues siempre obtendremos los mismos valores para cada celda y creemos que dificulta el manejo e interpretación de los datos para el lector. Asimismo comentamos al final del siguiente apartado por qué estos datos no son válidos si nos atenemos a los criterios de calidad impuestos para nuestros resultados.

4 Retardo del trayecto Dublín -> Londres

Una vez que hemos discutido sobre el retardo que sufrirán los paquetes en ambos sentidos nos centramos ahora en el tiempo que transcurre desde que la central de Dublín emite un paquete hasta que éste se recibe en la central de Londres. El trayecto implica a tres enlaces diferentes y, siguiendo el mismo razonamiento que antes optaremos por modelarlos como si de un sistema $M/M/1$ se tratara. Tal y como podemos esperar al lidiar con paquetes cuya longitud se distribuye exponencialmente podremos hablar de retardos medios, nunca deterministas. Así, el retardo total se puede obtener como: $E[T_t] = E[T_1] + E[T_2] + E[T_3]$ siendo $E[T_x]$ el retardo medio de tránsito asociado al enlace x .

Antes de llevar a cabo la simulación podemos esperar encontrarnos con un error mayor que en los casos anteriores. Ésto se debe a que estamos asumiendo que el proceso de salida de los paquetes de un enlace sigue siendo *Poissoniano* si las llegadas lo son cosa que **NO** es necesariamente cierta... No obstante y en ausencia de un modelo $G/G/1$ o mejor dicho, $G/M/1$ que nos permita caracterizar de manera exacta una distribución de llegadas genérica nos vemos obligados a trabajar con la aproximación más correcta aunque sepamos que ésta no es cierta. Asumiremos que los paquetes que son emitidos por un enlace para llegar al siguiente lo harán de acuerdo con un proceso de *Poisson* de igual tasa con la que llegaron a este primer enlace. Ya habíamos hecho esta suposición en el apartado anterior pero tan sólo en el paso del enlace A al enlace PPL puesto que estábamos interesados en los retardos en este último. Ahora sin embargo también se verá involucrado el enlace A' con lo que tendremos que tolerar este error una segunda vez.

Con todo aclarado pasemos a calcular el retardo en los enlaces A y A' , ambos con una tasa binaria $R_A = 64 \text{ kbps}$. Modelándolos como sistemas $M/M/1$:

$$E[T_1] = E[T_3] = E[T_A] = E[T_{A'}] = \frac{1}{\mu - \lambda} = \frac{1}{\frac{64 \text{ kbps}}{1000 \cdot 8 \text{ bit}} - 0,2 \frac{\text{msg}}{\text{s}}} = \frac{1}{8 - 0,2} = 128,205 \text{ ms}$$

Sumando todo:

$$E[T_t] = \sum_{i=1}^{i=3} E[T_i] = E[T_1] + E[T_2] + E[T_3] = 2 \cdot 128,205 \text{ ms} + 63,29 \text{ ms} = 319,7 \text{ ms}$$

Comparándolo con la media obtenida en la simulación, $356,914 \text{ ms}$, observamos que la diferencia es mayor que en el primer caso. Debemos destacar que éstos datos tienen su origen en el mismo experimento que los anteriores de manera que al compararlos no tenemos que contar con el "ruido" que supone la aleatoriedad intrínseca al escenario. También observamos que el valor, si bien no está tan ajustado a la media como antes, pertenece al entorno de centro la media y radio el intervalo de confianza sin duda alguna pues el mínimo se encuentra en $298,702 \text{ ms}$. De estos dos resultados se desgana que el intervalo de confianza es de $58,21 \text{ ms}$.

