Universidad Pública
de Navarra
Nafarroako
Unibertsitate Publikoa

E.T.S. de Ingenieros Industriales y de Telecomunicación
Departamento de Automática y Computación

Tesis Doctoral:

Aproximación al modelado y predicción de tráfico de Internet como múltiplex de conexiones de transporte

Autor: Daniel Morató Osés
Ingeniero de Telecomunicación
Director: Javier Aracil Rico
Dr. Ingeniero de Telecomunicación

Índice General

Índice de Figuras

Índice de Tablas

Resumen

El tráfico de datos en la Internet actual presenta un nuevo reto de caracterización y modelado para el correcto dimensionamiento de los equipos y enlaces que conforman la llamada "Red de Redes". En este trabajo presentamos una revisión de los modelos propuestos hasta la fecha, lo cual nos lleva desde los símiles con la telefonía clásica hasta los conceptos de *dependencia a largo plazo* y *autosimilitud*.

A partir de estos modelos abordamos la caracterización de una gran población de usuarios de Internet. Para ello nos hemos basado en trazas del tráfico del enlace IP sobre ATM para acceso a Internet de la Universidad Pública de Navarra. Dichas trazas han sido obtenidas mediante una novedosa herramienta de monitorización de enlaces ATM. Con estas trazas presentamos un análisis macroscópico de protocolos y servicios en el enlace que nos muestra a TCP como el principal protocolo y al Web como el servicio más utilizado al suponer más de tres cuartas partes del tráfico generado. A la vista de la predominancia de estas conexiones TCP realizamos una caracterización en base a procesos estocásticos para el multiplex de flujos TCP. Dicha caracterización se fundamenta en varias características observadas del tráfico, concretamente que la tasa de las conexiones TCP depende fuertemente del retardo extremo a extremo (RTT) de la conexión y que la intermitencia de las mismas no sigue la progresión exponencial que se esperaría del algoritmo de *slow-start*. Esto nos lleva a un modelo mediante restricciones (σ, ρ) que permite el empleo de tecnologías de conmutación de circuitos para la reserva de ancho de banda por flujo.

Con el conocimiento obtenido del funcionamiento de los flujos TCP en la red actual realizamos una revisión del modelo $M/G/\infty$ de flujos. Este es uno de los modelos más empleados tanto para generación de tráfico sintético de datos como en el estudio analítico de las características del mismo. Confirmamos dos de las hipótesis en que se basa (proceso de llegadas de Poisson y duraciones de flujos con varianza infinita) pero vemos que la hipótesis de tasa constante de transferencia dista de lo observado en el tráfico real. Por ello proponemos una alteración del modelo mediante la incorporación de una variable aleatoria de Weibull para la tasa de los flujos. Esta modificación permite que el tráfico resultante ajuste mejor la variabilidad de la marginal del mismo. El modelo $M/G/\infty$ clásico subestima la variabilidad del tráfico aunque modele correctamente su dependencia a largo plazo. Mostramos sin embargo que en las futuras redes de alta velocidad que formarán la próxima generación de Intenet el efecto de dependencia a largo plazo tenderá a desaparecer a costa de un aumento en la variabilidad del tráfico, que pasará a ser el factor que condicione las prestaciones de la red. Esto último es un fuerte apoyo para aceptar modificaciones del modelo en la línea de la propuesta.

Finalmente, empleamos la caracterización de flujos TCP obtenida para proponer un algoritmo de estimación de ancho de banda basado en el RTT de las conexiones. La estimación está orientada a la reserva de ancho de banda en enlaces de proveedores de acceso a Internet. Los resultados muestran que la estimación en base a parámetros conocidos *a-priori* es realizable y mejora los resultados obtenidos con asignadores basados en tasa de pico, asignaciones estáticas o best-effort. Esto abre numerosas posibilidades de estudio de algoritmos de asignación así como de cálculo dinámico de los parámetros de los mismos.

Abstract

This work presents a review of the proposed models for Internet data traffic. These models range from the classical telephony ones to modern concepts like *long-range depencence* and *self-similarity*.

Based on these models we aim at characterizing a large population of Internet users. We use the traffic traces recorded from Public University of Navarra IP over ATM access link. The traffic traces are obtained with a novel measurement equipment for high speed ATM links. We present a macroscopic analysis of our traces which shows that the ATM link bandwith is dominated by the connections from TCP protocol. Most of these connections are from the Web service, then, we propose an stochastic model for the multiplex of these TCP connections. The model depends on two characteristics of the traffic traces: first, the connection throughput has a strong dependence on the connection round-trip time (RTT) and second, the burstiness from these connections doesn't follow an exponential increase progression as the *slow-start* algorithm predicts. With these parameters we propose a (σ, ρ)-constrained stochastic model that can be used in per-flow resource allocation for switched circuit environments.

The understanding of the TCP dynamics on real links obtained from this work leads us to a revision of the $M/G/\infty$ traffic model. This is one of the most popular traffic models for traffic generation or theoretical analysis. We verify the first two hypotesis needed for the model (namely Poisson arrival process and connection durations with infinite variance) but the hypotesis about constant connection throughput is far from true. So we propose a revised version of the model that takes account into the variability of the connection thoughput by using a random variable with a Weibull distribution for this parameter. As a result, the traffic presents a higher variability in the marginal distribution of the process. The $M/G/\infty$ classical traffic model captures the long-range redependence of the real traffic but underestimates the variability of its marginal distribution. Such traffic variability affects queueing performance heavily and may have a stronger influence on Internet performance than long-range dependence as the backbones evolve to high-speed optical networks.

