

DetECCIÓN DE CONGESTIÓN EN LA INTERNET EUROPEA

A. Hernández, E. Magaña, *Member IEEE*, M. Izal, *Member IEEE* y D. Morató, *Member IEEE*

Resumen—En este artículo se presenta un estudio sobre la utilización de las medidas de retardo unidireccional para detectar y caracterizar la congestión de la red en la Internet europea. Los experimentos se han hecho utilizando la plataforma ETOMIC que permite realizar medidas de retardo unidireccional con marcas temporales de alta precisión. Durante el estudio se ha encontrado un enrutador (*router*) con un comportamiento extraño en el que el cuello de botella no es el ancho de banda disponible sino la capacidad de procesamiento del enrutador. Este enrutador ha sido caracterizado mediante diversos parámetros de red. Algunos de ellos son la dependencia de la limitación con la tasa de datos de entrada en paquetes por segundo, el tamaño de las ráfagas de paquetes perdidos medido en paquetes o en tiempo y la ausencia de un algoritmo específico de planificación en el enrutador que podría afectar a los flujos más demandantes.

Palabras clave—Prestaciones en redes de ordenadores, Diagnóstico de fallos en redes, Efectos de retardo

I. INTRODUCCIÓN

En general, los enlaces de Internet que interconectan las distintas redes nacionales y de proveedores de servicio están correctamente dimensionados de manera que se intenta evitar situaciones de congestión o por lo menos que no se prolonguen en el tiempo. Sin embargo, la variedad de infraestructuras y administradores existentes hacen que nos encontremos a veces con segmentos de red que suponen un verdadero cuello de botella. A la hora de encontrar estos puntos conflictivos nos podemos fijar en diferentes parámetros de red como el *throughput*, las pérdidas o el retardo [1]. De todas ellas, en este trabajo vamos a utilizar las medidas de retardo extremo a extremo. El retardo es un parámetro de red importante porque a partir de un análisis del retardo podemos obtener información de topología de la red, congestión y cambios de rutas [2].

El retardo extremo a extremo se puede separar en cuatro componentes [2]: transmisión, propagación, procesamiento y de tiempo en cola. Los retardos de transmisión y propagación son constantes para determinado camino ya que dependen de la capacidad del enlace y de la distancia del mismo respectivamente. Sin embargo, los retardos de procesamiento

y de tiempo en cola son variables aleatorias debido a la variabilidad en el número de tareas de un enrutador (*router*) y en las condiciones de la red respectivamente. Es cierto que en las arquitecturas modernas de enrutadores se tiende a un retardo de procesamiento constante por paquete asistido por el hardware adecuado que es capaz de procesar a velocidad del enlace, por lo que el tiempo en cola será el único factor variable en la medida de retardo extremo a extremo. Éste será por tanto un factor a tener en cuenta en el análisis a realizar. Ambas componentes de transmisión y de propagación darán una buena aproximación al mínimo del retardo extremo a extremo, que además tendrá que ser de nuevo un valor constante para un camino dado [3].

Habitualmente el retardo extremo a extremo se aproxima con la mitad del *Round Trip Time* (RTT), debido a la sencillez de su obtención a partir de la aplicación *Ping* que usa paquetes *ICMP Echo Request/Reply* y que corre en la máquina origen del experimento. Sin embargo, esta estimación no es siempre válida debido a la asimetría de los caminos. Por un lado, los caminos de ida y vuelta pueden atravesar enrutadores y enlaces diferentes debido a la naturaleza del protocolo de nivel de red IP. Por otro lado, aunque el camino sea el mismo, podemos encontrarnos condiciones de red diferentes, por ejemplo, según el grado de congestión de los enlaces [4]. Por tanto, para evitar estos efectos será de interés medir el retardo extremo a extremo en un sólo sentido (*One-Way Delay*, OWD), de manera que podamos tener caracterizado el retardo en cada uno de los sentidos independientemente.

La medida del OWD es mucho más compleja que la del RTT [5]. Exige ser capaz de enviar un paquete con una marca temporal (*timestamp*) en el origen y compararlo con el *timestamp* en el destino, es decir, necesita controlar ambos extremos del camino. Además, para que el *timestamp* sea significativo es necesario que los relojes de ambos extremos estén sincronizados. Esta sincronización requiere contar con equipos GPS (*Global Positioning System*) que ofrezcan una señal de referencia de reloj a las tarjetas de red directamente.

En este trabajo hemos usado la plataforma ETOMIC (*European Traffic Observatory Measurement Infrastructure* [6][7]) para realizar las medidas. Se trata de una plataforma de monitorización activa compuesta por 18 nodos distribuidos por Europa, cada nodo dotado de una tarjeta de red especial de monitorización (DAG Endace 3.6GE) y de sincronización GPS. Esta sincronización se consigue a partir de la señal PPS (*Pulse per Second*) generada por el GPS y ofrecida directamente a la tarjeta de red especial de monitorización. Esta plataforma permite medidas de alta precisión debido a que el *timestamp* lo inserta la propia tarjeta de red emisora,

Este trabajo ha sido financiado dentro del Proyecto Integrado Evergrow (No 001935) del Programa FP6/IST/FET de la Comisión Europea y del Proyecto del Plan Nacional PINTA TEC2004-06437-C05-03.

A. Hernández (correo-e.: ana.hernandez@unavarra.es), E. Magaña (correo-e.: eduardo.magana@unavarra.es), M. Izal (correo-e.: mikel.izal@unavarra.es) y D. Morató (correo-e.: daniel.morato@unavarra.es) pertenecen al Departamento de Automática y Computación de la Universidad Pública de Navarra, Campus Arrosadía, 31006 Pamplona, España.

evitando efectos del sistema operativo presente en los nodos (GNU Linux). La precisión que se consigue con ETOMIC está en el orden de centenares de nanosegundos. En cuanto a su ubicación, los nodos se encuentran conectados a redes de universidades, centros de investigación y grandes empresas. Por tanto, podrán dar una idea de la conectividad dentro de la red de alta velocidad europea Geant2. En concreto para el estudio se han utilizado entre otros los nodos que aparecen en la tabla I.