A lo largo de este primer ejercicio hemos introducido el escenario sobre el que experimentaremos a la vez que hemos comprobado que los modelos matemáticos si bien no capturan la totalidad de las características que definen el sistema, nos proporcionan estimaciones muy ajustadas aún cuando somos conscientes de estar cometiendo un error como en el segundo caso. Adjuntamos a continuación una tabla que contiene las medidas obtenidas y que permite una comparación más sencilla:

$[ms]$	Media	L. Inferior	L. Superior	Int. Confianza	Media Teórica
Retardo msgs D->L	356,914	298,702	415,127	58,21	319,7

Comprobamos ahora si las medidas recogidas tanto aquí como en el apartado anterior son de una calidad aceptable. La condición que debemos comprobar es que $\delta < 0,1 \cdot E[X]$ donde X es el parámetro estudiado. Vemos que, al contrario de lo que podríamos esperar ésto **NO** se cumple. Más adelante descubriremos el por qué de este resultado y cómo podemos intentar subsanarlo. Adjuntamos una comprobación "fallida" a modo de demostración:

$$9,26 \not< 0,1 \cdot 64,38 = 6,438$$

5 Variando el valor de α

Para comentar el impacto de llevar a cabo este cambio debemos recuperar la definición misma de un intervalo de confianza δ :

$$Sea \bar{x} = \frac{\sum_{i=1}^N x_i}{N}, m \rightarrow P(\bar{x} - \delta \leq m \leq \bar{x} + \delta) = 1 - \alpha$$

Esto es, la media real m está contenida en el intervalo $[\bar{x} - \delta, \bar{x} + \delta]$ con probabilidad $1 - \alpha$. A pesar de que ahora estemos modificando α de manera directa no debemos pensar en este parámetro como algo variable sino como en un valor impuesto. Cuanto mayor sea α asistiremos a un intervalo de confianza cada vez más permisivo pues la probabilidad de que m pertenezca a este intervalo será cada vez menor. En consecuencia podremos dar unos intervalos de confianza mucho más ajustados (esto es, δ decrecerá) pero no por ello más útiles ya que la información con la que trabajamos es en el fondo la misma al seguir efectuando tan solo 10 iteraciones del experimento. Si tenemos un intervalo de confianza tremendamente pequeño veremos que la probabilidad de que el valor real pertenezca al mismo es muy reducida con lo que no estamos sacando nada en claro. ¿Qué más nos da que el intervalo de confianza sea tan preciso si es muy probable que no contenga el valor buscado?

Así, se acepta un valor de $\alpha \leq 0,1$ como válido y éste es el que hemos empleado al obtener todos nuestros resultados. Vemos pues que si llevamos a cabo la operación contraria, esto es, reducimos α , el intervalo de confianza tenderá a crecer (δ se hará mayor). Esta interacción entre α y δ nos da la sensación intuitiva de que ambas están tensando una cuerda por extremos diferentes en el sentido de que si tenemos una probabilidad muy alta de encontrar un valor en un intervalo éste debe ser muy extenso mientras que si la probabilidad se reduce podemos hacer lo mismo con el radio de este entorno. ¿Cuál es la mejor opción? Dependerá de la validez que queramos conferir a nuestros datos así como de las exigencias que tengamos que cumplir y la información que queramos extraer. Finalmente nos gustaría señalar que dada la naturaleza de las variables aleatorias que estamos estudiando estos intervalos de confianza serán simétricos, esto es, serán entornos de centro \bar{x} y radio δ .

Adjuntamos pues los resultados obtenidos al variar el valor de α . La probabilidad de que el valor real esté en el intervalo buscado será de $P(m \in [\bar{x} - \delta, \bar{x} + \delta]) = 1 - 0,3 = 0,7$:

$\alpha = 0,3 [ms]$	Media	L. Inferior	L. Superior	Int. Confianza	Media Teórica
Retardo PPL D->L	64,38	58,82	69,94	5,56	63,29
Retardo msgs D->L	356,914	321,992	391,838	34,923	319,7

Tal y como comentábamos se ha reducido el intervalo de confianza. Si tomamos por ejemplo la primera fila:

$$\Delta\delta = \delta|_{\alpha=0,3} - \delta|_{\alpha=0,1} = 5,56 - 9,26 = -3,7; \Delta\alpha = 0,3 - 0,1 = 0,2$$

Nótese que en el caso del enlace cuyo retardo de transmisión era determinista no observamos diferencia alguna pues $\delta_{min} = 0$ que es lo que ya teníamos anteriormente. Es decir, en caso de estudiar un fenómeno totalmente determinista observamos cómo variar α no tiene efecto alguno pues la medida no tiene un comportamiento determinista dada su propia naturaleza.