Finally, this TCP flows characterization serves for an RTT-based bandwidth allocation algorithm. This resource allocation is useful in ISP network access links. Our findings suggest that *a-priori* bandwidth estimation is feasible and outperforms a broad class of peak-rate, best effort and static allocation flow switching solutions. Other allocation algorithms and dynamic parameter calculation is the subject of future research.

Capítulo 1

Introducción

1.1 Antecedentes y necesidad de la tesis

Un gran número de estudios han comprobado que la mayor parte del tráfico de la Internet actual está formado por conexiones del nivel de transporte. Por lo tanto debe ser posible el modelado (fenomenológico) del tráfico de Internet como un múltiplex de transacciones. El objetivo de esta Tesis Doctoral será proponer y analizar un modelo de este tipo.

Para poder estudiar un entorno de Telecomunicaciones como la Internet actual se va a comenzar por establecer sus semejanzas y diferencias con otros servicios de telecomunicaciones en cuanto a generación de tráfico. El ejemplo tradicional de esta categoría es la telefonia básica. Se van a repasar brevemente las técnicas tradicionales de caracterización y dimensionamiento de este servicio. Tras esto se analizará la evolución de los modelos para el tráfico de datos en Internet desde los modelos Poissonianos a los basados en procesos fractales.

1.1.1 Modelo telefónico básico

En la actualidad, uno de los servicios de telecomunicación más difundido es sin duda la telefonía básica. Desde sus comienzos hace más de un siglo hasta nuestros días se ha convertido en un servicio de comunicaciones con gran importancia tanto a nivel de usuario residencial, usuario empresarial, como incluso a nivel estratégico y táctico gubernamental.

Este servicio presenta al usuario final una calidad de servicio determinista: una vez establecida la llamada con el otro extremo la comunicación goza de un canal reservado de ancho de banda constante a través del cual la voz no debe sufrir cortes y está limitada la variación de retardos. Esta característica del servicio telefónico se logra gracias al establecimiento de un circuito físico entre los dos extremos con los requisitos necesarios.

La reserva en telefonía de un circuito con un ancho de banda fijo se basa en el análisis de la señal que va a circular por este canal, la señal de voz. Ésta mantiene una calidad aceptable aunque se reduzca su ancho de banda a 4KHz. Igualmente se ha determinado el retardo extremo a extremo máximo que puede existir para que el usuario no perciba cortes molestos. Así pues, sabemos qué caracteristicas ha de tener el canal reservado para cada comunicación de voz.

Todos los circuitos que se establecen cumplen con los requisitos mencionados. Queda por determinar cuál es la disponibilidad del servicio. Una solución completamente mallada, en la cual cada par de abonados tiene un circuito individual, provee de una disponibilidad total pero resulta completamente desproporcionada debido a que durante el funcionamiento normal de la red la mayor parte del equipamiento no se emplearía. Por ello la topología empleada en la red telefónica pública es jerárquica. La comunicación entre abonados pertenecientes a diferentes centrales finales se realiza estableciendo un circuito que atraviesa otras centrales de tránsito, centrales en niveles superiores de la jerarquía. Los enlaces entre las centrales finales y los nodos de tránsito soportan un gran número de circuitos simultáneos gracias a técnicas de multiplexación en frecuencia o en el tiempo. El objetivo de dimensionamiento en estas redes es determinar la capacidad (en número de circuitos de voz) de estos enlaces que permita dar la mayor disponibilidad al menor coste.

Los problemas de dimensionamiento de la red telefónica se centran en torno a dos situaciones en las cuales es imposible establecer el circuito deseado entre dos abonados y por tanto se pierde una llamada por falta de disponibilidad. La primera se debe al bloqueo interno de alguno de los conmutadores en el trayecto del circuito, es decir, la imposibilidad de establecer ese circuito debido a falta de recursos internos del conmutador. Esta probabilidad dependerá de la arquitectura escogida para el conmutador. La segunda causa es la saturación de todos los canales existentes entre nodos finales y de tránsito o entre estos últimos en el trayecto del circuito.

Para la obtencion de reglas de dimensionamiento se asume que no se va a necesitar dar servicio a todos los abonados al mismo tiempo. Para la red los usuarios siguen un comportamiento aleatorio. Para predecir el comportamiento de la red ante estos usuarios se necesita un modelo de los mismos y por ello se comienza por obtener un descripción estadística del comportamiento de los usuarios respecto a instantes en los que solicitan servicio, la duración del mismo y distancia con el destino (bien sean llamadas locales a la central o no). A continuación describiremos este modelo de usuarios y continuaremos presentando la caracterización analítica de los modelos clásicos aplicados a sistemas de conmutación telefónica.

1.1.1.1 Modelo estadístico de los usuarios

Desde el punto de vista de la red interesa más modelar a un conjunto de usuarios que a los usuarios de forma individual. O por lo menos grupos de usuarios por categorías que determinen sus patrones de tráfico como puede ser usuarios residenciales y de empresas. Las características principales de estos usuarios son el patrón que siguen de generación de llamadas y la duración que suelen tener éstas.