TABLA I
LOCALIZACIÓN DE ALGUNOS NODOS ETOMIC UTILIZADOS EN EL ESTUDIO

Nodo	País	Ciudad	Centro
colbud	Hungría	Budapest	Collegium Budapest
elte	Hungría	Budapest	Eotvos Lorand University
ericsson	Suecia	Estocolmo	Ericsson Research Center
Paris	Francia	Paris	Paris-Sub11

Mediante las medidas de OWD podremos detectar posibles puntos de congestión o cuellos de botella en la red europea que une los nodos de la plataforma ETOMIC. Caracterizaremos esa limitación y la dependencia de parámetros como paquetes por segundo o tiempo entre paquetes.

El trabajo se estructura de la siguiente forma. En la siguiente sección empezaremos utilizando el OWD para la detección de congestión, seguido de la identificación del punto conflictivo en nuestro caso un enrutador. Se comprobará que el problema se encuentra en la limitación hardware de un enrutador cuyo comportamiento se caracterizará en las siguientes secciones a nivel de pérdidas y justicia del algoritmo de planificación. Finalmente se presentarán las conclusiones del trabajo.

II. USANDO ONE-WAY DELAY PARA LA DETECCIÓN DE CONGESTIÓN

Para observar la variación del OWD se realizan experimentos de medida del OWD formados por 24.000 paquetes UDP de 100 bytes (tamaño a nivel IP mientras no se diga lo contrario) con tiempo entre paquetes constante para cada velocidad, entre cada par de nodos de la plataforma ETOMIC. Las tarjetas de red de alta precisión Endace DAG 3.6GE de los nodos incorporan un procesador y con ello capacidad de programar la generación de estas ráfagas de paquetes en la propia tarjeta, con lo que la precisión del tiempo entre paquetes conseguida es elevada. Para poder medir el OWD, en cada paquete enviado se incrusta un *timestamp* de envío que restado del *timestamp* del momento en el que se recibe en el destino nos dará el OWD, siempre que ambos extremos están sincronizados vía GPS. El *timestamp* de envío y recepción los inserta la tarjeta DAG lo más cerca posible de la red consiguiendo una medida de alta precisión y sin interferencias del sistema operativo de propósito general.

Se observan dos comportamientos. El primero, el más habitual, que aparece en la Fig. 1 en la que se representa el histograma normalizado (función densidad de probabilidad) del OWD para distintas velocidades, obtenido entre los nodos

de *colbud* y *paris*. En la figura no se detectan variaciones significativas del OWD aunque se comparen velocidades de hasta 90 Mbps. La media se mantiene y la variabilidad sobre la media también, debido a que los experimentos a diferentes velocidades se realizan consecutivamente y todos ellos encuentran el mismo tráfico existente sobre la red. Este comportamiento es debido a que la troncal que une los nodos (Geant2 [8]) es gigabit y 10-gigabit, y no estaremos introduciendo congestión en ningún momento a velocidades inferiores a 100 Mbps. Los nodos de la plataforma ETOMIC están conectados en redes de acceso a 100 Mbps por lo que no podemos comprobar mayores velocidades.

Si el mismo experimento se realiza en otro momento obtendremos un perfil que puede variar según el tráfico pre-existente sobre la red, suponiendo que no se producen cambios de rutas o en la topología de la red. Esta suposición de estacionariedad será válida en la escala de algunas horas [1].

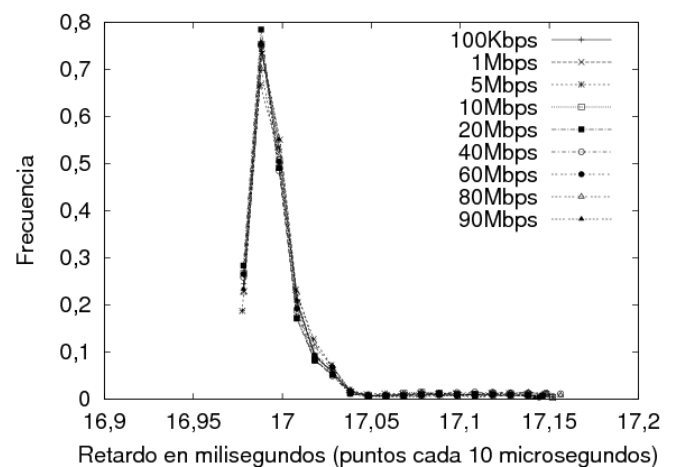


Fig. 1. OWD en un camino sin congestión

Todos los OWD de todos los caminos entre los nodos ETOMIC siguen un patrón similar al de la Fig.1 variando la media del OWD y la anchura de la distribución según la pareja de nodos, a excepción de las medidas realizadas con origen o destino en el nodo *ericsson* que siguen el patrón de la Fig.2. En ella de nuevo tenemos el histograma normalizado del OWD, obtenido para medidas realizadas desde el nodo de *colbud* al nodo de *ericsson*. Se observa como al aumentar la velocidad se mueve la media del OWD y se ensancha la distribución. Éste es un claro indicador de que en un enlace cercano en el camino al nodo *ericsson* (y común en el camino desde otros nodos a *ericsson*) se alcanza estado de congestión en torno a la velocidad de 16-17 Mbps según las condiciones de la red.