Concluimos pues que variar α no es algo que debamos hacer en busca de intervalos de confianza más ajustados sino un ejercicio que nos permite tener una mayor intuición sobre su significado e importancia.

6 Aumentando el número de iteraciones

Si no podemos variar α debe existir otra manera de acotar nuestro valor dentro de un intervalo de confianza más restrictivo sin reducir probabilidad de que el valor buscado esté contenido en el mismo. Para entender cómo lograrlo debemos recuperar la expresión del estimador de la media de un parámetro x :

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$$

La estadística nos asegura que la muestra con la que trabajamos es una muestra aleatoria simple pues las variables aleatorias que representan los valores estudiados están igualmente distribuidas y son independientes en cada iteración. Nos centraremos en analizar muestras exponencialmente distribuidas pues es la que sigue el tamaño de los paquetes generados en Dublín. Así, la distribución que sigue la superposición de varias variables aleatorias exponencialmente distribuidas es una $n - Earlang$ siendo n el número de variables sumadas. En otras palabras:

$$Sea X_i \sim exp(\lambda) \rightarrow \sum_{i=1}^N X_i \sim N - Erlang(\lambda)$$

De la función de densidad de una distribución $n - Erlang(\lambda)$ se desgrana que la esperanza viene dada por $\frac{n}{\lambda}$ y que la varianza se obtiene como $\frac{n}{\lambda^2}$. Sabemos además que si $X \sim n - Erlang(\lambda) \rightarrow \alpha \cdot X \sim n - Erlang(\frac{\lambda}{\alpha})$. Por tanto vemos cómo la superposición de varias distribuciones exponenciales acaba por suavizar los datos y otorgarnos la misma media que una sola distribución exponencial con una varianza mucho menor. Para demostrarlo vemos que:

$$Sea Y \sim exp(\lambda) \rightarrow E[Y] = \frac{1}{\lambda}$$

Si sumáramos varias Y_i :

$$Y_T = \sum_{i=1}^N Y_i \sim N - Erlang(\lambda) \rightarrow E[Y_T] = \frac{N}{\lambda}; V[Y_T] = E[(Y_T - E[Y_T])^2] = \frac{N}{\lambda^2}$$

Aplicando la propiedad de la multiplicación de una constante por una variable aleatoria que sigue una $n - Erlang(\lambda)$:

$$Y_T = \frac{1}{N} \cdot Y_T \rightarrow Y_T \sim N - Erlang(N \cdot \lambda) \rightarrow E[Y_T] = \frac{1}{\lambda} = E[Y]; V[Y_T] = \frac{1}{N \cdot \lambda^2}$$

Esto implica que la media de esta superposición es la de la variable aleatoria estudiada y que a medida que el número de muestras aumente la varianza bajará con lo que nos acercaremos constantemente a la media real. En otras palabras:

$$\lim_{N \rightarrow \infty} V[Y_T] = 0 \rightarrow \lim_{N \rightarrow \infty} \frac{Y_T}{N} = \frac{1}{\lambda}$$

No hemos hecho más que confirmar de manera simbólica la intuición que hemos ido afianzando durante todas nuestras experiencias con procesos aleatorios: a mayor número de repeticiones más exactos serán los datos que recojamos y más se acercarán estos a la realidad. En nuestro caso esto se traducirá en que a mayor número de repeticiones más pequeño será nuestro intervalo de confianza pues tendremos información como para hacer unas estimaciones más correctas y certeras manteniendo la misma probabilidad de "acertar". Para confirmar nuestras "sospechas" mostraremos los datos obtenidos. No olvidemos que de nuevo estamos trabajando con $\alpha = 0,1$ al efectuar las simulaciones:

20 iteraciones [ms]	Media	L. Inferior	L. Superior	Int. Confianza	Media Teórica
Retardo PPL D->L	57,41	51,60	62,50	5,45	63,29
Retardo msgs D->L	311,312	277,793	344,831	33,519	319,7

Fijémonos en que ahora el primer valor cumple la condición de que la media sea menor que el 10% del intervalo de confianza pero no es así para el segundo. Incrementando las iteraciones a 40 podemos poner remedio a esta última pega:

