
7 Comportamiento de TCP en entornos móviles: Efecto de los errores

7.1 INTRODUCCIÓN

Los servicios de transmisión de datos en redes móviles constituyen un sector de las telecomunicaciones que está experimentando un auge importante. La demanda potencial de servicios móviles de datos es tan grande que está obligando a los operadores a suministrar este tipo de comunicaciones. Las razones principales por las que los usuarios están comenzando a solicitarlos cada vez más, son:

- Por un lado, la reducción en el tamaño y precio de los ordenadores portátiles ha extendido su uso de tal manera que ya no se conciben las redes sino integrando estos equipos junto con los ordenadores fijos convencionales. La aparición, además, de una interfaz estándar entre el ordenador y el canal radio, la tarjeta PCMCIA, ha sido determinante en el desarrollo de módems compactos de muy reducido tamaño que facilitan el intercambio de datos en entornos móviles.
- En segundo lugar, la utilización de aplicaciones como el correo electrónico y, en general, de servicios de acceso a Internet, está llevando a los usuarios de portátiles a solicitarlos a través del enlace móvil, siendo clave para el desarrollo de los servicios de comunicación de datos a través del canal radio móvil.
- Además, los esfuerzos en el desarrollo de nuevas soluciones WAP (*Wireless Application Protocol*) [<http://www.wapforum.com>], para acceso a Internet desde el terminal móvil están siendo una realidad.

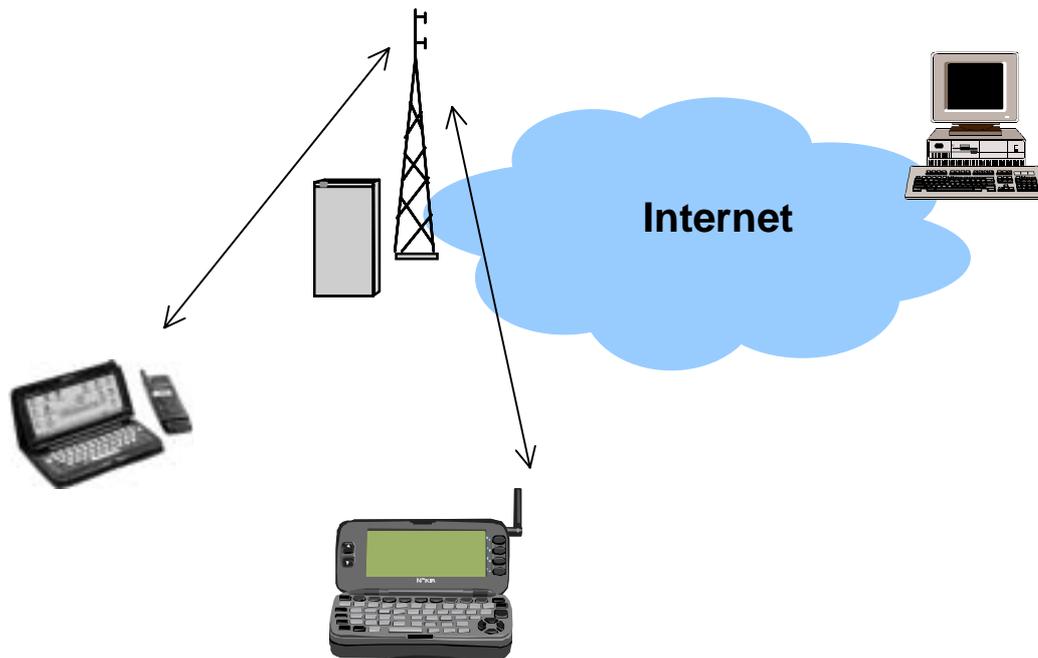


Figura 7.1 Acceso Internet a través de un enlace móvil

Internet, la mayor red de ordenadores del mundo, formada por millones de terminales repartidos por todo el planeta, está basada en la arquitectura de protocolos TCP/IP, cuyas aplicaciones más utilizadas son, sin duda, el correo electrónico, la transferencia de ficheros, el acceso remoto y el acceso a la Web, y todas ellas tienen en común el protocolo de transporte que utilizan, el TCP [Ste94, RFC793], que es extremo a extremo, fiable y orientado a conexión.

Los protocolos de la arquitectura TCP/IP fueron diseñados para ser utilizados en redes compuestas únicamente por ordenadores fijos y, en ningún momento, se planteó la posibilidad de que hubiesen de enfrentarse a un nuevo entorno, como es el canal radio móvil, que presenta características muy diferentes.

Por lo tanto, un aspecto crítico a la hora de determinar la viabilidad de ofrecer servicios de datos basados en aplicaciones TCP/IP en redes móviles es el del análisis de las posibilidades de acoplamiento de esta arquitectura de protocolos a este nuevo enlace de comunicación.

Puesto que en las redes fijas actuales las pérdidas de paquetes son debidas, fundamentalmente, a situaciones de congestión (ya que las tasas de error son muy bajas), las implementaciones modernas del TCP asocian cualquier error a este problema e incorporan toda una serie de mecanismos destinados a combatirlo [Jac88, Flo95, BrP95b, RFC2581].

Ahora bien, en los entornos móviles las pérdidas de paquetes son muy frecuentes debido a las altas tasas de error y a las desconexiones temporales asociadas al traspaso de las llamadas.

Por lo tanto, los algoritmos de control de la congestión que TCP activa al detectar los errores, pueden ocasionar un comportamiento muy poco eficiente, cuyos efectos deben ser evaluados.

Por esta razón, y teniendo en cuenta la gran popularización de las redes móviles, es importante el estudio del protocolo TCP tan extendido en la mayoría de redes actuales, de forma que pueda mejorarse su comportamiento en entornos móviles, tanto en la estación como en el terminal móvil.

Ya se ha comentado en capítulos anteriores que en la bibliografía existen varios estudios dedicados al análisis de aspectos relacionados con el comportamiento de TCP en los entornos móviles [MyS93, DCY93, DuR94, YaB94, BaB95, Cal95, BSA95, BSK95, ABS95, MGS95, BBK96, BBK97, CTG97, BKV96, BPS97].

El objetivo básico de este capítulo es analizar en profundidad los diversos factores que pueden incidir en la degradación del rendimiento del protocolo de transporte TCP cuando éste opera en estos entornos, y para ello hemos utilizado una implementación real de la arquitectura TCP/IP, encapsulado en PPP (*Point to Point Protocol*) [RFC1548], y se ha utilizado una aplicación de simulación de las características de transmisión del canal radio móvil.

Para este tipo de entornos se ha propuesto la optimización de diferentes aspectos como el tamaño de la unidad de transferencia. En este estudio se presenta además el efecto de las ráfagas de errores en las prestaciones del TCP.

7.2 PLATAFORMA EXPERIMENTAL

Este estudio se basa en la evaluación de una implementación real de TCP, *Linux* 1.2.13 basada en [RFC793], la cual incluye los mecanismos de “*Slow Start*” y “*Congestion Avoidance*” [Jac88], y los algoritmos de “*Fast Retransmit*” y “*Fast Recovery*” [Ste94].

La plataforma experimental se muestra en la Figura 7.2. La conexión TCP/IP se inicia usando la aplicación *ftp* en el PC que emula el terminal móvil. Los datos se encapsulan en PPP [RFC1548], y el Simulador de Canal introduce errores en la dirección del flujo de datos y/o en la de reconocimientos. Finalmente, otro PC recibe los datos emulando el terminal fijo. La conexión se traza mediante la aplicación *tcpdump*.

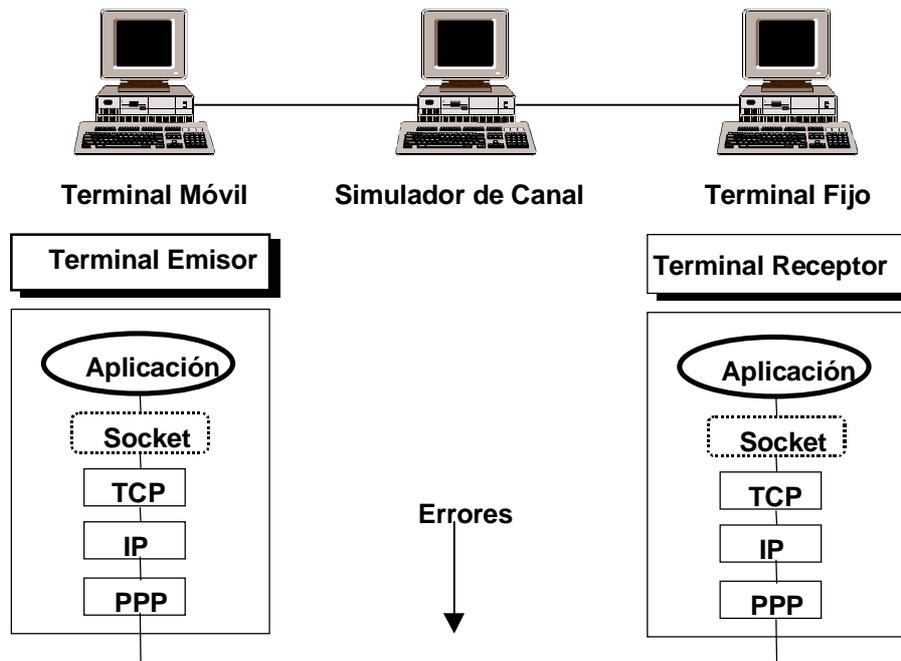


Figura 7.2 Plataforma experimental y arquitectura de protocolos

El Simulador de Canal introduce diferentes tipos de distribuciones de errores (uniforme, determinista y *Poisson*), así como los modelos de *Fritchman* [KaS78, Fri67, WaM95] y temporal, descritos en el Anexo B. En cada uno de los casos se define la BER a aplicar.

En la siguiente sección se discute la reacción del TCP sobre esta plataforma, emulando un enlace radio, teniendo especial importancia en los diferentes mecanismos que usa el TCP para garantizar una conexión extremo a extremo fiable.

7.3 EFECTO DE LA DISTRIBUCIÓN DE LOS ERRORES

El efecto combinado de los mecanismos generadores de error presentes en el canal radio móvil es tal que provoca que éstos no presenten una distribución uniforme, es decir, que no sean estadísticamente independientes sino que aparezcan en forma de ráfagas lo que hace imposible la comunicación durante periodos más o menos largos de tiempo.