Para ambas Fig. 1 y 2, el mínimo de OWD coincide para diferentes velocidades. Se trata de un valor fijado por el tiempo de transmisión, tiempo de propagación y tiempo de procesamiento, y por tanto constante para cada camino entre cada par de nodos, suponiendo constante el tiempo de procesamiento en arquitecturas de enrutadores actuales (o por

lo menos con una variación despreciable que se comprobará que no es cierto para la Fig. 2). Los paquetes que sufren este retardo mínimo no experimentan retardo por tiempo en cola, es decir, cuando llegan a la cola del interfaz de entrada de un enrutador no se van a encontrar ningún otro paquete en cola del mismo o diferente flujo, y por tanto pasan a ser atendidos inmediatamente por el enrutador. Esta variación del OWD respecto al mínimo es causada por tanto única y exclusivamente por el tiempo en cola en la Fig.1. Por tanto, esta medida se puede utilizar para caracterizar el tiempo en cola de los enrutadores en el camino entre dos nodos que podemos encontrar en estudios sobre tomografía de red [9].

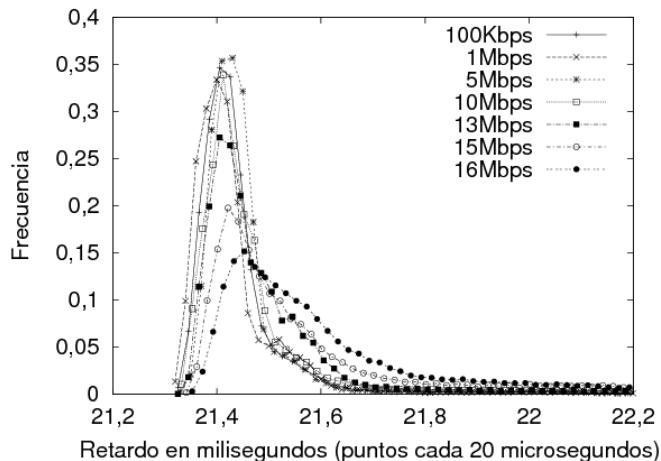


Fig.2. OWD en un camino con congestión

El máximo del OWD también es un valor constante que vendrá dado por los tamaños de las colas de los enrutadores en el camino bajo estudio. El peor caso con máximo OWD será aquel en el que todas las colas de los enrutadores se encuentren casi llenas (con espacio sólo para el nuevo paquete que llega). Se trata de una situación altamente improbable, debido a que sería una situación de congestión extrema a partir de la cual se perderían paquetes y que no es el punto de trabajo habitual de los enrutadores.

Si bien en general se ha comprobado la alta capacidad aún con tráfico interferente ya existente sobre el camino medido entre nodos conectados a Geant2, también se ha comprobado la existencia de un punto conflictivo cerca del nodo *ericsson* a relativa baja velocidad y que requerirá de un estudio más detallado en las próximas secciones. El OWD será un parámetro que se ve afectado por situaciones de congestión y por tanto nos podrá ayudar a detectarlas. En concreto, la variación del OWD se debe en principio a variaciones del tiempo en cola (el resto de componentes se pueden considerar constantes como se comentó anteriormente), el cual está relacionado directamente con la congestión en el enlace de salida del enrutador [10]. Sin embargo, podremos comprobar en próximas secciones como parte del retardo puede deberse a limitaciones de capacidad de procesamiento de los enrutadores.

III. IDENTIFICACIÓN DEL ORIGEN DE LA LIMITACIÓN

Se puede pensar que el origen de la variación en el OWD puede ser un enlace cuello de botella en el camino a *ericsson*. Para identificarlo, se realiza un *traceroute* para evaluar los enrutadores atravesados en el camino de *colbud* a *ericsson*. Una vez conocida esta lista de enrutadores que se atraviesan se trata de comprobar el retardo existente entre el origen *colbud* con cada uno de los enrutadores en el camino a *ericsson* mediante la herramienta *traceroute* en situación normal y generando un tráfico que dé síntomas de congestión en el OWD. Para la medición del retardo (esta vez RTT) se utiliza la propia estimación de tiempo del *traceroute* lanzado contra cada enrutador del camino al destino directamente, fijando el TTL de la medida, para poder realizar mediciones contra varios enrutadores del camino simultáneamente. Esta medida de retardo también se podría haber realizado utilizando *pings* pero se comprueba que existen algunos enrutadores que no contestan al *ICMP Echo Request* del *ping* cuando sí devuelven un *ICMP Destination Unreachable* por tiempo excedido en tránsito debido al *traceroute*, por configuración de reglas de filtrado. El *traceroute* es una herramienta de diagnóstico muy útil para los administradores de red por lo que normalmente está activada su respuesta en los enrutadores, sin embargo, podemos encontrarnos enrutadores que tampoco contesten a estos paquetes.

En concreto, para tener congestión se introduce un tráfico UDP con paquetes de 100 bytes a 17 Mbps durante varios segundos. De esta forma, tendremos el RTT para cada enrutador medido con y sin congestión. Detectando el primer enrutador en el que crece el RTT en situación de congestión respecto a la normal, habremos identificado el enrutador cuyo enlace de salida es el cuello de botella con el destino. Esto es posible porque el enrutador del cuello de botella estará encolando paquetes y por tanto se retarda la atención a todos los paquetes incluidos los de *traceroute*. En la Fig. 3 se observa el RTT con y sin congestión para cada enrutador en el camino entre *colbud* y *ericsson*. Se comprueba cómo el punto conflictivo se encuentra en el enrutador *ericsson-107599.k.se.telia.net* (213.65.55.238) (enrutador 14 en la figura), situado en Linkping (Suecia) si se utilizan herramientas de localización de IPs [11]. Para el resto de enrutadores se obtienen medidas orientativas del RTT porque el error cometido en la medida está en el orden de magnitud de las diferencias entre la mayoría de enrutadores por lo que no se ven diferencias importantes para la Fig. 3. Sólo hay un fuerte crecimiento del retardo en los enrutadores 7 y 8 que está indicando enlaces largos en distancia con elevado tiempo de propagación. Notar que el retardo con y sin congestión es el mismo para enrutadores por debajo del conflictivo porque ellos no se ven afectados.