40 iteraciones [ms]	Media	L. Inferior	L. superior	Int. Confianza	Media Teórica
Retardo PPL D->L	57,41	53,59	61,22	3,81	63,29
Retardo msgs D->L	307,604	286,493	328,715	21,111	319,7

Tal y como esperábamos obtenemos unos intervalos de confianza aún menores. Además hemos logrado que todos los valores sean "correctos" en los términos que hemos discutido anteriormente. Concluimos pues que incrementar el número de iteraciones supone obtener un intervalo de confianza menor sin tener por ello que renunciar a una gran probabilidad de que los datos reales pertenezcan al mismo. Los datos obtenidos van convergiendo hacia la media real. Otra medida que podríamos haber tomado es incrementar el tiempo de ejecución de los distintos experimentos para al final acabar incrementando el número de muestras en aras de lograr el mismo objetivo. Analizaremos esta opción en los siguientes apartados.

7 Bloqueo de llamadas

Para poder llevar a cabo un análisis más preciso y que se ajuste lo máximo posible a la realidad existen dos alternativas. Por un lado y tal y como hemos venido realizando en los casos anteriores, podemos aumentar el número de repeticiones realizadas sobre la simulación. En el caso que ahora nos atañe vamos a hacer uso de la

otra vía de la que disponemos para lograr el fin deseado. Así, aumentando el tiempo de duración de la repetición logramos recopilar un número mayor de datos lo que nos permite establecer una relación con lo que sucede realmente.

En el escenario del que disponemos (nótese que ahora aludimos a la infraestructura telefónica), pese a existir diversos posibles destinos de las llamadas realizadas desde **Dublín** pues esta sede posee conexión tanto con el nodo localizado en **Madrid** así como con el localizado en Londres, llevando a cabo un estudio minucioso de la información obtenida navegando por los distintos menús disponibles en el punto donde se generan las llamadas podemos cerciorarnos con total seguridad de que estas únicamente poseen como destino **Madrid** pues así resulta explicitado en uno de los parámetros de configuración observados.

El modelo empleado para poder llevar a cabo un estudio más sencillo del comportamiento de todo el esquema se corresponde con el de un sistema $M/M/N/N$, tomando como valor para el número de líneas o canales dedicados el más limitante, en nuestro caso 14, pues este es el valor que antes imposibilita el transcurso de nuevas llamadas. Aunque la central que en primer lugar recibe las llamadas dispone de capacidad para cursar un número mayor, si no existen recursos desde ella hacia el exterior las solicitudes no pueden transcurrir, lo que explica la selección del valor que establece el tope en la capacidad del sistema en conjunto.

En referencia a estas llamadas generadas en **Dublín** y con destino **Madrid**, el análisis y el estudio del fichero *reports* generado de forma automática nos proporciona un apartado específico referente a las probabilidades de bloqueo que sufren las llamadas que tiene lugar en el sistema y que son objeto de nuestra investigación.

De este modo, poniendo el foco de atención en las que en concreto tienen su origen en **Dublín** nos percatamos de que presentan una probabilidad de bloqueo en tanto por 1 de 0,132, un valor en cierto modo coherente y concordante con los cálculos matemáticos llevados a cabo en función del sistema planteado en primera instancia, es decir, la probabilidad de bloqueo para un sistema $M/M/N/N$ viene dada por $E_1(N, A)$ siendo $N = 14$ líneas y presentando las llamadas un tiempo de generación ($\frac{1}{\lambda}$) de 30 s así como un tiempo de duración de 3 min, ambos parámetros siguiendo un modelo de distribución exponencial y resultando el tráfico generado en cada sede igual a $\frac{\lambda}{\mu}$, es decir, 6 *Erlangs*, con lo que en total se tiene un tráfico de $A = 12$ *Erlangs* pues debemos considerar los dos sentidos ya que ambos transcurren por las mismas 14 líneas.

Haciendo uso de la calculadora de tráfico de la que disponemos e introduciendo los valores citados anteriormente correspondientes al tráfico total así como el número de líneas disponibles obtenemos un valor bastante próximo al observado mediante los ficheros *reports* generados con el simulador que empleamos. En nuestro caso, obtenemos una probabilidad de bloqueo teórica de 0.117, un valor relativamente cercano al anterior y que difiere de forma pequeña en las imperfecciones presentadas por el modelo empleado para tratar de imitar el funcionamiento de todo el conjunto en global.