Para evaluar el efecto especialmente nocivo de la distribución a ráfagas de los errores presentes en el canal radio, se han introducido errores simples de forma determinista y errores a ráfagas siguiendo el modelo simplificado de *Fritchman* (Anexo B) de manera que la tasa de error en bit en ambos casos fuese idéntica.

En la Figura 7.3 se muestra un gráfico comparativo del caudal en función de la tasa de error y de la distribución estadística de los errores. En todos los casos se transfirió un fichero de 100 Kbytes y se configuró la MTU del enlace de comunicación a 296 bytes. No se utilizó tipo alguno de compresión de cabeceras TCP/IP.

Los resultados muestran que las ráfagas de error degradan en mayor medida el comportamiento que los errores deterministas. En cualquier caso, queda claro que en presencia de errores la degradación en el comportamiento del protocolo debe ser evaluado.

Incluso para tasas de error en bit reducidas, la estadística de los errores tiene una influencia considerable. Como puede verse, el caudal alcanzado utilizando el modelo de Fritchman con una BER de 1 bit erróneo cada 100000 es, aproximadamente, un 10 % inferior al medido introduciendo errores simples de forma determinista. A medida que la BER del canal aumenta, la diferencia en el impacto sobre el rendimiento del protocolo entre ambas distribuciones se agudiza aún más. Para una BER de 1/15000 existe ya una diferencia del 20 %.

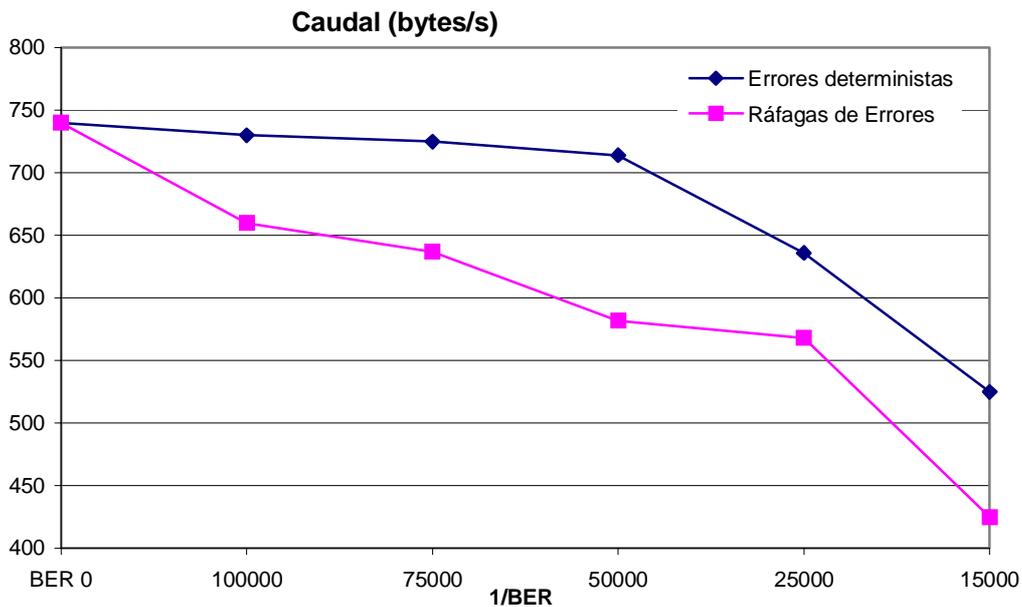


Figura 7.3 Evolución del caudal

Los errores simples suelen ser detectados a través del mecanismo de *fast retransmit*. Por esta razón, raramente necesita TCP esperar a la expiración del temporizador de retransmisión para poder retransmitir el segmento afectado y, en consecuencia, el canal de comunicación permanece, prácticamente la totalidad del tiempo que dura la comunicación, ocupado. El caudal en estos casos puede calcularse de manera aproximada como :

$$Caudal = \frac{IU}{IU + IC + IR} \times BW_{neto} \quad (7.1)$$

donde:

IU = Cantidad de información útil transferida.

IC = Cantidad de información correspondiente a cabeceras (encapsulado TCP/IP/PPP).

IR = Cantidad de información retransmitida.

BW_{neto} = Velocidad del enlace al 80% de eficiencia debido los bits de paridad de PPP.

No obstante también hay que tener en cuenta que la expiración del temporizador de retransmisión para detectar el error y la activación del mecanismo de *backoff* exponencial, resulta en que TCP no aprovecha toda la capacidad que le ofrece el enlace al generar pausas en la comunicación que, adecuadas para redes propensas a la congestión, son absolutamente innecesarias en enlaces como el radio en el que los errores responden a la existencia de intervalos específicos de mal funcionamiento del canal.

En estos casos el caudal responde a la expresión:

$$Caudal = \frac{IU}{IU + IC + IR + \Delta T \times BW_{neto}} \times BW_{neto} \quad (7.2)$$

donde ΔT indica el intervalo de tiempo que el canal permanece inactivo.

Para ver hasta qué punto la ecuación (7.1) se aleja mucho más de la realidad cuando los errores ocurren en ráfagas que cuando éstos aparecen de forma aislada, la Figura 7.4 compara los valores de caudal obtenidos experimentalmente en cada caso con los que esta ecuación predice, habiendo obtenido experimentalmente el valor de IR.

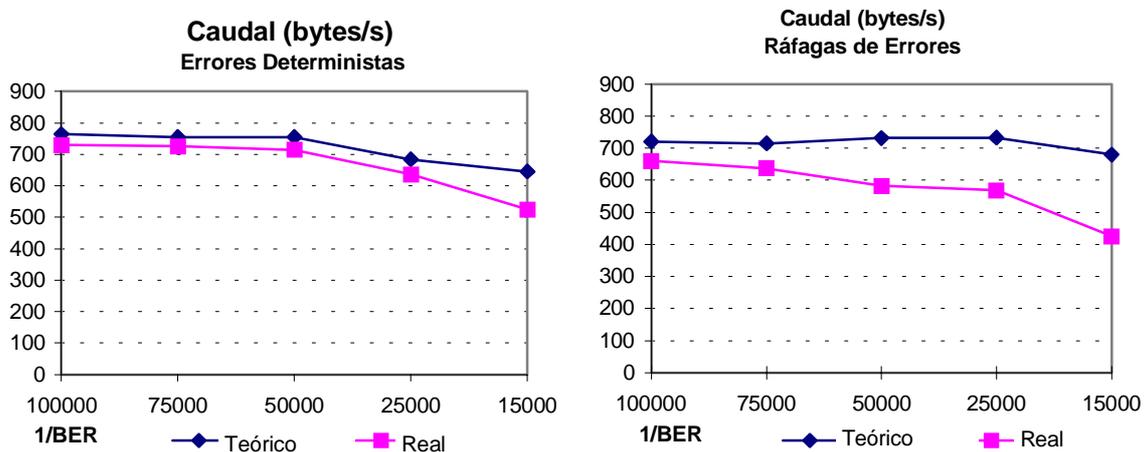


Figura 7.4 Comparación entre el caudal real y teórico

Como puede apreciarse, en ambos casos la discrepancia entre la curva teórica y la experimental aumenta a medida que lo hace la tasa de error. Sin embargo, independientemente de la BER del canal, los errores en ráfaga contribuyen a ampliar aún más esta diferencia. Esto es debido a que en estos casos es más frecuente que se produzca el algoritmo de *backoff*.

Puede resultar sorprendente observar que, para el caso de $BER=1/15000$, el caudal obtenido introduciendo errores simples deterministas se aleja considerablemente del que la ecuación (7.1) predice.

Un análisis pormenorizado del fichero *tcpdump* de esta transferencia nos permite comprobar cómo con esta tasa de pérdida se mantiene el valor de la ventana de congestión por debajo de los cuatro segmentos. El resultado es que TCP se ve forzado, en muchas ocasiones, a recurrir a la expiración del temporizador de retransmisión para detectar las pérdidas (no se recibe el número de duplicados suficientes para activar el algoritmo de Retransmisión Rápida). Estos periodos de espera asociados a la detección de los errores son los que contribuyen a alejar el valor teórico predicho por la ecuación (7.1) del valor realmente medido.

En la Figura 7.5 se compara la evolución, captada en recepción, de dos transferencias, una que introduce errores simples de forma determinista y la otra a ráfagas. En ambos casos la tasa de error en bit del canal se estableció en 1 bit erróneo cada 50000 bits de información transferidos.

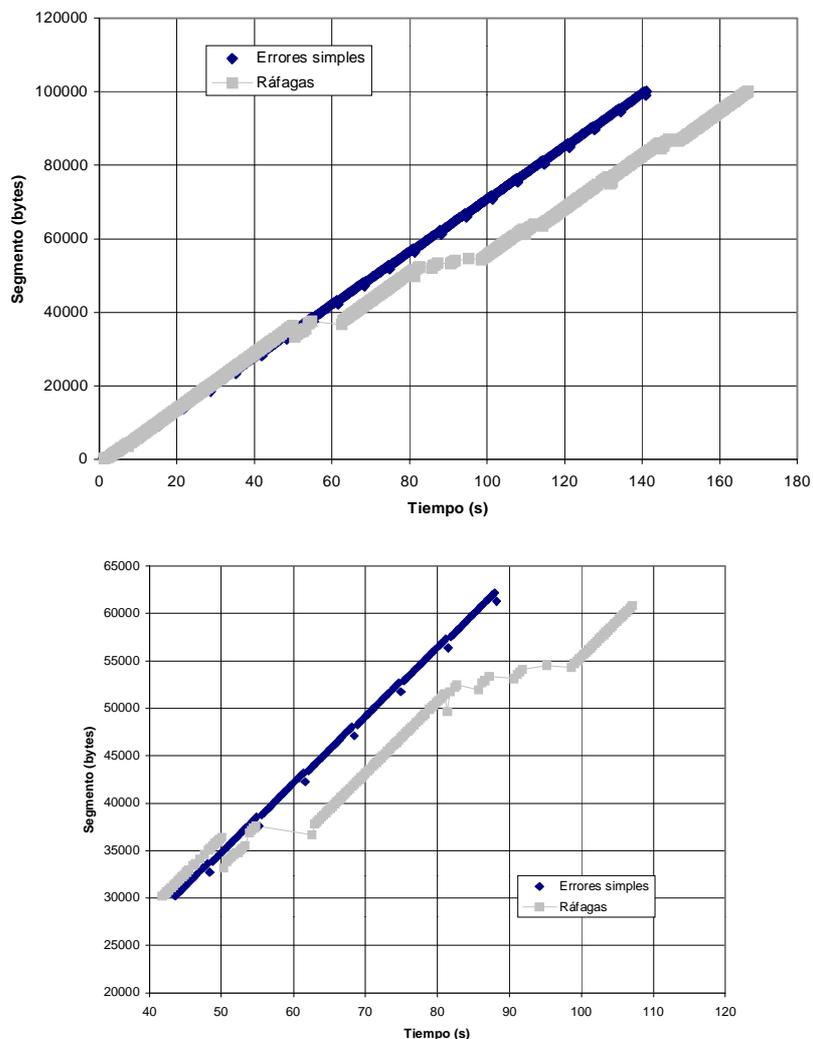


Figura 7.5 Evolución de la transferencia . Abajo: detalle de los errores

En el caso correspondiente a los errores simples, la traza presenta una pendiente prácticamente constante a lo largo de toda la conexión. El emisor, en todos los casos, detecta rápidamente la pérdida de cada paquete y se recupera de ella sin apenas dar lugar a discontinuidades en la curva; esto es debido a la aplicación de la estrategia de detección de errores a través de retransmisión rápida.