En ese punto del enrutador 14 se observa cómo se produce un crecimiento fuerte del RTT, obteniendo para este enrutador un RTT mayor de 3,7 sg que sale fuera de la gráfica. Este valor tan elevado del RTT nos está indicando que no se trata del aumento de retardo que sufren todos los paquetes por el tiempo en cola que tienen que esperar antes de ser servidos. Se

debe a que la respuesta a nuestro paquete de *traceroute* es una tarea de baja prioridad para el enrutador que hace que se posponga su envío hasta que el enrutador deje de estar en situación de congestión. Es entonces cuando el enrutador vuelve a tener tiempo para atender tareas que no son propias del enrutamiento básico que debe priorizar. Por eso el tiempo tan elevado del RTT no puede deberse a retardo en cola y sí a posponerse la respuesta a la finalización del flujo de alta velocidad. Los enrutadores que vienen a continuación en el camino a *ericsson* (15 y 16 de la Fig. 3) también tienen un RTT que ha crecido pero en órdenes de magnitud del tiempo en cola, ya que este tiempo ha sido heredado del tramo hasta el enrutador 14. Los paquetes que llegan a los enrutadores 15 y 16 vienen retrasados al tener que esperar un tiempo en cola importante en el enrutador 14 hasta ser enviados al siguiente enrutador.

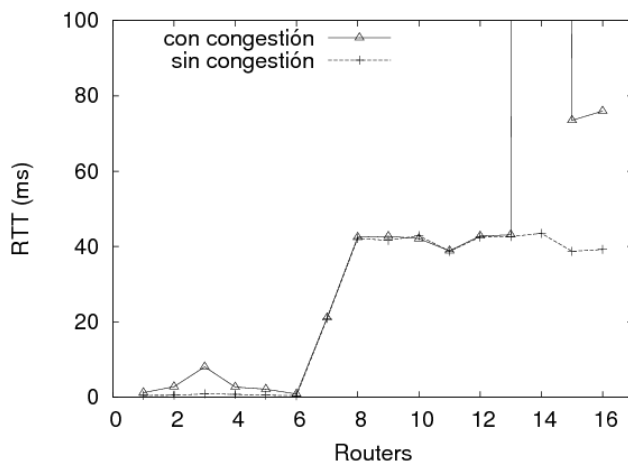


Fig. 3. RTT del *traceroute* en cada uno de los saltos en el camino entre los nodos *colbud* y *ericsson* sin y con tráfico generado simultáneamente que le lleve a situación de congestión

El hecho de posponer la respuesta al *traceroute* en ese enrutador es un claro indicador de que el enrutador no tiene capacidad de procesamiento suficiente para atender todas las tareas, posponiendo las tareas secundarias. Por tanto, se puede intuir que el punto de congestión no es debido al cuello de botella de un enlace sino más bien debido a la capacidad de procesamiento del enrutador. En la siguiente subsección trataremos de comprobar esta hipótesis.

A. Efecto sobre el OWD y pérdidas

En situación de congestión, la serie temporal del OWD medido con 48.000 paquetes UDP es el que se observa en la Fig. 4. El experimento se ha realizado manteniendo constante la tasa de paquetes por segundo en transmisión, variando el tamaño de paquete para 50 y 100 bytes, resultando velocidades de transmisión de 8,5 y 17 Mbps. En esta figura se observa una primera parte con pendiente ascendente durante la que se empiezan a encolar cada vez más paquetes en la cola de entrada del enrutador (no en la cola de salida como podría pensarse en un primer momento y que luego justificaremos) hasta llegar a un punto en el que esa cola se

llena por completo. Entonces todos los paquetes ya no pueden acomodarse en la cola por lo que aparecerán pérdidas de paquetes. Por tanto, el retardo máximo de OWD se encuentra acotado por este tamaño de cola de entrada del enrutador.

Si la limitación fuera por la capacidad del enlace de salida del enrutador, a diferentes velocidades como las mostradas en la Fig. 4 se debería obtener diferente pendiente en la primera parte del retardo. En concreto, para el caso de paquetes de 50 bytes no debería producirse crecimiento del retardo porque la velocidad de entrada 8,5 Mbps sería menor que el punto de congestión de 17 Mbps encontrado en el apartado anterior. Es decir no habría congestión. Sin embargo, hay congestión en ambos casos, a 8,5 y 17 Mbps. La pendiente es la misma porque es constante la tasa de paquetes transmitidos, con lo que se demuestra que la limitación se debe a la capacidad de procesamiento en paquetes por segundo del enrutador [12]. Por tanto, el punto de congestión no es debido al cuello de botella de un enlace sino más bien debido a la capacidad de procesamiento del enrutador, es decir, de la capacidad de reenvío de paquetes por segundo de la CPU o *backplane* del enrutador. Por tanto, el retardo en cola observado es de la cola de entrada y no de salida del enrutador. Si la limitación estuviera en la capacidad del enlace de salida sí entraría en juego el tiempo en la cola de salida, pero no es el caso.

En la Fig. 4 se observa que el OWD crece a valores mayores para tamaño de paquete pequeño debido a que se pueden encolar más paquetes para un mismo tamaño de la cola de entrada al enrutador. Se puede observar que la cola se llena durante más tiempo. Este tiempo no es el doble (aunque sea la mitad de tamaño de paquete) porque el tamaño útil de buffer no es directamente proporcional al tamaño de paquete, debido a cómo se indexan los paquetes en la arquitectura interna del enrutador.