En conclusión, nos percatamos de que se está produciendo un bloqueo de un porcentaje importante de las llamadas que se originan en la sede de **Dublín** con destino **Madrid**. Se trata de algo más del 10% más específicamente, un valor un tanto elevado para lo que sería un valor apropiado y deseado en la realidad, en la que se busca reducir al mínimo esta probabilidad, siendo prácticamente de obligación situarla como máximo en una cifra porcentual, es decir, un valor inferior a la decena y siempre lo más cercano a 0.

8 Ocupación del enlace Dublín --> Londres

El programa en el cual nos apoyamos para llevar a cabo un análisis más detenido del transcurso del tráfico tanto de paquetes como de llamadas entre varios nodos, COMNET III, nos permite ir un poco más allá en este estudio y realizar incluso simulaciones sucesivas que difieran en un parámetro en concreto, el cual sufre unas variaciones explicitadas de antemano, lo que posibilita conocer la influencia y trascendencia del mismo en el funcionamiento global de todo el sistema que nos ocupa.

De este modo, en nuestro caso hemos decidido realizar una serie de simulaciones centrándonos en la trascendencia que presenta el tiempo entre llegadas de mensajes, es decir, $\frac{1}{\lambda}$, en cada uno de los nodos que son capaces de generar esta información, es decir, tanto en las sedes de **Dublín** como de **Madrid**.

En todas las variantes del experimento a realizar el parámetro de interés sigue una distribución exponencial, la cual se encuentra caracterizada en cada simulación por uno de los valores incluidos en la lista especificada al llevar a cabo la configuración previa del escenario correspondiente a través de los distintos menús disponibles. Estos datos en concreto son 0,5, 1, 2, 4, 6, 8 y 10, respectivamente, para cada uno de los 7 experimentos de los que consta la cuestión a estudiar.

Llevando a cabo esta configuración podemos analizar cómo responde nuestro esquema ante los diferentes escenarios que tienen lugar, percibiendo qué sucede en los enlaces cuando los paquetes se generan con una frecuencia mayor o cómo a medida que este ritmo de generación disminuye la congestión así también lo hace. De forma análoga ocurre con el retardo que sufren los paquetes desde que son generados hasta que llegan al destino, circunstancia que también podremos apreciar gracias a las particularidades del caso que abordamos.

Tras poner en marcha la prueba y esperar a la conclusión de la misma, nos es posible obtener diversos documentos de estadísticas así como ficheros *reports*, resultando más precisos los primeros y ajustándose en mayor medida a los valores obtenidos tras llevar a cabo el correspondiente análisis matemático, el cual, mediante el parámetro $\rho = \frac{\lambda}{\mu}$ nos informa del factor de utilización u ocupación teórico que debería presentar el canal por el que transcurre todo el tráfico.

Al establecer el punto de mira en los valores obtenidos y plasmados en el documento de estadística apreciamos que la variación en el parámetro $\frac{1}{\lambda}$, es decir, el tiempo entre llegadas de mensajes, influye de manera inversamente proporcional en el factor de ocupación del canal. Disminuyendo la frecuencia con la que se generan los paquetes, o lo que es lo mismo, aumentando el tiempo entre la llegada de cada paquete, logramos que cada vez vaya disminuyendo más el número de paquetes lo que descongestiona el canal por el que circulan los mismos. De forma inversa sucedería de manera similar, por lo que una reducción del tiempo entre cada mensaje aumentaría la ocupación del medio transmisor.

En nuestro caso, con el transcurso de los experimentos se produce una disminución en el porcentaje de utilización del canal. Este cambio se produce de forma similar en ambos sentidos.

Pero la importancia de la fluctuación del parámetro explicitado no se queda aquí. Como era de esperar, su influencia va más allá y presenta su trascendencia en cuanto al tiempo que tardan los paquetes en llegar desde su origen hasta su destino. Esto tiene su lógica pues cuanto mayor sea la congestión que presente el canal, más tiempo debe esperar el paquete para viajar a través de él.