Mediante medidas experimentales [1, capítulo 10] se ha comprobado que los tiempos entre llamadas de una población de usuarios se ajusta correctamente a un proceso estocástico $T = (T_k : k = 0, 1, 2...)$ semi-infinito de tiempo discreto y estado continuo donde los T_k son variables aleatorias i.i.d.s (independientes e idénticamente distribuidas) que siguen una distribución exponencial de parámetro λ (1.1).

$$P(T_i < t) = 1 - e^{-\lambda t} \quad , t \geq 0, \lambda \geq 0 \tag{1.1}$$

Se define un proceso estocástico X que describe el *Número de llegadas en el intervalo de tiempo t*. $X = (X_t : t = 0, 1, 2...)$ es un proceso estocástico estacionario en el sentido amplio, discreto, con media λ y varianza σ^2 donde los X_t siguen una variable aleatoria de Poisson (1.2).

$$P(X_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \tag{1.2}$$

En el proceso real de llamadas existen fluctuaciones horarias de la tasa de llegadas, las cuales se introducen considerando el modelo de Poisson homogéneo solo para intervalos de una hora de duración con la tasa correspondiente. Estas tasas horarias siguen patrones diarios predecibles, destacando una hora del día que presenta la mayor actividad y para la cual se realiza el dimensionamiento de los enlaces. Existen igualmente patrones periódicos debidos a las estaciones del año así como suaves variaciones globales entre años.

El segundo parámetro necesario para describir el comportamiento de los usuarios es la duración de las llamadas. La distribución que mejor se aproxima a los datos experimentales es la de una variable aleatoria exponencial de media $1/\mu$. Una propiedad muy importante de las distribuciones exponenciales es que cumplen que $P(T \leq t + t_0 | T > t_0) = P(T \leq t)$, lo cual se suele expresar diciendo que *carecen de memoria*, es decir, el conocimiento de la historia de la variable no ayuda en la predicción de su valor futuro. En el caso de una variable aleatoria que describe la duración de una llamada, si es exponencial, el conocer la duración hasta el momento de la conexión no reduce la incertidumbre sobre cuándo finalizará.

1.1.1.2 Modelos de sistemas

Una central de conmutación con m circuitos de salida y usuarios descritos mediante un proceso de Poisson de llegadas y duraciones de las llamadas exponenciales corresponde en la teoría de colas [2] a un modelo $M/M/m/m$.

En este sistema se calcula la probabilidad de rechazo de llamadas, es decir, de que una nueva petición de servicio no encuentre ninguna línea libre, la cual viene dada por la fórmula B de Erlang (1.3) [2].

$$B\left(m, \frac{\lambda}{\mu}\right) = E_{1,m}\left(\frac{\lambda}{\mu}\right) = \frac{\left(\frac{\lambda}{\mu}\right)^m / m!}{\sum\limits_{k=0}^{m} \left(\frac{\lambda}{\mu}\right)^k / k!} \tag{1.3}$$

El modelo $M/M/m/m$ supone una población infinita de usuarios. En los casos en los que esta hipótesis no resulta razonable pero se puede mantener el modelo de Poisson de llegadas adaptando la tasa al número de usuarios a los que no se está dando servicio se emplea el modelo $M/M/m/m/L$ donde L es el tamaño de la población. Para este sistema la probabiliad de bloqueo viene dada por (1.4).

$$B = \frac{\binom{L-1}{m}\left(\frac{\lambda}{\mu}\right)^m}{\sum\limits_{i=0}^{m} \binom{L-1}{i}\left(\frac{\lambda}{\mu}\right)^i} \tag{1.4}$$

En estos modelos los posibles reintentos de llamada tras bloqueo se modelan tradicionalmente incrementando la tasa del proceso de llegadas, despreciando la correlación existente entre las llamadas bloqueadas y los reintentos.

Estos modelos apoyan la estructura jerárquica de la red telefónica. Se comprueba que la concentración de usuarios hacia niveles superiores de la jerarquía permite obtener la misma probabilidad de bloqueo con un número menor de enlaces. Esto se debe a que al aumentar la multiplexación de usuarios el proceso de generación de tráfico resultante se asemeja cada vez más a un ruido blanco y gaussiano, cuya varianza disminuye con $1/n$.

Los cálculos de probabilidad de bloqueo en rutas completas a través de esta red se encuentran con que generalmente existirán varios caminos posibles entre origen y destino. Mediante varias hipótesis de independencia y otras simplificaciones se consigue resolver el problema reduciéndolo a conjuntos de enlaces serie o paralelo donde cada uno se modela como ya se ha mencionado [1, capítulo 10].

Vemos pues que ante tráfico con características similares al telefónico el dimensionado de los elementos de conmutación se puede realizar en base a modelos sencillos. Gracias a estos modelos podemos diseñar las centrales de conmutación para cumplir unos requisitos de calidad de servicio. Dado que el ancho de banda requerido por conexión es fijo la red se dimensiona en base a la disponibilidad, definida por la probabilidad de rechazo de llamada. Mediante el dimensionamiento para el caso de mayor intensidad de tráfico (la hora cargada) se logra prevenir la congestión en la red.

1.2 Modelos del tráfico de datos en Internet

Entre las décadas de los 80 y los 90 se ha dado un crecimiento exponencial del tráfico en redes de datos [3, 4]. El fenómeno de interconexión de redes de area local conocido como Internet ha evolucionado desde sus inicios militares pasando por el ámbito académico universitario hasta su salida al entorno comercial.

Los servicios clásicos de la Internet inicial como el terminal remoto (Telnet), el correo electrónico (SMTP) o la transferencia de ficheros (FTP) han dado paso a una plétora de nuevos servicios y aplicaciones que están viendo su auge con el avance de Internet en países desarrollados a niveles similares al teléfono.

Servicios como el World Wide Web, la difusión de video y radio, la videoconferencia o la voz sobre IP, junto con las capacidades de transmisión de las nuevas redes ópticas y de acceso al usuario de banda ancha han llevado al tráfico de datos a representar una parte muy importante del tráfico de información mundial. Por tanto se ha pasado a necesitar conocer y controlar el comportamiento de la red y poder asegurar tanto su funcionamiento como unos parámetros contratados de calidad, al igual que sucedió con la telefonía básica.