En cambio, en la gráfica correspondiente al modelo a ráfagas, destacan los tramos prácticamente planos de la traza debidos a intervalos de inactividad del canal de comunicación. Son estos intervalos los que degradan de forma muy significativa el rendimiento del protocolo de transporte.

El desigual impacto que tienen estos dos modelos de generación de errores se advierte también en la Figura 7.6 en la que se muestran los valores adoptados por el temporizador RTO, en ambos casos, a lo largo de la comunicación.

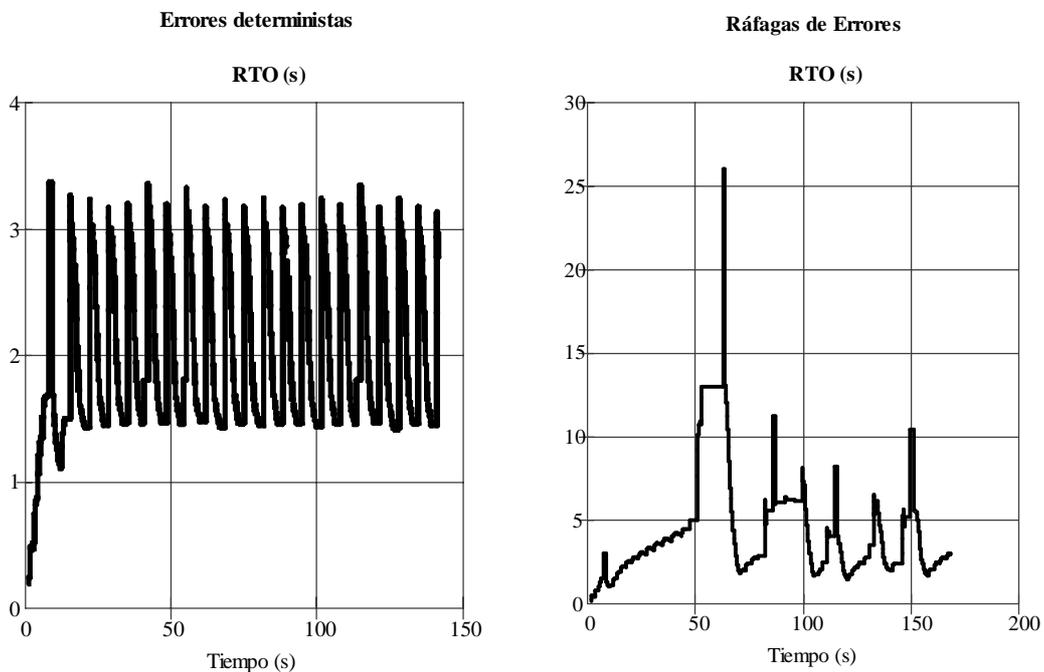


Figura 7.6 Evolución del valor de RTO a lo largo de la transferencia

En la Figura 7.6, con errores deterministas, vemos que tras detectar el error y retransmitir el segmento afectado, el emisor TCP dobla RTO y vuelve, rápidamente, a transmitir con normalidad tomando nuevas muestras de RTS y calculando los nuevos valores de RTT y RTO. En cambio, en el caso de errores a ráfagas, el proceso de recuperación es mucho más traumático. Cabe notar la diferencia en la escala de los valores de RTO de un caso y del otro. Los elevados valores que adopta RTO son una clara indicación de lo que ocurre: las sucesivas retransmisiones de segmentos están forzando al emisor a recurrir, con frecuencia, a la activación del mecanismo de *backoff* exponencial, es decir, a frenar la inyección de datos durante largos periodos de tiempo. **Una vez más vemos cómo esta estrategia, idónea para equilibrar el flujo de datos en una red congestionada, es exactamente la que menos**

conviene cuando los problemas provienen de los frecuentes errores de transmisión que ocurren en el canal radio móvil.

En conclusión, en los errores deterministas, las pérdidas se detectan rápidamente mediante los reconocimientos positivos del TCP. No obstante, en el otro caso, los errores afectan una ventana de transmisión TCP, y puede verse el largo periodo de inactividad lo que produce la degradación en el comportamiento. Especialmente, la evolución del RTO en el caso de ráfagas de errores muestra la activación del "backoff" exponencial [Ste94], lo que frena la velocidad de transmisión de TCP.

7.4 EFECTO DE LAS DESCONEXIONES

Otro de los aspectos importantes de la degradación de TCP en entornos móviles es el efecto de las desconexiones. Al producirse esta situación, el flujo extremo a extremo tanto de datos como de reconocimientos se detiene. Los mecanismos anteriormente descritos de TCP, actuarán provocando que la fuente retransmita el primer segmento no reconocido, aplicando el algoritmo de *backoff* exponencial e iniciando los algoritmos contra la congestión (Inicio Lento). No obstante, todos los esfuerzos serán en vano hasta que se restablezca la conexión, ya que ninguno de los segmentos llegará al destino. Durante el periodo de desconexión, el algoritmo de *backoff* irá espaciando las retransmisiones (cabe recordar que este mecanismo está pensado para solucionar las situaciones de congestión), doblando el valor del temporizador RTO.

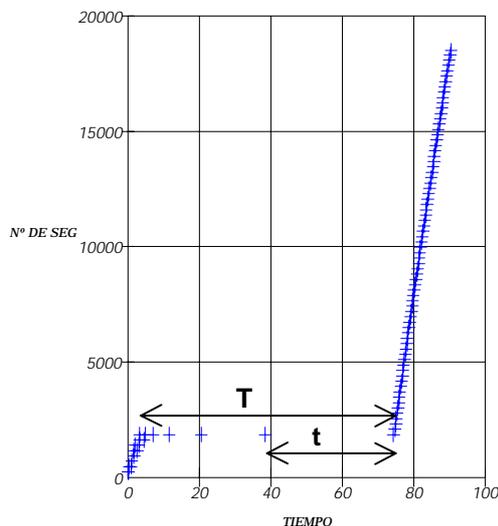


Figura 7.7 Efecto de las desconexiones

La Figura 7.7 muestra la evolución de los segmentos enviados por la fuente. El tiempo durante el cual se produce la pausa en la comunicación es de T segundos. En la figura puede apreciarse que se retransmite seis veces el primer segmento no reconocido. Mientras dure la desconexión del canal, como que igualmente no podrán transmitirse datos, no importa si el

segmento se retransmite o no. No obstante, una vez se restablece el canal, **interesaría que la retransmisión del paquete se realizara lo antes posible**. Si suponemos que el canal se restablece inmediatamente después de la penúltima retransmisión, el canal permanecerá ocioso durante RTO segundos más, valor que al haberse producido el *backoff* puede ser elevado (la mayoría de implementaciones tienen un valor RTO máximo de 120 o 240 segundos). En este ámbito, deben proponerse mecanismos que detecten la actividad del canal.

Normalmente, en situaciones de congestión o errores simples en el canal, es poco probable que se pierda completamente el flujo de datos en ambos sentidos de la comunicación, con lo cual, en caso de no recibir ningún reconocimiento durante un cierto espacio de tiempo, sería adecuado volver a poner el temporizador de retransmisión a su valor antes de la retransmisión, de forma que la pérdida se recupere lo antes posible después del restablecimiento.

Este algoritmo es el que proponemos y tanto es adecuado para casos de congestión o pérdidas por errores como para los casos de desconexiones.

El algoritmo opera de la siguiente forma:

- Se produce una expiración del temporizador de retransmisión:
 - Se guarda el valor de RTO y se pone el sistema en estado de *backoff*
 - Se aplica el algoritmo de *backoff* (doblar RTO y retransmitir el segmento en cuestión)
 - Se activa un temporizador de recepción de reconocimientos, t_{ack} . El valor de t_{ack} debe ser el valor de RTO al producirse la expiración del temporizador, ya que es el valor que nos indica el retardo existente en el enlace en el momento de producirse la retransmisión.
- Si durante t_{ack} no se recibe ningún reconocimiento, entonces
 - Causado por desconexión
 - Se pasa al estado de desconexión
 - Se retransmite el segmento cada RTO segundos, siendo RTO el valor guardado en el instante de producirse la retransmisión.
 - Sino, causado por congestión o pérdida
 - Si el *ack* reconoce nuevos datos
 - Salir del estado de *backoff*. Ya se ha recuperado la pérdida
 - Sino, es un duplicado.

- Una vez el sistema ha entrado en el estado de desconexión, se irán retransmitiendo los segmentos hasta que se reciba un reconocimiento. En ese instante:
 - Si es un duplicado
 - Mantenerse en el estado de desconexión
 - Si reconoce a datos nuevos
 - Salir del estado de desconexión. Ya se ha recuperado la pérdida.

Los nuevos estados se definen como:

- Estado de backoff: Periodo durante el cual se aplica el mecanismo normal de TCP de *backoff* exponencial.
- Estado de desconexión: Periodo durante el cual se retransmite el segmento a intervalos RTO.