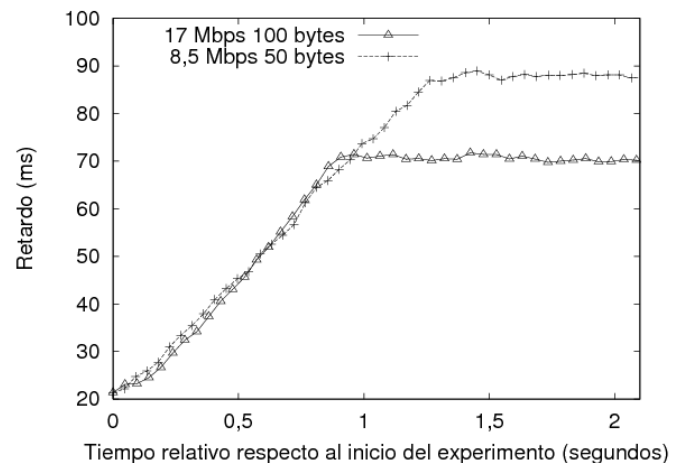


Fig. 4. OWD en situación de congestión en el camino de *colbud* a *ericsson*

Se ha encontrado por tanto un punto de congestión debido a limitaciones hardware del equipo de enrutamiento, lo cual deja abiertas preguntas de interés como si esta limitación de paquetes por segundo de capacidad del enrutador es constante, cómo es el patrón de pérdidas de paquetes o el esquema de

planificación entre flujos del enrutador.

La serie temporal de la pérdida de paquetes acumulada asociada a la Fig. 4 se muestra en la Fig. 5. En la primera parte no hay pérdidas porque los paquetes se van encolando en el buffer de entrada del enrutador (crecimiento del OWD de la Fig. 4), pero tan pronto como se llena esta cola se empiezan a perder los paquetes que no caben. Como los paquetes llegan a tasa constante idéntica para cada tamaño de paquete y también es constante la tasa de paquetes por segundo que es capaz de procesar el enrutador para un tamaño de paquete determinado, la diferencia entre ambos fija la pendiente de los paquetes perdidos que es lineal con el tiempo.

Las pendientes de las dos líneas para distinto tamaño de paquete de la Fig. 5 no son iguales debido al diferente coste de transmisión que le supone al enrutador el envío de un paquete según su tamaño. Cuando haya paquetes encolados podrá enviar más rápidamente (menor tiempo de transmisión) los paquetes más pequeños, consiguiéndose menos tasa de pérdidas respecto al caso del tamaño de paquete más grande. Es decir, se consigue una mayor tasa efectiva de transmisión en paquetes por segundo con el tamaño de paquete pequeño, como era de esperar. Este fenómeno se comprobará en el apartado siguiente haciendo un estudio de la máxima tasa de paquetes recibidos en el destino en función del tamaño de paquete.

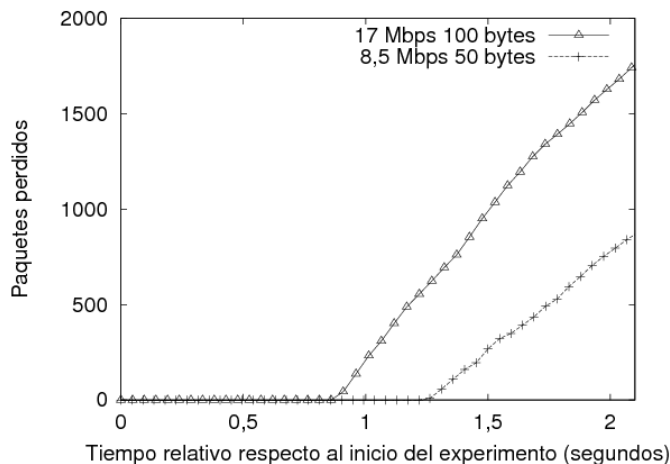


Fig. 5. Pérdidas acumuladas en situación de congestión en el camino de *colbud a ericsson*

Se ha comprobado que la congestión detectada en un primer momento se debe en verdad a la limitación en la capacidad de procesamiento del enrutador. Hacer notar que estamos apurando las prestaciones del enrutador por ser paquetes pequeños (100 bytes a 17 Mbps). Las transferencias de datos habituales si pretenden optimizar su tasa de transferencia harán uso en la práctica de tamaños de paquete limitados por la MTU, habitualmente de 1500 bytes en tecnologías Ethernet. La limitación de velocidad del enrutador en paquetes de 1500 bytes sería bastante superior a la considerada para 100 bytes.

IV. CAPACIDAD DE PROCESAMIENTO DEL ENRUTADOR

A la hora de analizar la capacidad de procesamiento del enrutador bajo estudio, será de interés comprobar si existe un límite constante en el número de paquetes por segundo que sea capaz de procesar este enrutador y reenviar al destino. En la Fig. 6 se presenta los paquetes recibidos en función del número de paquetes por segundo transmitidos desde el origen, para experimentos de 24.000 paquetes UDP de 100 bytes y diferentes tasas de envío (en paquetes por segundo).

En la Fig. 6 se comprueba cómo para bajo número de paquetes por segundo el enrutador es capaz de procesar todos los paquetes y reenviarlos al destino. El enrutador saca tantos paquetes por segundo como recibe, observándose la pendiente de la figura. Sin embargo, al llegar a la capacidad máxima del enrutador (cerca de 20.000 paquetes por segundo), no sólo no se mantiene a esa capacidad máxima sino que baja. Esta bajada en la capacidad de procesamiento del enrutador conforme aumenta la tasa de paquetes de entrada es debida a que el tirar paquetes también supone un coste para el enrutador. El paquete es leído por el interfaz de entrada del enrutador y aunque no se almacene en la cola, tendrá que decidir tirarlo, con lo que se dedica un tiempo de procesamiento a cada paquete tirado. Por tanto, conforme crezca la tasa de transmisión también crecerá la necesidad de tirar paquetes, observándose la caída en la tasa de paquetes procesados de la figura. El resultado muestra los paquetes recibidos para todo el camino que atraviesa varios enrutadores, pero el efecto predominante es el del enrutador bajo estudio.