Este hecho puede apreciarse en cierto modo y de forma un tanto pequeña al observar los tiempos transcurridos en el sentido desde **Dublín** hasta **Londres** mediante las estadísticas disponibles. Sin embargo, la disminución apreciada resulta en cierta medida más bien mínima, siendo esta del orden de algún milisegundo o, de forma equivalente, algunos cientos de microsegundos.

En cambio, no sucede de forma similar en el sentido contrario, es decir, desde **Londres** hasta **Dublín**. Esto es debido, tal y como hemos comentado en cuestiones anteriores, a que los paquetes que viajan en esta dirección presentan un tamaño constante y, además, pasan de un enlace más lento a uno más rápido por lo que en el segundo no se forman colas y los tiempos permanecen constantes y invariantes.

A través de este caso hemos podido constatar la trascendencia que presenta así como la influencia de la variación de un parámetro del sistema global como es el caso del tiempo entre llegadas de mensajes para el grado de ocupación o utilización del canal por el que transcurren los paquetes y su consiguiente relevancia en los retardos sufridos por estos últimos al viajar por el medio transmisor.

9 Aumentando la velocidad de los enlaces

Llegado a este punto, procedemos a incidir un ápice más en la situación que se nos había presentado en un primer momento. En ella se tiene un sistema caracterizado por un parámetro $\alpha = 0,1$ el cual define los intervalos de confianza sobre cada una de las mediciones en torno al modelo llevadas a cabo.

Para poder extraer una conclusión más precisa y que se asimile lo máximo posible a la realidad con anterioridad decidimos llevar a cabo una simulación incrementando el número de repeticiones que se iban a producir, llevando esta cifra hasta el valor de 40 lo que nos permitía obtener un intervalo de confianza más reducido y acotado.

Como comentamos en primera instancia en el momento de realizar un análisis un poco más profundo del esquema del que disponemos, el enlace existente entre Londres y Dublín puede asimilarse a un modelo de colas $M/D/1$ por cuanto los mensajes que son generados en el origen presente una longitud fija y constante que no sufre variaciones en ningún momento por lo que el resultado que produce es siempre el mismo, resultando también fijo el tiempo que tarda en atravesar un paquete el enlace establecido entre los routers de las dos localizaciones citadas previamente. Por su parte, en el sentido contrario el escenario resulta un tanto diferente ya que se corresponde con un sistema $M/M/1$ pues la longitud de los paquetes sigue una distribución exponencial, lo que se traduce en cambios en la misma y, de forma consecuente, en los tiempos que se pueden medir respecto a la disposición existe.

En este último apartado de estudio la variación presentada respecto a lo presentado en un primer momento reside en la modificación de la capacidad que poseen los enlaces que unen los routers presentes entre si. Este valor de la tasa binaria pasa a incrementarse hasta los 256 *kbps*, hecho que tendrá su trascendencia y que procedemos a comprobar y a analizar.

Si tomamos el documento de estadísticas que se genera después de la ejecución del programa y tras llevar a cabo el consiguiente filtrado de la múltiple información que presenta, nos centramos en el retardo que sufren los mensajes en el enlace situado entre los routers de **Dublín** y **Londres** así como en el retardo de los paquetes entregados por la fuente de mensajes de **Dublín**.

Analizando estos datos, su concordancia en cuanto a los cálculos matemáticos también resulta correcta, de forma similar a lo que sucedía en las primeras cuestiones que abordamos en el documento que nos ocupa.

Llevando a cabo una comparación entre los resultados obtenidos en este último caso con los que se obtuvieron en las simulaciones iniciales podemos apreciar ciertas variaciones, tal y como era de esperar. Esta suposición previa es debida a que ambas situaciones difieren en la tasa binaria del enlace, con lo que nos encontramos con algo que preveíamos.