Con este algoritmo se mejora la recuperación del protocolo después de una desconexión. Las mejoras que pueden obtenerse con el algoritmo dependerán de:

- Duración de las desconexiones
- Valor de RTO al producirse la desconexión
- Valor máximo de RTO
- Duración de la transferencia. Cuanto más corta sea la transferencia (como accesos a Web), peor será el efecto ya que el tiempo de pausa será relevante respecto a la duración total de la transmisión.

El inconveniente del algoritmo es que mientras dure la desconexión, se introducirán en el enlace más segmentos de los que se enviarían en caso de aplicar el *backoff* exponencial, pudiendo producir congestión en los enlaces. Además, en el peor de los casos, si actúan algoritmos a nivel de enlace con recuperación de errores, el hecho de enviar más paquetes producirá la saturación del enlace. No obstante, la interacción de los mecanismos de recuperación de errores a nivel de enlace quedan fuera del ámbito de esta Tesis, ya que no afectarían únicamente en este algoritmo, sino al resto de mecanismos que utiliza TCP para la recuperación de errores.

En el caso de actuar erróneamente (es decir, actuar como si fuese desconexión y que en realidad sea congestión), empeorará la situación, no obstante, la probabilidad de que esto suceda es muy pequeña.

7.5 INFLUENCIA DEL TAMAÑO DEL PAQUETE

Otro aspecto importante es el análisis del tamaño del paquete más idóneo para TCP en enlaces con error.

En este apartado, constataremos cómo el tamaño de los paquetes que circulan por el enlace de comunicación tiene un enorme impacto sobre el rendimiento, en términos de caudal, que puede llegar a alcanzarse en una transferencia de datos TCP a través del canal móvil.

En primer lugar, estudiaremos el caso de un canal completamente libre de errores. Seguidamente presentamos los resultados obtenidos al experimentar con *errores simples deterministas*. En los siguientes apartados generaremos secuencias de error a ráfaga a través del *Modelo Simplificado de Fritchman* y *Modelo Temporal*.

7.5.1 MEDIDAS EN AUSENCIA DE ERRORES

Cuando el canal no introduce errores, es evidente que cuanto mayor es la MTU del enlace mayor es, también, la eficiencia conseguida en la transferencia de los datos. Esto es así puesto que el tamaño de la información no útil (cabeceras, *trailers*, corrección de errores, ...) se mantiene constante independientemente de la cantidad de información útil incluida en el campo de datos del segmento TCP.

En la Figura 7.8 se muestra el caudal obtenido experimentalmente, en función de la MTU del canal. La velocidad del enlace es de 9600bps (dando una velocidad efectiva de 960 bytes/s) y el tamaño de fichero transferido es de 100Kbytes. Puede observarse que el caudal máximo se obtiene para la MTU mayor.

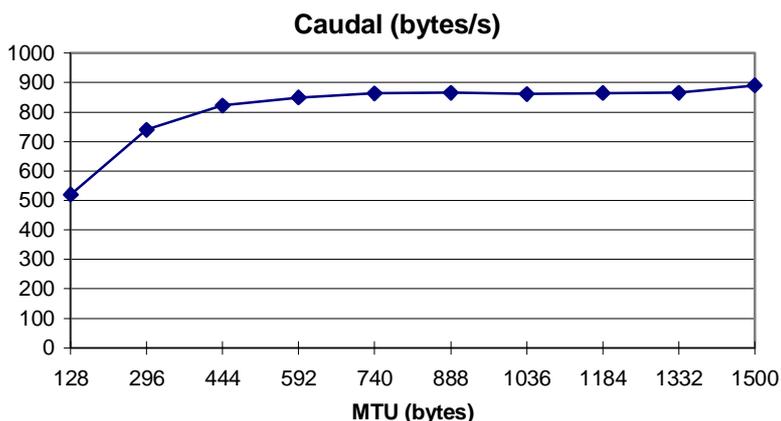


Figura 7.8 Caudal obtenido en ausencia de errores

7.5.2 ERRORES DETERMINISTAS

En este apartado someteremos la transferencia *ftp* de un fichero de 100 kbytes al patrón de errores deterministas y evaluaremos el efecto que la variación en el tamaño de paquete utilizado

tiene sobre el caudal alcanzado en la comunicación. La velocidad del enlace sigue siendo de 9600bps. La simplicidad de este modelo de error permite describir y analizar, los resultados obtenidos.

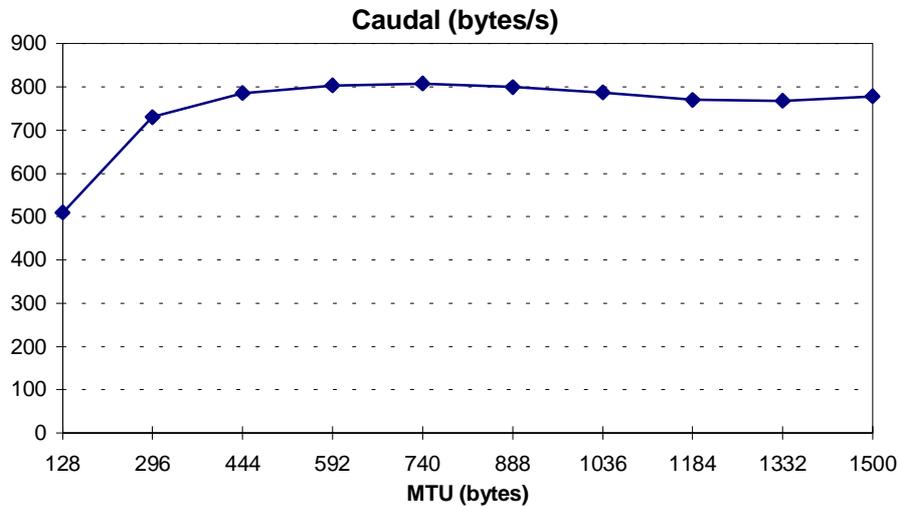


Figura 7.9 Caudal en función de la MTU del enlace para una BER de 1/100000

En la Figura 7.9 se muestran las medidas de caudal conseguidas al introducir exactamente 1 bit erróneo cada 100000 bits.

De la gráfica puede observarse que la curva ya no muestra una tendencia creciente como ocurría en el caso anterior, que existe un valor de MTU para el cual el caudal medido es máximo, y que a medida que la MTU del enlace se aleja de esta MTU *óptima*, el caudal muestra una tendencia decreciente.

Podemos decir que la MTU que proporciona un caudal máximo es el resultado de un compromiso:

- Por un lado, cuanto mayor es el tamaño de los paquetes que circulan por el enlace, tanto mayor es la probabilidad de que alguno de los bits que lo constituyen resulte afectado por un error. Teniendo en cuenta que un sólo bit erróneo se traduce, indefectiblemente, en la pérdida de todo el paquete y, por tanto, en la ulterior retransmisión de su contenido al completo, vemos que el caudal es mucho más sensible a los errores cuando se utilizan paquetes de gran tamaño que cuando éstos son reducidos.
- Por otro lado, la eficiencia en la utilización del enlace puede verse resentida si se utilizan paquetes de tamaño reducido debido a la excesiva cantidad de ancho de banda empleada en transmitir información no útil.

Este compromiso quedaba ya reflejado en la ecuación (7.1). En ella se muestra la expresión del caudal para un canal como el que estamos considerando en el que los errores se producen de forma aislada.

El compromiso en la elección del tamaño de paquete se hace patente, y queda reflejado gráficamente en la Figura 7.10, donde se muestran los valores de la Información de Control (IC) y de los datos retransmitidos (IR), para cada valor de MTU. Asimismo, la suma de ambos factores también se ha considerado interesante.

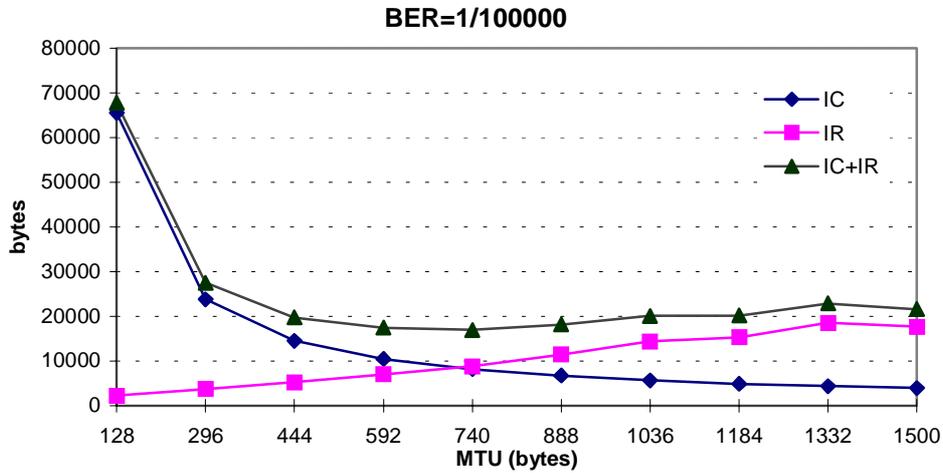


Figura 7.10 Evolución de los factores IC e IR

A medida que aumenta la MTU del enlace, **IC** adquiere valores decrecientes, puesto que la proporción información no-útil/información útil va disminuyendo. Por otro lado, **IR** mide el número de bytes retransmitidos y, por tanto, aumenta conforme la MTU lo hace. La suma de **IC** e **IR** es mínima y, en consecuencia, el caudal máximo. En este caso el valor de MTU óptima es de 740 bytes.

No obstante, a medida que la BER del canal aumenta, ya hemos visto que el caudal debe verse afectado por el tiempo que el canal permanece ocioso debido al tiempo de recuperación de los errores. Además, cuando experimentaremos con patrones de error a ráfaga, esta influencia se verá más acusada. En la Figura 7.11 puede verse como se diferencia el caudal obtenido experimentalmente y el obtenido mediante la ecuación (7.1) a medida que aumenta la BER del canal. Los valores de MTU no representados en la figura significan que la transferencia no pudo realizarse debido a la gran cantidad de errores.