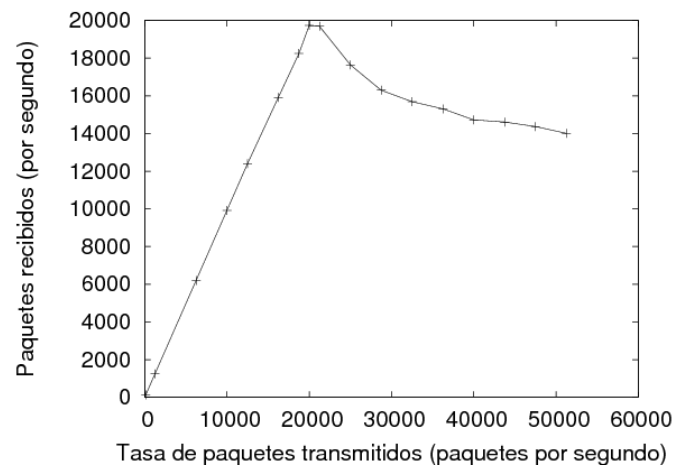


Fig. 6. Paquetes reenviados por el enrutador bajo estudio en función del número de paquetes por segundo transmitidos para tamaño de paquete de 100 bytes

Para comprobar el efecto del tamaño de paquete, en la Fig. 7 se presenta la máxima tasa de paquetes procesados conseguida para cada tamaño de paquete. Es decir, se representan los puntos de pico de la Fig. 6 pero para varios tamaños de paquete porque ya se ha visto en esa misma figura que las tasas máximas se obtienen alrededor del punto de congestión. Los experimentos se han realizado de nuevo con

24.000 paquetes UDP entre *colbud* y *ericsson*. En la Fig. 7 se observa como conforme aumenta el tamaño de paquete esta tasa máxima de paquetes recibidos baja debido al mayor coste de reenvío de paquetes más grandes. Este coste será mayor por un lado debido a la necesidad de manejar mayor número de bytes en los buses y copias internas del enrutador. Por otro lado, el mayor coste se deberá al mayor tiempo de transmisión que deberán soportar los paquetes más grandes a una tasa de salida constante en bits por segundo. De esta forma también se justifica la diferente pendiente de las curvas de la Fig. 5 ya revisada en la que teníamos diferentes pendientes en la tasa de pérdidas según el tamaño de paquete. Diferentes tasas efectivas de reenvío según el tamaño de paquete implicarán también diferente tasa de pérdidas según el tamaño de paquete.

Resumiendo, la limitación del enrutador es debido a su capacidad de procesamiento en paquetes por segundo, pero esta limitación no es constante y depende de la tasa y del tamaño de paquetes a su entrada.

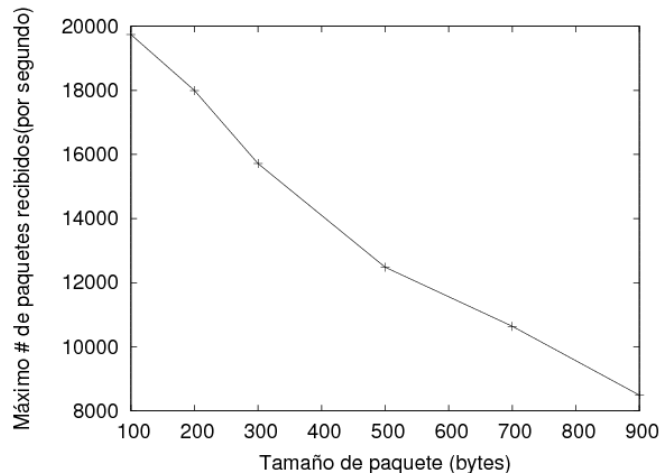


Fig. 7. Máxima tasa de paquetes recibidos en función del tamaño de paquete transmitido para los puntos donde empieza a producirse la congestión

V. CARACTERIZACIÓN DE LAS PÉRDIDAS

Debido a las peculiaridades de este enrutador en el que la capacidad está limitada por su capacidad de procesamiento y no por el enlace de salida que le corresponde, el patrón de pérdidas generado resulta de interés. Para ello, realizamos un experimento de 24.000 paquetes UDP de 100 bytes con cada velocidad (es decir, diferente tiempo entre paquetes) con destino *ericsson*, obteniendo el resultado mostrado en la Fig. 8. En la figura se muestra el histograma del número de paquetes consecutivos perdidos para cada velocidad, es decir, el tamaño de ráfagas de paquetes perdidos que se producen en el escenario. Se observa como conforme aumenta la velocidad aumenta el número de paquetes perdidos consecutivos dentro de cada ráfaga, ensanchándose las distribuciones. Para 17 Mbps casi todas las pérdidas son de paquetes sueltos y sólo un pequeño porcentaje corresponde a pérdidas de 2 paquetes consecutivos. El número de paquetes consecutivos perdidos crece conforme aumenta la velocidad, es decir, la tasa de paquetes por segundo a la entrada del enrutador.

Considerando el resto del tráfico que circula por ese enrutador como tráfico interferente, se puede sospechar que el enrutador tira paquetes a intervalos de tiempo dependientes de este tráfico interferente, por lo que a mayor tasa de paquetes por segundo las ráfagas de pérdidas contienen mayor número de paquetes perdidos.

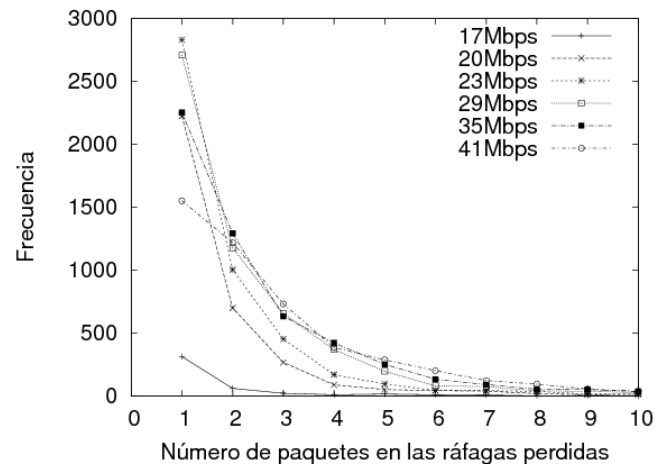


Fig. 8. Histograma del tamaño de las ráfagas de paquetes perdidos en número de paquetes