Profundizando más en los valores recolectamos percibimos que al aumentar el parámetro que varía conseguimos reducir los tiempos resultantes. Las mediciones referentes al enlace entre los router de **Dublín** y **Londres** en ambos sentidos se ven disminuidos a la mitad, algo en cierto modo de esperar por cuanto la capacidad del medio se ha visto incrementada también duplicándose. Respecto al tiempo retardo de los paquetes entregados por la fuente de mensajes de **Dublín** esta medición sufre una disminución de forma similar, aunque esta no resulta tan abultada ya que no solo depende del enlace que ha aumentado su régimen binario si no que implica a otros enlaces más que son trascendentes al ser más lentos, lo que marca el resultado final.

Respecto a los intervalos de confianza, estos se ven más concretados y reducidos, estableciendo unos valores más acotados y precisos.

De este modo, podemos observar la trascendencia de modificar un valor y poniendo el foco en un punto en concreto, donde podemos apreciar cambios significativos, pero que, al generalizar para todo el esquema, no tienen una influencia tan grande por cuanto son más los elementos involucrados en el cómputo global.

A modo de resumen para este último caso, incluimos una tabla para poder visualizar más explícitamente esta variaciones comparando ambos casos. Los cálculos matemáticos realizados para corroborar lo visualizado a través de la simulación del escenario definitivo son exactamente iguales al precedente anterior con la excepción de la variación de la tasa binaria.

40 iteraciones [ms]	Media	L. Inferior	L. Superior	Int. Confianza	Media Teórica	Media Ej. 3
Retardo <i>PPL</i> D->L	28,7	26,8	30,61	1,91	31,45	57,41
Retardo msgs D->L	277,082	258,473	295,69	18,608	313,82	356,914

10 Conclusión

La realización de la práctica que nos ocupa nos he permitido profundizar mucho más en diversos aspectos de las redes de comunicación tanto de datos como de voz de los cuales ya éramos conocedores pues los habíamos estudiado en diversas asignaturas a lo largo de nuestro grado universitario, pero sobre los que, realmente, no llegábamos a asimilar su verdadera influencia e importancia en todo el entramado que constituyen las mencionadas redes.

Así, hemos podido experimentar con la modificación de parámetros que caracterizan tanto los paquetes que viajan a nuestro alrededor como las llamadas que se producen entre la población, además de los enlaces por los cuales transcurre toda esta información.

Existen diversas circunstancias muy importante para poder relacionar con mayor precisión las simulaciones llevadas a cabo con la estricta realidad. Entre ellas sobresalen dos, las cuales han sido puestas en práctica con instancia a lo largos de los diferentes escenarios planteados relacionados con el modelo de red proporcionado. La primera de ella se basa en aumentar el número de repeticiones que deben realizarse sobre el escenario conformado. De forma similar, incrementar el tiempo de duración de estas repeticiones también es un acontecimiento que nos ayuda a un lograr un mejor análisis. La razón de todo ello es que se consigue recolectar un mayor número de muestras y datos lo cual posibilita acotar muchos más las mediciones y reducir los intervalos de confianza de las mismas, aspecto de suma importancia de igual modo.

Este experimento implementado nos ha posibilitado, además, comprender la relación existente entre la realidad y los modelos que se emplean para conectar esta con las puras matemáticas que facilitan todo el estudio numérico existente. Ello nos ha conducido a apreciar los mínimos pero existentes errores que surgen al establecer esta relación. Del mismo modo, la asociación con un esquema de colas nos permite calcular de forma teórica algunas mediciones para compararlas posteriormente con la propia simulación, apreciando además cómo influyen los tipos de distribución estadística en la caracterización de los parámetros implicados, viendo como en algunos casos, por extraño que parezca, los tiempos permanecen constantes, mientras que en otros casos el aumento o disminución de ciertos valores influye muy significativamente en los resultados finales. Hemos podido cerciorarnos de las limitaciones existentes en función de las capacidades de los medios a través de los cuales, al fin y al cabo, viaja toda la información intercambiada y cómo estas influyen en las mediciones e incluso hasta en el consiguiente bloqueo de llamadas.

Por tanto, la práctica abordada resulta de gran utilidad y curiosidad para comprender un poco más todo lo que sucede en las redes de comunicación existentes a nuestro alrededor en cualquier lugar y de las cuales hacemos uso prácticamente en cualquier momento, incluso sin llegar a ser conscientes de ellos y de la importancia y trascendencia de todos y cada uno de sus componentes.