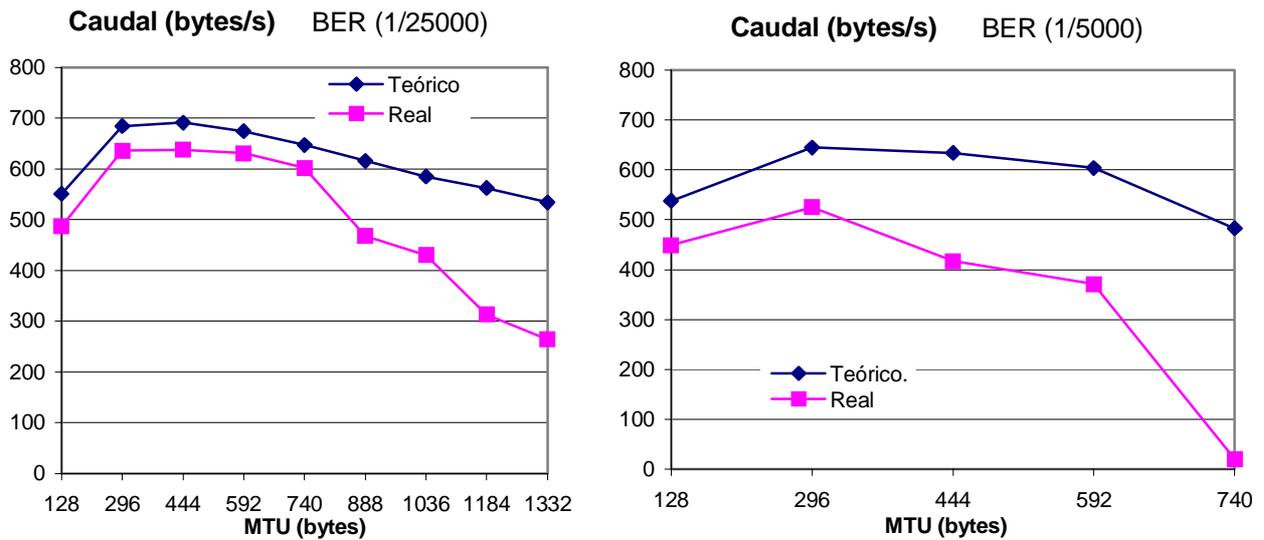


Figura 7.11 Comparación entre el caudal real y teórico

Cabe recordar aquí que, para una tasa de error en bit dada, la probabilidad de pérdida de un paquete es tanto mayor cuanto mayor es el tamaño del paquete. Esta es la razón por la que las diferencias entre caudal real y el obtenido teóricamente se agudizan cuando la MTU del enlace adopta valores crecientes. A medida que aumenta la tasa de error del canal, la menor vulnerabilidad de los paquetes de tamaño reducido adquiere importancia convirtiéndose en el factor que determina el rendimiento final de la transferencia. A altas tasas de pérdida de paquete, TCP se ve obligado, con frecuencia, a esperar a la expiración del temporizador de retransmisión (ya que la ventana es tan estrecha que no permite la detección mediante la recepción de reconocimientos correspondientes a segmentos enviados con posterioridad al erróneo) y, en general, a frenar la inyección de datos mediante estrategias destinadas a combatir la congestión.

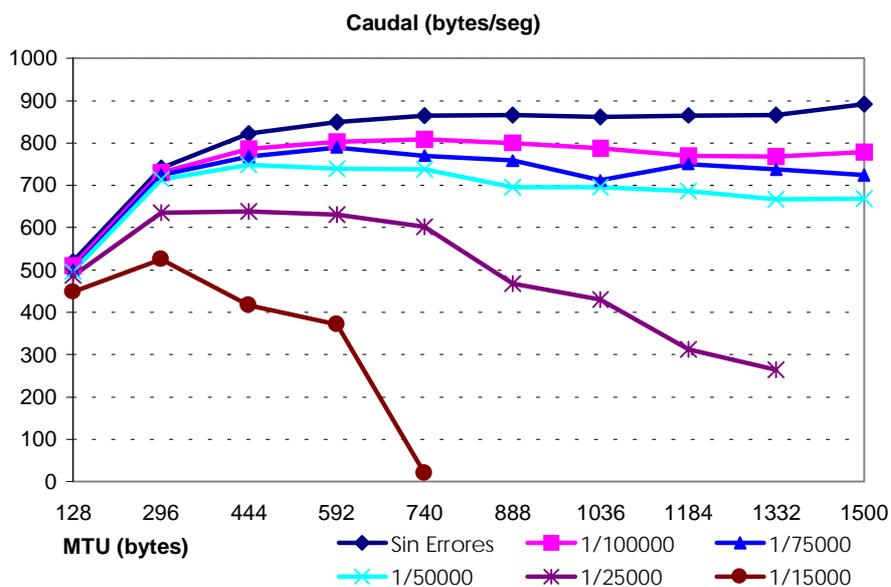


Figura 7.12 Medidas de caudal en función de la BER del canal y la MTU del enlace

Finalmente, la Figura 7.12 presenta los resultados de caudal en función de la MTU del enlace obtenidos para el caso de errores deterministas. En ella se pone de manifiesto la importancia de transmitir con la MTU óptima para las características del enlace.

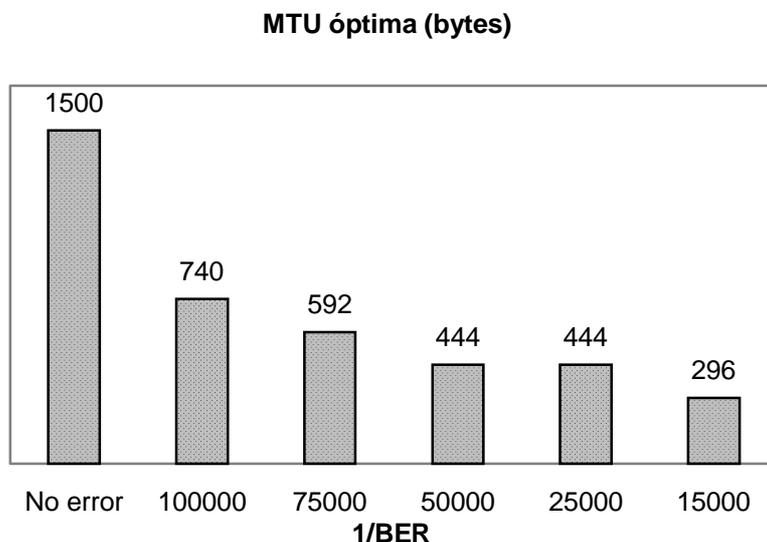


Figura 7.13 Valores de MTU óptima en función de la BER del canal

A modo de resumen se presentan los valores de la MTU óptima para cada una de los casos de error, con la que se obtiene un caudal máximo. De la Figura 7.12 puede verse la ganancia que se obtiene en cada caso por el hecho de utilizar la MTU óptima.

7.5.3 CANALES CON MEMORIA: MODELO SIMPLIFICADO DE FRITCHMAN

El modelo de *errores simples deterministas* nos ha permitido llevar a cabo un estudio detallado de los resultados debido a que los errores se producen de forma aislada. Este estudio ha puesto de manifiesto la decisiva importancia de la elección de un valor adecuado para la MTU del enlace.

En este apartado pretendemos comprobar si este carácter crítico del tamaño de paquete empleado es igualmente extensible al caso más realista de canales con memoria, es decir, canales como el móvil en los que los errores muestran una cierta correlación dando lugar a ráfagas. Los valores asociados al modelo se encuentran en el Anexo B (parametrización del modelo de Fritchman). En todos los casos se transfirió un fichero de 100Kbytes sobre un enlace de 9600bps.

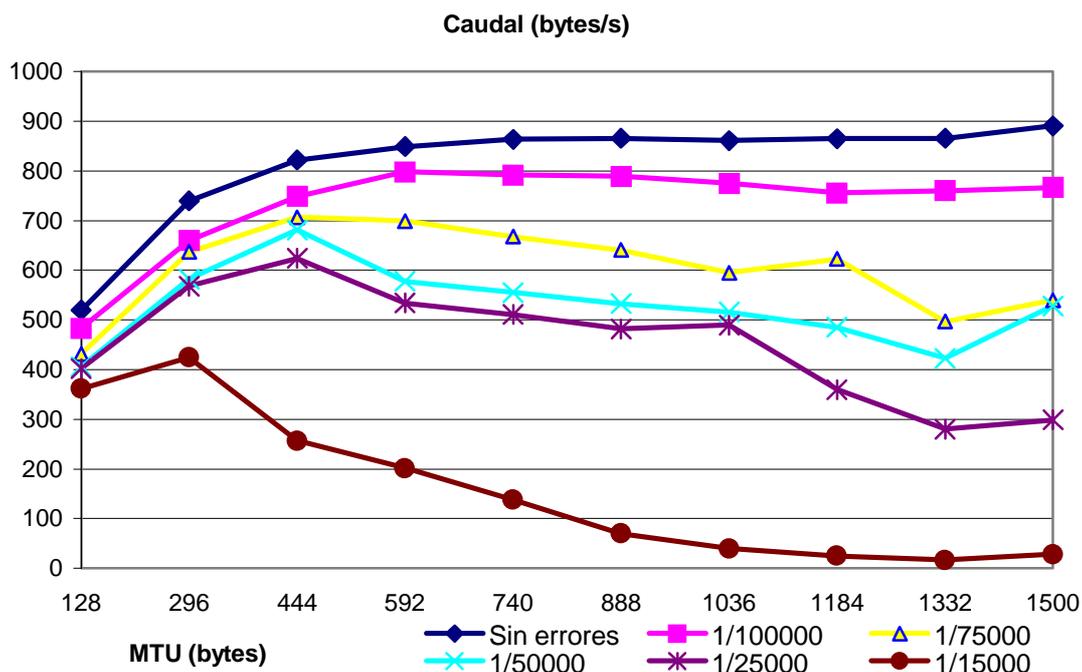


Figura 7.14 Medidas de caudal

Antes de pasar a evaluar el efecto que tiene la variación de la MTU del enlace en este caso, nótese que la comparación de la Figura 7.14 con la Figura 7.12 permite confirmar las conclusiones extraídas en el apartado anterior acerca de la influencia de la distribución no uniforme de los errores y hacerlas extensibles a cualquier tamaño de paquete. Como puede apreciarse, los valores de caudal de la Figura 7.14 son, prácticamente en todos los casos, inferiores a los equivalentes de la Figura 7.12. La única excepción a esta regla general la constituyen casos como, por ejemplo, el correspondiente a BER=1/25000 y $MTU \geq 888$ en los que, pese a todo, los valores de ambas gráficas son muy similares. Hay que tener en cuenta, no obstante, que cuando el tamaño de paquete es de 888 bytes y se produce un error cada 25000 bits, uno de cada cuatro paquetes, en media, resulta corrompido en el caso de errores simples deterministas por lo que la hipótesis de errores aislados que, en cierta manera, presupone la capacidad del emisor de recuperarse ante los errores sin necesidad de recurrir a la expiración del temporizador o a la activación de mecanismos de control de la congestión, comienza a perder validez. Esto es aún más claro cuando el tamaño de paquete o la BER del canal aumentan. En cualquier caso, se demuestra que TCP es un protocolo extremadamente sensible a la correlación en los errores.