1.2.1 Modelos clásicos de sistemas de colas

Los nodos de transmisión de datos (routers, conmutadores, etc) operan generalmente mediante estrategias de almacenamiento y retransmisión (store-and-forward). Estos sistemas almacenan en memoria el paquete de datos recibido, calculan cuál es el enlace de salida que debe tomar y esperan hasta disponer de los recursos necesarios para reenviarlo hacia su destino. Esta característica puede incorporarse a los modelos de teoría de colas mencionados en el estudio del sistema de conmutación telefónico. En estos casos los parámetros de mayor importancia para el dimensionamiento serán el tiempo en el sistema y la probabilidad de que dicho tiempo exceda un valor. Si la cola de la que dispone ha de modelarse como finita se deberá tener en cuenta de nuevo la probabilidad de rechazo o pérdida, en este caso de un paquete.

El modelo más simple de teoría de colas supone que el proceso de llegada de mensajes es de Poisson, los tiempos de servicio siguen una distribución exponencial y todos los procesos son estacionarios e independientes.

En el caso de transmisión de datos, el tiempo de servicio exponencial implicaría que los tamaños de los paquetes, o más correctamente sus tiempos de servicio, siguieran esa distribución. Se comprueba sin embargo que en su lugar presentan una clara bimodalidad debida a protocolos como TCP que alternan paquetes del máximo tamaño permitido por la red con paquetes de confirmación sin datos. Sin embargo, aunque el tiempo de servicio no siga una distribución exponencial resultaría de gran utilidad para el modelado que el proceso de llegadas sí fuera un proceso de Poisson, dado que en ese caso aún se pueden dar soluciones analíticas para los momentos de las diferentes distribuciones de interés.

1.2.2 *The failure of Poisson modeling*

En los comienzos de ARPANET los modelos poissonianos daban resultados acordes con las mediciones reales. Se ha comprobado sin embargo que el tráfico de datos actual presenta unas características muy diferentes de las de un proceso de Poisson, las cuales se verán a continuación.

Comenzaremos dando una serie de definiciones para los procesos de tráfico y sus estadísticos que emplearemos en el resto del trabajo.

Definición: Sea el proceso de llegadas de paquetes descrito por los instantes de llegada $a = (a_k : k = 0, 1, ...)$ o equivalentemente por $T = (T_k : k = 1, 2, ...)$ donde T_i es el tiempo entre la llegada del paquete i-ésimo y el (i-1)-ésimo. Se define el proceso de cuenta de paquetes $X = (X_i : i = 1, 2, ...)$ de forma que cada X_i es el número de llegadas que se producen en un intervalo de anchura δ, siendo estos intervalos contiguos. Si X es un proceso estocástico estacionario en el sentido amplio llamaremos $\mu = E[X_i]$ a su media, $\sigma^2 = var(X) = E[(X_t - \mu)]^2$ a su varianza finita, $cov(X_t, X_{t+k}) = E[(X_t - \mu)(X_{t+k} - \mu)] = \gamma(k)$ a su función de auto-covarianza y $\rho(k) = \gamma(k)/\gamma(0)$ a su función de autocorrelación.

X es un proceso de Poisson si las variables X_i son independientes e idénticamente distribuidas (i.i.d.) con distribución de Poisson de parámetro λ. En este caso las variables aleatorias T_i son también variables i.i.d. con distribución exponencial de parámetro λ.

1.2.2.1 Los tiempos entre llegadas no son exponenciales

Para un proceso de Poisson la marginal del proceso T será una exponencial y la función de autocorrelación nula salvo para desplazamiento cero, donde debe valer uno. Diversos estudios

[5, 6, 7, 8, 9, 10] han comprobado que no se cumple ninguna de estas características en el caso de tráfico de Internet.

Jain y Routhier [5] distinguieron diferentes zonas en la marginal de T, lo cual les llevó a formular un modelo diferente basado en lo que denominaron *trenes de paquetes*. En este modelo el tráfico entre pares de nodos se da en forma de ráfagas de paquetes muy próximos entre si. Los tiempos de separación entre ráfagas se diferencian en varios ordenes de magnitud de los tiempos de separación entre paquetes. Un método para generar tráfico de este tipo se basa en una cadena de Markov de dos estados, donde uno de los estados corresponde a la generación de paquetes y el otro a la espera entre trenes de paquetes [11]. Se debe generar por separado el proceso correspondiente a cada par de nodos dado que la suma de procesos de trenes de paquetes no va a dar un nuevo proceso de este tipo.

Sin embargo para este modelo se consideró que el tráfico no destacaba por valores altos de autocorrelación. En la década de los 90 se revela que aunque los valores concretos de la autocorrelación sean pequeños tiene una gran importancia en las propiedades del tráfico la velocidad a la que decaen.

1.2.2.2 Ráfagas en todas las escalas

Con el aumento de capacidad de las redes y el empleo de nuevos servicios en la década de los 90 se ha producido un cambio en la naturaleza del tráfico en Internet [11, 4]. El desarrollo de nuevas herramientas de elevada precisión para la monitorización de redes LAN y WAN [7, 12, 13] ha permitido que saliera a la luz que el proceso de llegadas de tráfico presenta ráfagas en un amplio rango de escalas temporales. Examinando detalladamente una ráfaga de tráfico a cualquier escala nos encontramos con que está compuesta por periodos de baja actividad separados por unos pocos periodos muy activos. Medidas tradicionales de intermitencia como la relación entre el ancho de banda pico y el medio no son fáciles de interpretar ante este fenómeno, dada la fuerte dependencia que presentan con el intervalo en el cual se calculen. El coeficiente de variación del proceso de llegadas no converge rápidamente hacia cero al aumentar el intervalo de observación, a diferencia de su comportamiento en los modelos más empleados como el de Poisson, llegadas en bloque de Poisson, hiperexponenciales o procesos de Poisson modulados por Markov (MMPPs). Otra medida de la intermitencia del proceso, el índice de dispersión de cuentas para el proceso de llegadas (relación entre la varianza del número observado de llegadas en un intervalo respecto a la media), crece a lo largo de varios órdenes de magnitud de agregación del proceso, a diferencia de los procesos mencionados que se estabilizan en escalas de tiempo relacionadas con sus parámetros de definición [8].