En la Fig. 9 se representa la función de supervivencia de la duración de las ráfagas de paquetes perdidos, contabilizando el tiempo entre el último paquete recibido anterior a la ráfaga perdida y el primer paquete tras la ráfaga perdida. Existen ráfagas de duración mayor a las que se presentan en la figura pero con muy pocas ocurrencias por lo que no es representativa su función de supervivencia y por tanto se descartan para el estudio. Las líneas obtenidas en la figura son similares para todas las velocidades si descartamos las de 17 y 20 Mbps que están muy próximas al punto de congestión. Para las velocidades en congestión viene a indicar un esquema de planificación de pérdida de paquetes en el enrutador que tira los paquetes que llegan durante un intervalo de tiempo que no depende de la velocidad del tráfico inyectado y dependerá de la tipología del tráfico interferente en el enrutador. Los intervalos de tiempo durante los cuales se pierden paquetes son próximos en todos los casos.

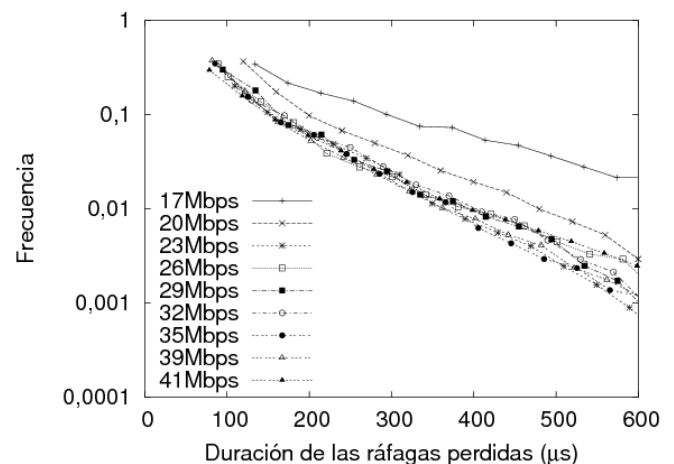


Fig. 9. Función de supervivencia de la duración de las ráfagas de paquetes perdidos.

Con tráfico interferente similar y por tanto la misma distribución temporal de pérdidas, el número de paquetes que compone cada ráfaga de pérdidas dependerá de la tasa de envío desde el origen. Por eso, en la Fig. 8 salía diferente número de paquetes perdidos por ráfaga según la velocidad de envío.

VI. JUSTICIA EN EL ESQUEMA DE PLANIFICACIÓN DEL ENRUTADOR

Por los análisis anteriores puede parecer que el enrutador esté aplicando un esquema de planificación que esté perjudicando, mediante la pérdida de sus paquetes, en mayor medida al flujo inyectado de alta velocidad. Sin embargo comprobamos que no es así mediante la generación de 2 flujos desde diferentes orígenes con el mismo destino *ericsson*, uno de baja velocidad a 5Mbps desde el nodos de *elte* y otro de mayor velocidad desde el nodo de *colbud*. De nuevo, los flujos consisten en paquetes equiespaciados UDP de 100 bytes con tiempo entre paquetes marcado por la velocidad en cuestión. Para conseguir que ambos flujos lleguen al mismo tiempo al enrutador estudiado, al ser los retardos desde el nodo origen al enrutador distintos para cada uno de los caminos, prolongamos el flujo de menor velocidad enviado desde *elte* en el tiempo (más paquetes enviados) y retrasamos el envío del flujo desde *colbud*. Con ello aseguramos que los 24.000 paquetes del flujo desde *colbud* lleguen al enrutador a la vez que los provenientes de *elte*, quedándonos con el intervalo de tiempo en el que coinciden ambos flujos. En la tabla II se comprueba que el efecto de pérdida de paquetes es proporcional en ambos flujos, obteniéndose aproximadamente el mismo porcentaje de paquetes recibidos con éxito en destino.

TABLA II
PORCENTAJE DE PAQUETES RECIBIDOS EN DESTINO PARA DISTINTAS COMBINACIONES DE VELOCIDADES DE 2 FLUJOS SIMULTÁNEOS

Velocidad colbud (Mbps)	Velocidad elte (Mbps)	Porcentaje colbud	Porcentaje elte
15	5	65,19 %	65,77 %
20	5	46,74 %	46,74 %
25	5	39,07 %	38,59 %

De esta forma se ha comprobado que todo el tráfico que atraviesa el enrutador se encuentra en igualdad de condiciones en cuanto a la planificación y por tanto las pérdidas son proporcionales a la velocidad de cada flujo. No se aplica ninguna disciplina de conformación de tráfico o calidad de servicio más allá de la FIFO en la cola de entrada del enrutador, que pudiera estar falseando nuestras medidas.

VII. CONCLUSIONES

El retardo en un sólo sentido es útil para detectar situaciones de congestión en un camino concreto al destino, sin interferencias en la medida causadas por el camino de vuelta presentes en las medidas clásicas por RTT. El poseer

una plataforma de medida de alta precisión con sincronización GPS ha permitido un estudio en detalle de las variaciones del OWD en función de parámetros como la velocidad o el tamaño de paquete. Dentro del estudio realizado, se ha detectado un punto conflictivo con limitación próxima a los 17 Mbps cercano al nodo *ericsson*. Es un punto aislado, cerca de la red de acceso del nodo *ericsson* y por tanto no representativo de la alta capacidad de la red europea Geant2 para el resto de combinaciones de caminos.

El análisis realizado ha ofrecido resultados interesantes en cuanto a que la limitación encontrada en el camino a *ericsson* no es debida a la capacidad física de los enlaces sino a las capacidades hardware de un enrutador que hemos podido localizar, en este caso haciendo uso del RTT. Se ha estudiado el OWD obtenido según diferente tamaño de paquete y cómo la limitación viene por la capacidad de procesamiento del enrutador. Este comportamiento es variable porque al aumentar la velocidad (el número de paquetes por segundo a tamaño constante) disminuye el valor límite de capacidad de procesamiento en paquetes por segundo del enrutador debido al coste que le supone al enrutador tirar paquetes. El enrutador no puede dejar de leer paquetes que le llegan por un interfaz, los debe leer, cargarlos en memoria y luego decidir qué hacer con ellos, lo que supone un coste.