Cuando los errores se distribuyen en ráfagas, la ecuación que permite calcular el caudal alcanzable en una transferencia deja de depender exclusivamente del volumen de información de cabecera y del número de retransmisiones que TCP se ve obligado a realizar. En este caso, la ecuación queda también determinada por los intervalos de inactividad en el canal (ΔT), como ya vimos en la ecuación (7.2), que se generan como consecuencia de la aplicación de los algoritmos de control de la congestión y retransmisión de TCP.

La ecuación (7.2), por tanto, establece que el caudal es el resultado de un compromiso triple. Además de los dos factores analizados en el apartado anterior (**IC** e **IR**), aparece un tercero ΔT como consecuencia de las pausas en la comunicación.

Durante el intervalo ΔT podemos medir el número de bytes que TCP deja de transmitir por culpa de las pausas que se producen en la comunicación. La duración de estos intervalos depende de factores tales como: el valor adoptado por la variable RTO en el momento de producirse una ráfaga de error (ya que este valor indica el tiempo que invertirá TCP en detectar el error a través de la expiración del temporizador de retransmisión) y la distribución de los errores que constituyen la ráfaga (es decir, su impacto concreto sobre la ventana del emisor).

Una vez determinadas las causas que contribuyen a incrementar el valor de ΔT , cabe preguntarse cómo influye la configuración de la MTU del enlace de comunicación en el valor de este parámetro.

El factor **IR** (a diferencia de **IC** que solamente depende del volumen de información de control), sí se ve afectado por la distribución de los errores. En la Figura 7.15 se indica el número de bytes retransmitidos en función de la MTU del enlace y la tasa de error en bit introducida.

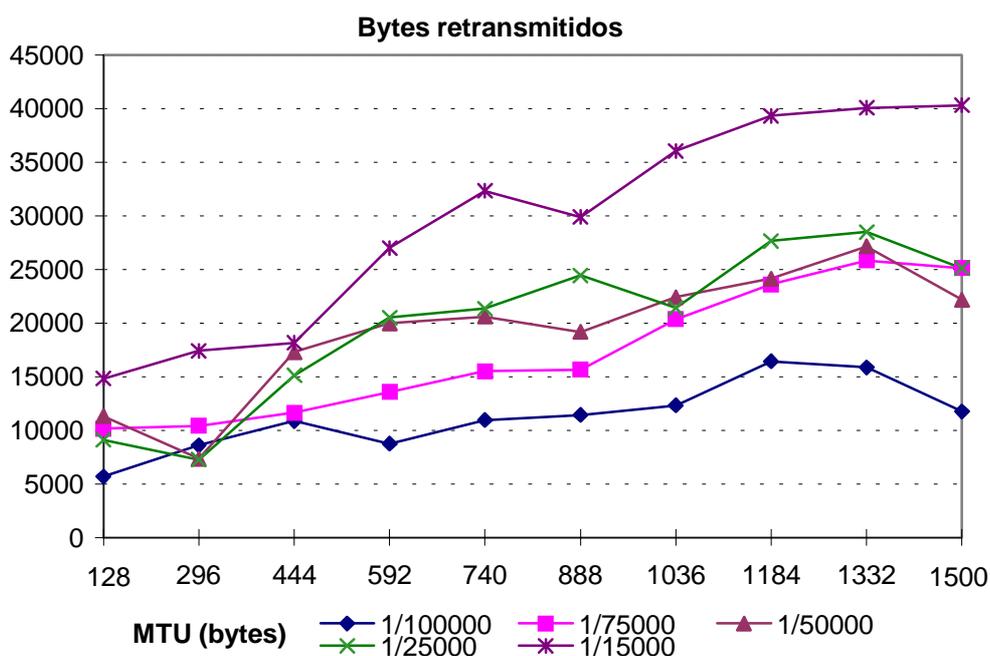


Figura 7.15 Número de bytes retransmitidos en función de la MTU del enlace y la tasa de error en bit del canal

En general, el volumen de información retransmitida aumenta a medida que lo hace el tamaño de paquete y la BER del canal. Sin embargo se observan algunos casos a priori sorprendentes cuya justificación hay que buscarla en el tercero de los factores, ΔT .

Por un lado, RTO constituye una estimación (sobre dimensionada) del tiempo de ida y vuelta de los paquetes que envía el emisor TCP; esto significa que si los paquetes utilizados son de gran tamaño, los valores adoptados por esta variable serán también elevados y, en consecuencia, la expiración del temporizador de retransmisión tras una pérdida no detectable mediante

reconocimientos tardará más en producirse. En este sentido, las pausas en la comunicación asociadas a la detección de la pérdida de segmentos serán más largas cuanto mayor sea el tamaño de paquete empleado. Por otro lado, cuanto mayor es la MTU del enlace, más tarda la ventana del emisor en ensancharse (los reconocimientos tardan más en llegar); esto implica que, dada una ráfaga de errores, la probabilidad de que el emisor pueda recibir un número suficiente de ACK como para percatarse de la situación disminuye, lo cual contribuye a incrementar el número de intervalos de espera. Además, como es obvio, el incremento en la tasa de error del canal favorece todo este tipo de situaciones en las que TCP desaprovecha la capacidad del canal durante ciertos periodos de tiempo.

La Figura 7.16 pone en evidencia que, a medida que la BER y/o el tamaño de paquete empleado aumentan, la cantidad de información que TCP se ve obligado a retransmitir deja de ser la causa principal del deterioro del caudal alcanzado en las transferencias cediendo el lugar a los intervalos en los que la comunicación permanece interrumpida. En estas gráficas se compara el caudal teórico que se hubiese alcanzado en cada uno de los casos estudiados en este apartado si este parámetro respondiese a la expresión indicada en la ecuación (7.1), es decir, si el canal no permaneciese nunca ocioso y, en consecuencia, sin tener en cuenta el factor ΔT .

En todos los casos, el caudal teórico es superior al real, lo que demuestra que TCP no sólo "pierde tiempo" retransmitiendo sino también detectando las situaciones de error y tratando de combatir estados que erróneamente asocia a la congestión del enlace de comunicación. En el apartado anterior, en cambio, pudimos comprobar cómo la ecuación (7.1) describía con mayor aproximación la realidad obtenida experimentalmente cuando los errores se producían de forma aislada y sólo a medida que éstos comenzaban a concentrarse debido al incremento en la tasa de error del canal, comenzaba a ser menos válida. En el caso que contemplamos ahora, pese a estar utilizando los mismos valores de BER, las diferencias entre el caudal teórico y el real son mucho más significativas y se agudizan enormemente a medida que la tasa de error aumenta.

Volviendo a la Figura 7.14, observamos cómo para cada situación concreta de error sigue existiendo un valor de MTU que denominamos óptimo para el cual el caudal conseguido es máximo. A medida que el tamaño de paquete se aleja del óptimo, tanto en valores crecientes como decrecientes, el caudal tiende a disminuir.

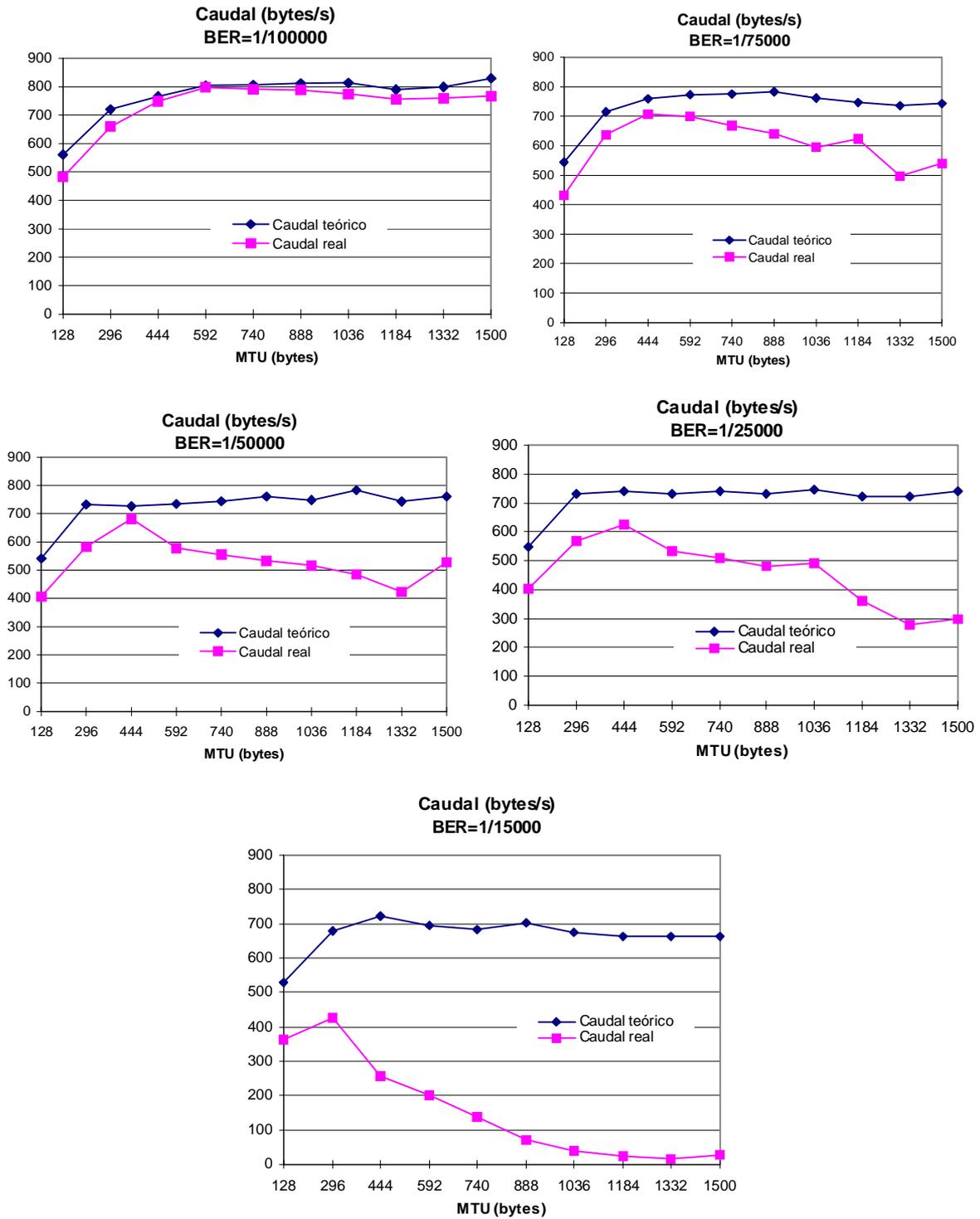


Figura 7.16 Comparación entre el caudal real y el teórico

Además, cuanto mayor es la tasa de error del canal, menor es el tamaño de paquete que proporciona la mayor eficiencia en la comunicación. Este efecto se muestra en la Figura 7.17.