A diferencia del caso de telefonía y de la aplicación de un modelo de Poisson para el tráfico de datos, un aumento en los buffers para absorber las ráfagas del tráfico real conlleva un decremento de la probabilidad de desbordamiento de los mismos considerablemente inferior al caso de tráfico poissoniano. Se llegaría a necesitar tamaños de buffer que sobrepasaran el límite impuesto por los requisitos de máximo retardo en el nodo. Se puede acotar el porcentaje de pérdidas en grandes intervalos de tiempo pero éste no se mantiene en intervalos inferiores sino que las pérdidas se concentran en los intervalos de alta actividad y desaparecen en los de baja actividad. Este fenómeno de ráfagas en todas las escalas lleva a un sobredimensionamiento de la red si se busca minimizar la congestión, con la consiguiente pérdida de economía de escala [8].

1.2.2.3 ¿Sigue siendo válido un modelo de tráfico de Poisson?

El modelo de Poisson se mantiene como válido tan solo para modelar procesos de llegada de conexiones. Concretamente conexiones TCP de servicios que mantengan una fuerte relación entre una nueva conexión y la incorporación de un nuevo usuario a la red. Esta característica conlleva la propiedad de falta de memoria de la exponencial en el tiempo entre llegadas. De esta forma el proceso de Poisson vuelve a modelar la llegada de usuarios a la red, igual que hacía en el caso de la telefonía, pudiéndose considerar un proceso homogéneo en intervalos en torno a una hora de duración. Cumplen estas características las conexiones de Telnet y de control de FTP [14] mientras que servicios como NNTP, SMTP, FTP-data o WWW [15, 16] son muy diferentes de un modelo de Poisson. Sin embargo, aun para Telnet o FTP-ctrl, una distribución exponencial no va a modelar bien la llegada de nuevos paquetes dentro de las conexiones de estos servicios ya que subestiman la capacidad que tiene el tráfico para formar ráfagas [14].

Otros modelos matemáticamente abordables que se han intentado aplicar son por ejemplo los procesos de renovación. Estos procesos engloban al de Poisson dado que se caracterizan porque sus tiempos entre llegadas son independientes e idénticamente distribuidos [17]. Sin embargo la función de autocorrelación de estos procesos es nula a cualquier desplazamiento, mientras que el tráfico real presenta una fuerte correlación [18]. Estos procesos tienden a un ruido blanco y gaussiano a medida que son observados a mayor distancia.

Se ha comprobado que tanto procesos derivados del de Poisson, como los procesos compuestos de Poisson (compound Poisson), procesos de llegadas en bloque de Poisson, hiperexponenciales o procesos de Poisson modulados por Markov (Markov-modulated Poisson processes) tampoco son capaces de capturar las características de las ráfagas que se producen en el tráfico de datos. Todos estos procesos carecen de la suficiente variabilidad para representar el tráfico analizado, el cual presenta ráfagas importantes en escalas que van desde los milisegundos a los cientos de segundos [7]. Estos modelos pueden ser ampliados para capturar las nuevas propiedades del tráfico, sin embargo esto requiere incrementar considerablemente su complejidad reduciendo su tratabilidad analítica y añadiendo gran número de parámetros cada vez más difíciles de interpretar. Se necesita un nuevo modelo que tenga la misma estructura de ráfagas que el tráfico real pero con un reducido número de parámetros.

1.2.3 Dependencia a largo plazo y autosimilitud

Estudios de tráfico más recientes han mostrado que el tráfico en redes modernas de datos presenta lo que se denomina dependencia a largo plazo [18], efecto Joseph [19] o LRD (Long Range Dependence).

Diremos que el proceso X estacionario en el sentido amplio presenta dependencia a largo plazo si existe un $0 < \beta < 1$ y una constante $c_p > 0$ tal que se verifica (1.5) [20].

$$\lim_{k\to\infty} \frac{\rho(k)}{c_p k^{-\beta}} = 1 \tag{1.5}$$

Se puede asumir que la función de autocorrelación de X es de la forma $\rho(k) \sim c_p k^{-\beta}$ para $k \to \infty$. Entonces es sencillo ver que la función de autocorrelación de este proceso no es sumable, es decir, $\sum_k \rho(k) \to \infty$. Una consecuencia directa de esta divergencia es la presencia de un polo en el origen de la función de densidad espectral, dado que es la serie de Fourier de la autocorrelación (1.6).

$$S(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho(k) e^{-ik\omega} \tag{1.6}$$

No importa que los valores de la autocorrelación puedan ser pequeños, lo determinante es que su suma pueda ser indefinidamente grande. En comparación, para procesos con dependecia a corto plazo la autocorrelación decae a cero lo suficientemente rápido para ser sumable.