También se ha verificado cómo las pérdidas se producen a ráfagas cuyo tamaño es mayor en distribución conforme aumentamos la velocidad de transmisión del flujo (variando el tamaño de los paquetes) o se reduce el tiempo entre paquetes (con tamaño de paquete constante). En todo caso, estas pérdidas dependen del tráfico interferente que ya exista sobre la red, y que hace que se produzcan pérdidas durante tiempos establecidos por este tráfico interferente, y por tanto común para cualquier combinación de velocidad del flujo de los experimentos.

Se ha comprobado para el enrutador bajo estudio cómo las pérdidas afectan por igual en porcentaje a todos los flujos establecidos independientemente de su velocidad, es decir, que no incorpora ningún mecanismo de calidad de servicio para limitar los flujos más demandantes.

VIII. REFERENCIAS

- [1] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker. "On the constancy of internet path properties". In *ACM SIGCOMM Internet Measurement Workshop (IMW '2001)*, San Francisco, California, USA, November 2001.
- [2] C. J. Bovy, H. T. Mertodimedjo, G. Hooghiemstra, H. Uijtervaal, and P. Van Mieghem. "Analysis of end-to-end delay measurements in Internet". In *Passive and Active Measurement Conference (PAM 2002)*, Fort Collins, CO, USA, March 2002.
- [3] N. Hu and P. Steenkiste. "Quantifying internet end-to-end route similarity". In *Passive and Active Measurement Conference (PAM 2006)*, Adelaide, Australia, March 2006.
- [4] D. Constantinescu and A. Popescu. "Modeling of one-way transit time in IP routers". In *Advanced Int'l Conference on Telecommunications and Int'l Conference on Internet and Web Applications and Services AICT-ICIW '06*, February 2006.
- [5] T. Iwama, A. Kaneko, A. Machizawa, and H. Toriyama. "Real-time measurement of one-way delay in the internet environment". *The Institute of Electronics, Information and Communication Engineers*, 2004(B-16-1), 2004.

- [6] E. Magaña, D. Morató, M. Izal, J. Aracil, F. Naranjo, F. Astiz, U. Alonso, and et al. "The European traffic observatory measurement infrastructure (ETOMIC)". In *IEEE International Workshop on IP Operations & Management (IPOM 2004)*, Beijing, China, October 2004.
- [7] G. Vattay and E. Magaña. "European Traffic Observatory Measurement Infrastructure (ETOMIC) web page". <http://www.etomic.org>. Collegium Budapest and Public University of Navarra, 2007.
- [8] K. Ullman. "The GEANT network". <http://www.geant.net>. Dante, Cambridge, 2007.
- [9] G. Simon, P. Hága, G. Vattay, and I. Csabai. "A flexible tomography approach for queueing delay distribution inference in communication networks". In *Proceedings of Internet Performance, Simulation, Monitoring and Measurement (IPS-MoMe 2005)*, Warsaw, Poland, March 2005.
- [10] M. Garetto and D. Towsley. "Modeling, simulation and measurements of queueing delay under long-tail internet traffic". In *SIGMETRICS 2003*, San Diego, USA, June 2003.
- [11] A. McElligott. "Geobytes Inc. IP address locator tool". <http://www.geobytes.com/iplocator.htm>. Geobytes, 2007.
- [12] M. Paredes-Farrera, M. Fleury, and M. Ghanbari. "Router response to traffic at a bottleneck link". In *2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, TRIDENTCOM 2006*, March 2006.

BIOGRAFÍAS



Ana Hernández es Ingeniera de Telecomunicación por la Universidad Pública de Navarra, Pamplona, España, en 2006.

En la actualidad es contratada de investigación y se encuentra realizando estudios de master y doctorado. Su principal tema de interés en investigación son las medidas activas para monitorización de red.



Eduardo Magaña (M'00) es Ingeniero y Doctor Ingeniero de Telecomunicación por la Universidad Pública de Navarra, Pamplona, España, en 1998 y 2001 respectivamente.

En la actualidad es profesor titular en la Universidad Pública de Navarra. Durante el año 2002 trabajó como investigador postdoctoral en el Department of Electrical Engineering and Computer Science de la Universidad de California en Berkeley. Sus principales temas de interés en investigación son el análisis de tráfico y el estudio de eficiencia en redes de conmutación óptica de ráfagas.



Mikel Izal (M'00) es Doctor Ingeniero de Telecomunicación por la Universidad Pública de Navarra en 2002.

Durante 2003 realiza estancia postdoctoral como visitante científico en el Institute Eurecom, Sophia-Antipolis, Francia. Profesor Titular de Universidad en la Universidad Pública de Navarra desde 2006. Desde 1997 forma parte del Grupo de investigación de Redes, Sistemas y Servicios Telemáticos de la UPNA, ha participado en proyectos del Plan Nacional (Español) de I+D así como en proyectos integrados (proyecto EVERGROW) y en redes de excelencia (ePhoton-ONE) del VI Programa Marco de la Unión Europea. Sus intereses de investigación actuales incluyen el análisis de tráfico y prestaciones, especialmente en redes ópticas de nueva generación y sistemas peer-to-peer.



Daniel Morato (M'00) obtuvo el título de Ingeniero de Telecomunicación y de Doctor en Ingeniería de Telecomunicación en 1997 y 2001 respectivamente.

Durante el año 2002 trabajó como visitante postdoctoral en el Departamento EECS de la Universidad de California en Berkeley. Hoy en día es Profesor Titular de Universidad en el Departamento de Automática y Computación de la Universidad Pública de Navarra. Sus temas de interés incluyen el análisis de tráfico y prestaciones de servicios de Internet y redes de alta velocidad.