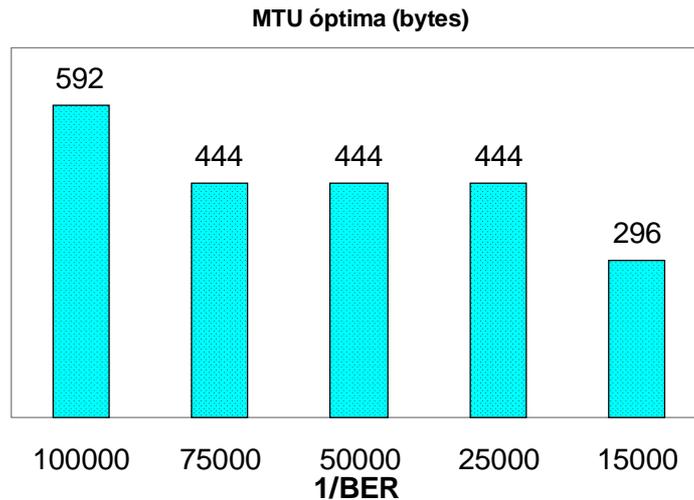


Figura 7.17. MTU que proporciona el caudal máximo

Según indica la Figura 7.17, la MTU óptima para $BER=1/75000$, $1/50000$ y $1/25000$ es siempre 444 bytes. Podemos constatar la tendencia decreciente de la MTU óptima. Para la menor de estas tasas de error en bit, $BER=1/75000$, el caudal conseguido con $MTU=444$ y $MTU=592$ es muy similar mientras que para $MTU=296$ es sensiblemente inferior, tal y como vimos en la Figura 7.14. Sin embargo, a medida que la BER aumenta, la tendencia se invierte y $MTU=296$ proporciona mejores resultados que $MTU=592$ por lo que aunque el valor discreto óptimo de MTU se mantiene fijo, la tendencia a valores menores es evidente y se confirma para $BER=1/15000$ donde $MTU=296$ da lugar al mayor rendimiento.

A tenor de los resultados obtenidos podemos concluir que la tendencia constatada en el apartado anterior respecto a la influencia que la variación de la MTU del enlace tiene sobre el caudal alcanzado en las transferencias de datos, es extensible al caso real de canales con memoria. En el apartado siguiente, tendremos ocasión de afianzar aún más esta conclusión cuando presentemos los resultados obtenidos al utilizar el Modelo Temporal descrito en el Anexo B que permite generar ráfagas de error cuya duración sigue una distribución exponencial.

7.5.4 CANALES CON MEMORIA: MODELO TEMPORAL

Para completar el análisis de la influencia del tamaño de paquete, hemos utilizado el Modelo Temporal (Anexo B). La experimentación con este otro modelo nos permitirá verificar la validez de los resultados expuestos en los apartados anteriores.

En la Figura 7.18 se muestran los valores de caudal obtenidos en función de la MTU del enlace y la duración media del intervalo de ráfaga (la duración media del intervalo entre ráfagas, $1/\lambda$, se mantuvo fija en 10 segundos). En todos los casos se transfirió un fichero de 200 kbytes de información, sobre un enlace de 9600bps.

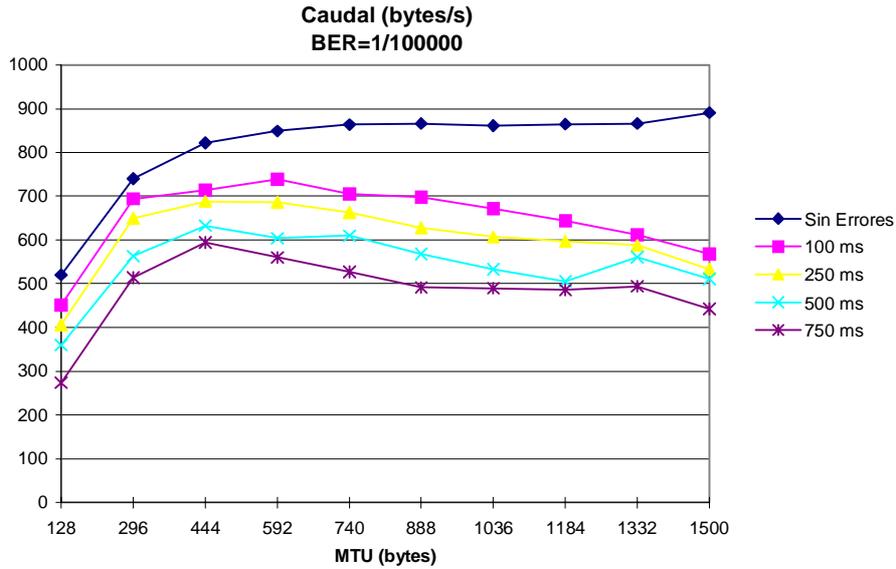


Figura 7.18 Medidas de caudal

En la Figura 7.18 se refleja, nuevamente, el compromiso entre la mejor utilización del ancho de banda del canal que proporcionan los paquetes de gran tamaño y la menor vulnerabilidad a los errores de los paquetes pequeños.

En términos de cantidad de información retransmitida, volvemos a constatar cómo, en ocasiones, transferencias en las que el caudal es menor que en otras dan lugar a un número menor de bytes retransmitidos. Esto puede verse en la Figura 7.19 en la que, al igual que en la Figura 7.15, pese a observarse una tendencia general al incremento en el volumen de información retransmitida a medida que aumentan la tasa de error y el tamaño de paquete es posible distinguir varios casos en los que esta tendencia no se verifica. Estos casos son más frecuentes a medida que la tasa de error del canal y/o la MTU del enlace aumenta.

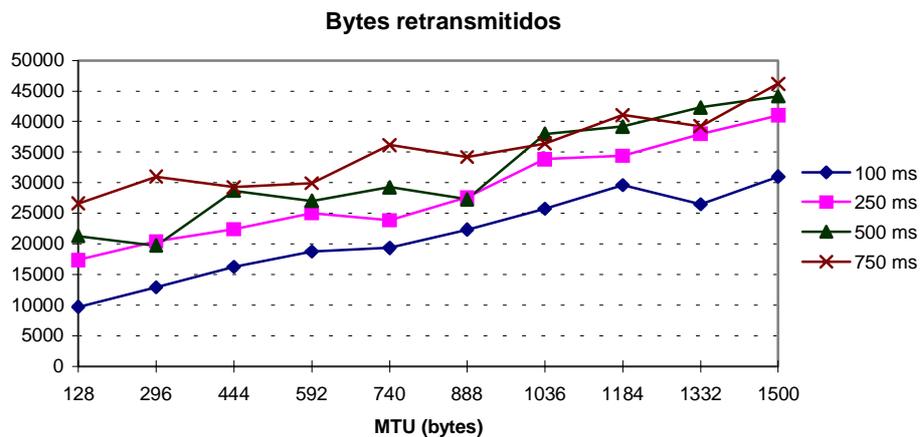


Figura 7.19 Cantidad de datos retransmitidos

Vamos a ver ahora como podemos comparar el caudal obtenido con el teórico, en el caso de tener el modelo temporal.

En este caso, durante los periodos de desconexión, seguro que no podremos enviar información, de forma que el caudal máximo teórico será:

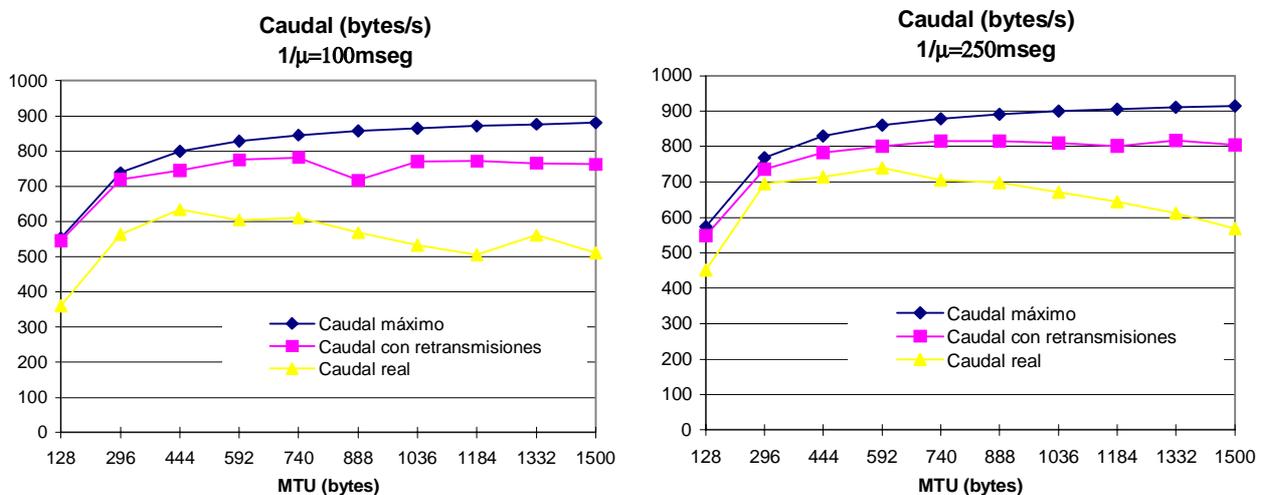
$$Caudal_{max\ te\acute{o}rico} = \frac{\frac{1}{\lambda}}{\frac{1}{\lambda} + \frac{1}{\mu}} \times Caudal_{max} = \frac{\mu}{\lambda + \mu} \times Caudal_{max} \quad (7.3)$$

donde $Caudal_{max}$ es el correspondiente a la ecuación (7.2), en la que se tiene en cuenta la cantidad de datos retransmitidos y el tiempo de inactividad del canal para la recuperación de la pérdida.