Para describir las propiedades de este proceso X definimos una familia de procesos $X^{(m)}$ que sean versiones promediadas del mismo. Tomamos $X^{(m)} = (X_k^{(m)} : k = 1, 2...)$ donde $X_k^{(m)} = (X_{km-m+1} + ... + X_{km})/m$ para $m = 1, 2, 3....$ Estos nuevos procesos son estacionarios en el sentido amplio con varianza σ_m^2, función de autocovarianza $\gamma^{(m)}(k)$ y función de autocorrelación $\rho^{(m)}(k)$.

Una propiedad fundamental de los procesos dependiendes a largo plazo es que la varianza del agregado decae con $m^{-\beta}$, es decir, decae más lentamente que m^{-1} $(0 < \beta < 1)$ lo cual implica que $m\sigma_m^2 \to \infty$ con $m \to \infty$. Esto contrasta con el caso de procesos estacionarios formados por variables i.i.d.s para los cuales el Teorema Central del Límite indica que la varianza del agregado va a decaer con m^{-1}. Incluso para el caso de procesos con dependencia a corto plazo Mandelbrot y Van Ness demostraron que la varianza decae con $\beta = 1$ [21].

Se puede demostrar que para un proceso X con dependencia a largo plazo la estructura de autocorrelación de $X^{(m)}$ tiende a una función no degenerada (1.7), es decir, a un comportamiento asíntótico diferente de un ruido blanco. Los procesos que cumplen esta propiedad, $\rho^{(m)}(k) \to \rho(k)$ para k y m lo suficientemente grandes, se dice que son *asintóticamente auto-similares de segundo orden* con parámetro de Hurst $H = 1 - \beta/2$ [11, 22]. Se dice que un proceso X es *exactamente auto-similar de segundo orden* o fractal si la autocorrelación $\rho^{(m)}(k) = \rho(k) = g(k)$ para todo $m \geq 1, k \geq 0$ [18]. Definiciones equivalentes para fractal son que $\sigma_m^2 = \sigma^2 m^{-\beta}$ o que la densidad espectral siga la forma (1.8). Finalmente un proceso es *exactamente auto-similar* si $X =_d m^H X^m$ donde $=_d$ representa igualdad en sentido de distribuciones de dimensión finita. Dado que la autocorrelación de estos procesos no es sumable tenemos que $0 < \beta < 1$ y por tanto $1/2 < H < 1$. Para el caso de procesos con dependencia a corto plazo o variables independientes el parámetro H vale $1/2$.

$$\lim_{m\to\infty} \rho^{(m)}(k) \to \rho(k) = g(k) \equiv \frac{1}{2}\left[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}\right] \tag{1.7}$$

$$0 < \beta < 1, \quad k \in \{1, 2, ...\}$$

$$f(\lambda) = c\left|e^{2\pi i\lambda} - 1\right|^2 \sum_{l=-\infty}^{\infty} \frac{1}{|\lambda + l|^{3-\beta}}, -\frac{1}{2} \leq \lambda \leq \frac{1}{2} \tag{1.8}$$

Para estos procesos dependientes a largo plazo la estadística R/S o estadística de rango ajustado reescalado, definida por (1.9), sigue la forma $E[R(n)/S(n)] \approx n^H$ con $n \to \infty$. Esta medida es capaz de capturar las fluctuaciones en el proceso, destacando los procesos autosimilares por no tender a $n^{0.5}$ como hacen todos los dependientes a corto plazo.

$$R(n)/S(n) = 1/S(n)[max(0, W_1, W_2, ..., W_n) - min(0, W_1, W_2, ..., W_n)] \tag{1.9}$$

$$W_k = (X_1 + X_2 + ... + X_k) - k\overline{X}(n), 1 \leq k \leq n$$

$$\overline{X}(n) \equiv \text{media muestral de } X_i,\ S^2(n) \equiv \text{varianza muestral de } X_i$$

Para un proceso exactamente auto-similar de segundo orden la estructura de autocorrelación se mantiene aunque cambiemos de escala de observación agregando en los procesos $X^{(m)}$. El efecto visual de una realización del proceso es el mismo independientemente de la escala de observación, lo cual se traduce en que la presencia de ráfagas se pueda dar en cualquier escala y no exista un tamaño "típico" de ráfaga.

Los modelos matemáticamente abordables que se han intentado aplicar para modelar el tráfico de datos poseen autocorrelaciones que decaen rápidamente. Los procesos de renovación, caracterizados por que sus a_k son independientes e idénticamente distribuidos y entre los que se engloba el de Poisson [17], los procesos compuestos de Poisson (compound Poisson), procesos de llegadas en bloque de Poisson, hiperexponenciales, procesos de Poisson modulados por Markov (Markov-Modulated Poisson Processes) o los procesos ARMA, todos ellos carecen de la suficiente variabilidad para representar el tráfico de datos. Estos procesos presentan dependencia a corto plazo, con lo que a partir de cierto nivel de agregación tienden a un ruido blanco gaussiano de varianza finita.

1.2.3.1 Características autosimilares del tráfico de datos

Los resultados respecto a la presencia de ráfagas en todas las escalas del tráfico de datos en redes de area local [7, 8] pueden ser justificados por un carácter autosimilar del proceso de tráfico. El crecimiento del índice de dispersión de cuentas en un intervalo de varios órdenes de magnitud concuerda con el comportamiento para procesos autosimilares. Igualmente la varianza de los procesos $X^{(m)}$ del tráfico real decaen con m de forma más lenta que m^{-1}, la función de densidad espectral del proceso de tráfico diverge cerca del origen y la estadística R/S da un $0.5 < H < 1$.