En la Figura 7.20 se muestra, para cada valor de μ la comparación entre:

- Caudal Máximo: es el valor máximo teórico más optimista que el que se presenta en la ecuación (7.3), ya que asume que TCP se adapta perfectamente a las condiciones del canal y sólo transmite cuando éste está en buenas condiciones.
- Caudal con Retransmisiones: es el valor correspondiente al caudal que TCP conseguiría si su rendimiento sólo se viera afectado por las retransmisiones asociadas a los errores introducidos por el canal, adaptándose también perfectamente a las condiciones del canal.
- Caudal Real: caudal obtenido experimentalmente.
-

Puesto que el Caudal Máximo no tiene en cuenta las retransmisiones, su valor es siempre creciente a medida que el tamaño de paquete aumenta. De hecho, no es más que el resultado de multiplicar el caudal alcanzable en ausencia de errores por una constante independiente del tamaño de paquete. En cualquier caso, resulta evidente a partir de la Figura 7.20 que el Caudal Máximo constituye una sobrevaloración excesiva del Caudal Real.



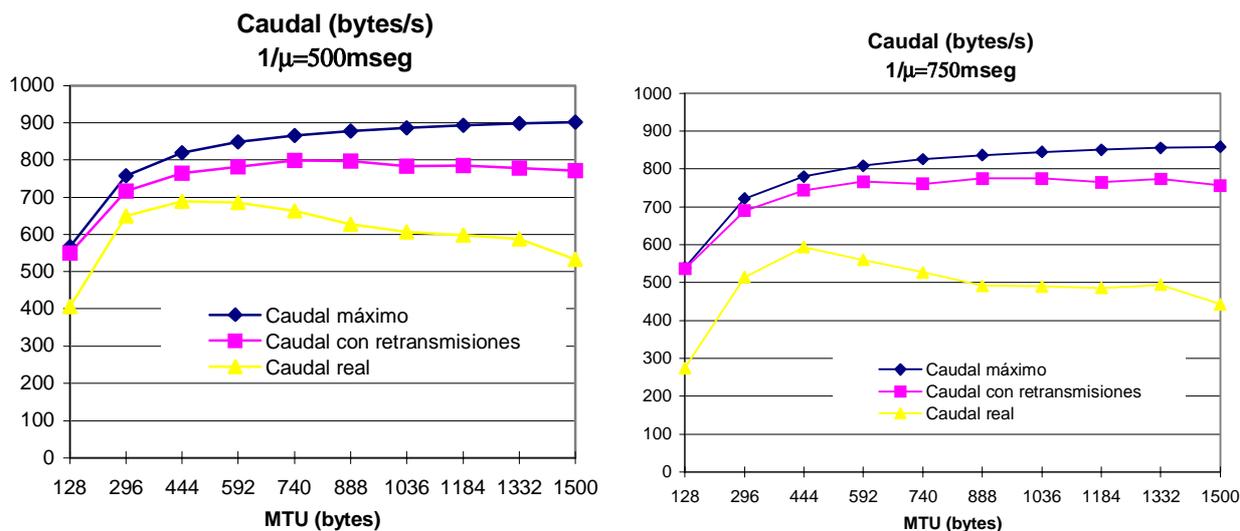


Figura 7.20 Comparación de Caudal

También el Caudal con Retransmisiones ofrece unos valores que, contrastados con la realidad, se revelan demasiado optimistas incluso en las condiciones de canal más favorables. La correlación que presentan los errores en este caso contribuye a distanciar notablemente los valores en ambos casos.

7.6 CONCLUSIONES

A lo largo de este capítulo hemos visto el efecto de los errores en el rendimiento del protocolo TCP. Su degradación ha quedado demostrada tanto en el caso de errores simples como con errores a ráfagas. Una vez vista esta degradación se ha evaluado el efecto del valor de MTU en cada caso.

A partir de los resultados obtenidos podemos extraer varias conclusiones:

1. En ausencia de errores (caso ideal), como es lógico, cuanto mayor es la MTU del enlace de comunicación (es decir, cuanto mayor es la proporción de información útil a información de cabecera), tanto mayor es el caudal alcanzado en la transferencia ya que, en este caso, la única penalización que sufre este parámetro es la debida a la transmisión de información no útil asociada a cabeceras y *trailers*.
2. En todos los casos se comprueba cómo, independientemente de la MTU del enlace, a medida que aumenta la tasa de error del canal el caudal disminuye ya que el número de retransmisiones aumenta y las pausas en la comunicación son cada vez más significativas.
3. El caudal alcanzado en una transferencia depende del compromiso entre tres factores fundamentales:
 - *Volumen de información de control*: Cuanto mayor es el tamaño de paquete empleado, tanto menor es la proporción de información no útil encapsulada en cada uno de los

paquetes que circulan por el enlace. Esto redundará en un mejor aprovechamiento del ancho de banda disponible en el canal.

- *Volumen de información retransmitida:* Los paquetes de tamaño reducido son menos sensibles a los errores. Teniendo en cuenta, además, que cada error fuerza la retransmisión de todo un paquete, el número total de bytes retransmitidos depende directamente del tamaño de paquete empleado.
- *Pausas en la comunicación:* En ocasiones, el canal permanece ocioso porque el módulo TCP emisor interrumpe la emisión de segmentos. Esto ocurre, por ejemplo, cuando se hace necesario recurrir a la expiración del temporizador de retransmisión para detectar las pérdidas o bien cuando TCP activa mecanismos, como el de *backoff* exponencial, para frenar la inyección de datos en la red con el objetivo de combatir lo que el protocolo interpreta como una posible situación de congestión.

También la reducción del tamaño de la ventana de congestión tras cada error, contribuye a ralentizar la transmisión, al provocar un peor aprovechamiento de la capacidad del canal. Este efecto, poco significativo en enlaces con un ancho de banda reducido, como el nuestro, en el que la ventana adopta valores muy pequeños, puede tener un gran impacto en redes con un producto ancho de banda * retardo elevado.

Cuanto mayor es el tamaño de paquete utilizado, mayores son los valores adoptados por la variable RTO, los cuales determinan el intervalo de tiempo que TCP debe esperar para detectar la pérdida de un paquete mediante la expiración del temporizador. Además los paquetes de gran tamaño frenan el crecimiento de la ventana de congestión de manera que el número de reconocimientos que TCP puede recibir tras un error es menor y, en consecuencia, menores son también las posibilidades de que tal error se detecte mediante el mecanismo de Retransmisión Rápida. En general, valores reducidos de la ventana del emisor fuerzan, con mayor frecuencia, la activación de mecanismos de control de la congestión como el *backoff* exponencial.

Hemos visto que, cuando los errores se producen de forma aislada, son los factores **IC** e **IR** los que inciden de forma más contundente en el caudal final alcanzado. No obstante, cuando los mecanismos generadores de errores son tales que dan lugar a ráfagas de error, entra en juego el tercero de estos factores, **ΔT** , convirtiéndose, a medida que las condiciones de error y el tamaño de paquete empleado aumentan, en el parámetro realmente decisivo.

4. Como resultado del compromiso entre los tres factores analizados existe siempre una MTU óptima cuya elección da lugar al caudal máximo.

Hemos comprobado cómo, a medida que el canal introduce más errores, el valor de esta MTU óptima tiende a reducirse debido a la menor vulnerabilidad a los errores de los paquetes de tamaño pequeño.

5. A partir de todo lo visto, podemos afirmar que la elección del tamaño de paquete que debe circular por el enlace de comunicación tiene una repercusión enorme sobre el rendimiento conseguido por TCP en su servicio de transmisión masiva de datos a través del canal móvil.

En las Tablas 7.1, 7.2 y 7.3 se comparan, para cada uno de los modelos de canal estudiados y para cada condición de error, el caudal alcanzado utilizando la MTU óptima con el caudal proporcionado por la MTU que da lugar al rendimiento más bajo. Los valores de MTU se indican, en todos los casos, en bytes y los de caudal en bytes/s.

BER	MTU óptima	Caudal máximo	MTU peor	Caudal mínimo	% mejora
0	1500	891	128	520	42%
1/100000	740	808	128	510	37%
1/75000	592	790	128	500	37%
1/50000	444	748	128	496	34%
1/25000	444	638	1332	264	58%
1/15000	296	525	740	20	96%

Tabla 7.1 Errores aislados

1/ μ	MTU óptima	Caudal máximo	MTU peor	Caudal mínimo	% mejora
100 ms	592	739	128	451	24%
250 ms	444	688	128	406	41%
500 ms	444	633	128	360	43%
750 ms	444	594	128	274	54%

Tabla 7.2 Modelo temporal

BER	MTU óptima	Caudal máximo	MTU peor	Caudal mínimo	% mejora
1/100000	592	798	128	483	39%
1/75000	444	707	128	432	39%
1/50000	444	681	128	406	40%
1/25000	444	624	1332	280	55%
1/15000	296	425	1332	17	96%

Tablas 7.3 Modelo de Fritchman

Resulta especialmente interesante observar cómo a medida que las condiciones de error del canal empeoran, la elección del tamaño de paquete se hace más crítica.

En este capítulo se ha analizado el comportamiento de TCP en entornos con errores. Las conclusiones obtenidas nos servirán, además, como base para el estudio del protocolo TCP tanto en entornos asimétricos con pérdidas, como en enlaces de baja velocidad en caso de que el canal introduzca errores, tal y como se detallará en los siguientes capítulos.