Se ha comprobado que el tráfico en LANs durante las horas de alta actividad se comporta de forma acorde con un proceso autosimilar de segundo orden con parámetro de Hurst en torno a 0.8 mientras que en periodos de baja intensidad de tráfico es asintóticamente autosimilar. Este valor de H resulta ligeramente superior durante las horas de mayor intensidad de tráfico. Dado que un mayor valor de H indica un mayor alejamiento de un comportamiento de dependencia a corto plazo o independencia (para el que $H = 0.5$) vemos que cuando se intensifica el tráfico el proceso no tiende a estructuras más suaves, con menores variaciones, sino hacia procesos que pueden presentar mayores ráfagas e intermitencia [18]. Este comportamiento es radicalmente diferente del esperado por analogía con el tráfico telefónico, donde la concentración de usuarios permitía un diseño más ajustado al reducirse más rápidamente la varianza.

En el caso de redes de area extensa (WANs) se ha comprobado que también presentan dependencia a largo plazo, si bien las trazas analizadas tan solo pueden considerarse asintóticamente autosimilares [12]. Para el caso de situaciones de muy baja actividad, correspondientes a periodos de tiempo sin usuarios humanos, se puede llegar a presentar un $H \approx 0.55$ con lo que podría modelarse por procesos con dependencia a corto plazo [18].

1.2.3.2 Causas de la LRD

Un modelo de tráfico autosimilar va a capturar la característica de dependencia a largo plazo del tráfico y por tanto proveerá de resultados más cercanos a la realidad que los modelos clásicos [23, 24, 25, 26], los cuales subestiman parámetros como el retardo medio o el tamaño que puede alcanzar la cola. Modelos como los procesos de poisson modulados por Markov pueden ajustarse a las características del tráfico pero a costa de aumentar considerablemente el número de parámetros necesarios, con lo que resulta mucho más complicado interpretar el significado de cada uno [17]. Un razonamiento basado en la Cuchila de Occam lleva a aceptar un modelo con dependencia a largo plazo. Una de las ventajas de un modelo autosimilar es precisamente poder capturar el carácter de dependencia a largo plazo con un reducido número de parámetros y dar por tanto un modelado parsimonioso [27]. Sin embargo los procesos estocásticos autosimilares se encuentran a primera vista desligados de su causa, el modelo no es capaz de dar una explicación física de la autosimilaridad del tráfico.

Se han encontrado sin embargo métodos constructivos que a partir de procesos con interpretación física plausible den un proceso autosimilar mediante agregación. Tal es el caso del modelo de fuentes ON/OFF [18, 27, 19, 12]. Una fuente ON/OFF se considera de forma análoga a un tren de paquetes como la consecución de intervalos de actividad e inactividad de una fuente y representaría el proceso de intercambio de datos entre dos nodos. Durante los intervalos ON se considera una transmisión según un modelo de fluidos con tasa unidad. Se demuestra [19] que la superposición de un gran número de estas fuentes tiende a un proceso autosimilar, de hecho a un FBM (Fractional Brownian Motion) [20], el único proceso autosimilar gaussiano con incrementos estacionarios.

El proceso de incrementos de un FBM se conoce como FGN o Fractional Gaussian Noise. Existe un requisito fundamental para la convergencia a FBM o equivalentemente a FGN y se trata de que los intervalos ON y OFF, o al menos uno de ellos, sigan una distribución con varianza infinita, lo que se conoce como efecto Noé. Una forma simple de considerar este efecto es mediante distribuciones de cola pesada con varianza infinita. Estas distribuciones cumplen la relación (1.10).

$$P(X > x) \sim cx^{-\alpha}, x \to \infty, 1 < \alpha < 2 \tag{1.10}$$

Si $\alpha < 2$ entonces el segundo momento de la distribución (y por tanto la varianza) no existe. Si $\alpha < 1$ ni tan siquiera existe la media [27]. Un caso simple de distribuciones de este tipo es la familia de distribuciones de Pareto definidas por (1.11) [28].

$$P(X > x) = \left(\frac{K}{x}\right)^{\alpha} \tag{1.11}$$

Se comprueba que los procesos de tráfico entre pares de nodos presentan efecto Noé con lo que esta explicación resulta adecuada [27]. El parámetro H del proceso resultante y por lo tanto el carácter autosimilar del tráfico depende del valor del α más pequeño de las distribuciones de las fuentes ON/OFF mediante la relación $H = (3 - \alpha)/2$, por lo que es importante la estimación de este parámetro. En el caso de que haya diferentes patrones de fuentes con diferente α el que va a predominar en el agregado va a ser siempre el que dé mayor valor de H.

Este modelo es un caso particular del de trenes de paquetes, sin embargo ni en la definición inicial [5] ni en estudios posteriores se habían empleado distribuciones de cola pesada. Con anterioridad a la introdución de modelos autosimilares en el modelado de tráfico los trenes de paquetes siempre se habían tomado con distribuciones de varianza finita, tales como exponenciales o geométricas [27].

Para completar este modelo queda determinar el motivo por el cual los procesos ON/OFF presenten efecto Noé. Dado que hoy en día la mayor parte del tráfico es TCP tiene sentido abordar primero los flujos pertenecientes a este protocolo. Un hecho íntimamente relacionado con los flujos de datos es la distribución que siguen los ficheros que suelen encontrarse en servidores de Internet. Una gran cantidad de servicios de Internet se basan en la transferencia de ficheros entre computadoras y dado que la distribución de los tamaños de los mismos presenta varianza infinita [6, 16] es lógico pensar que también la presenten los tamaños de los ficheros transferidos y las duraciones de dichas transferencias [12]. Los principales servicios de Internet presentan estas características. Así, la transferencia de ficheros grandes mediante FTP se modela correctamente con la cola de una distribución de Pareto [9]. Los flujos del servicio WWW, el cual representa el torno al 75% del tráfico WAN [13], se ajustan a colas de Pareto con exponentes entorno a 1.2 (media finita pero varianza infinita) tanto para la duración de los flujos como para sus tamaños e igualmente los intervalos OFF presentan efecto Noé pero en menor grado [16].

El hecho de que las transferencias en la red se realicen sobre un protocolo fiable como TCP, el cual realizará retransmisiones en caso de necesidad y por lo tanto alargará los flujos, apoya al efecto Noé en la duración de los flujos [29, 24] pero aun sin él, la distribución de cola pesada de los ficheros transferidos conlleva el efecto Joseph de forma relativamente independiente de cambios en la topología de la red, el tipo de tráfico, la distribución de tiempos entre llegadas de flujos o los recursos de la red [24].

El protocolo que sigue a TCP en cuestión de tráfico generado es UDP. Dado su carácter básico no fiable su empleo tiende a centrarse en torno a servicios multimedia. Para este tipo de servicios la mayor componente va a ser la difusión de video a tasa variable (VBR), que independientemente del codificador, su naturaleza (videoconferencia, videoteléfono, video con calidad de estudio) y a una escala mayor que una escena se ha comprobado que presenta nuevamente dependencia a largo plazo [30, 31].

1.2.3.3 Consecuencias de la LRD

Como se ha visto, el tráfico en redes LAN y WAN es fractal, es decir, puede modelarse adecuadamente con un proceso estacionario con dependecia a largo plazo. Un proceso autosimilar no tiene un tamaño típico de ráfaga, lo cual permite que en cualquier escala de observación el proceso no se suavice. Los efectos de esta estructura de correlación sobre las prestaciones de un servidor con cola son muy diferentes de los obtenidos con un proceso de Poisson u otros procesos con dependecia a corto plazo. Se comprueba que es precisamente la existencia de dependencia a largo plazo en las trazas de tráfico el parámetro que perjudica gravemente al sistema [23, 22].

Con una entrada autosimilar la distribución de la ocupación de la cola decae más lentamente que una hipérbola y por tanto de forma más lenta que con un proceso de Poisson, para el que la caída es exponencial [27, 23, 22]. Igualmente la probabilidad de desbordamiento del buffer también decae hiperbólicamente con el tamaño del mismo [22]. Aumentar los tamaños de los bufferes de la cola para aumentar el throughput tiene el perjudicial efecto de aumentar exponencialmente el retardo en cola mientras que la probabilidad de pérdida desciende más lentamente. Una política de incremento de ancho de banda obtiene mejores resultados en el aumento del throughput y descenso del retardo [25].

La multiplexación de fuentes autosimilares tiene el problema de mantenerse el H de las fuentes o en el caso de que sean diferentes, de la que lo tenga mayor. Esto implica que ante multiplexación de usuarios se va a mantener la presencia de ráfagas. Sin embargo, el modelo basado en FGN predice también una reducción del parámetro de variabilidad $a = \sigma^2/\mu$ con $\sqrt{n}$, con lo que si la capacidad del servidor es lo suficientemente grande la reducción de las fluctuaciones permitirá obtener ganancia estadística [23].

1.2.3.4 Modelos de generación de tráfico autosimilar

Los modelos de tráfico fractal presentan serias dificultades a la hora de resolver de forma analítica problemas de dimensionamiento de enlaces, tamaños de bufferes, técnicas de control de admisión, etc. Generalmente estos modelos permiten poner de manifiesto factores que pueden ser la causa de las diferentes características de los procesos de tráfico reales. Ante la dificultad de resolver analíticamente los sistemas de interés recurriremos a trazas de tráfico generadas en base al modelo, con las mismas características principales que el tráfico real, que permitan poner a prueba los sistemas o realizar extensivos estudios numéricos.

Los modelos que se han empleado principalmente en la literatura se detallan a continuación.

FBM Para construir un proceso estocástico fractal se parte de una condición para la autocorrelación (1.7) del mismo. Dado que es una condición referente a momentos de segundo orden resulta adecuado escoger procesos gaussianos dado que por definición vienen completamente caracterizados por sus momentos de primer y segundo orden.

Se parte del ruido browniano fraccional o FBM y de su proceso gaussiano de incrementos (FGN). Cuando el parámetro $H = 1/2$ este ruido gaussiano fraccional se reduce al ruido browniano estándar. Tomando un FBM se puede definir un proceso de tráfico de la forma (1.12).

$$A_t = mt + \sqrt{am} Z_t \tag{1.12}$$

donde A_t es el proceso acumulado de tráfico en el intervalo $t \in (-\infty, \infty)$, Z_t un FBM normalizado, $t \in (-\infty, \infty)$, $m > 0$ es la media del tráfico, $a > 0$ un coeficiente de varianza y H el parámetro de Hurst de Z_t.

Este modelo presenta dos características muy interesantes. La primera se basa en que un modelo de fuentes ON/OFF converge básicamente a un FBM, con lo que existe una relación causal con procesos más similares a fuentes reales de tráfico. La segunda característica es que se trata de un modelo simple y tratable de forma analítica, lo cual ha permitido demostrar teóricamente que si un tráfico de este tipo ataca a un servidor con cola la ocupación de la misma sigue aproximadamente una distribución de Weibull [32].

Para la generación de realizaciones de un proceso de este tipo se han desarrollado diversos métodos aproximados [33, 34].

ON-OFF El método más empleado para la generación de trazas de tráfico es el basado en fuentes ON/OFF o procesos de renovación con distribuciones de cola pesada [27, 19]. Los motivos más claros se basan en la justificación fenomenológica del modelo y la posibilidad de introducir valores de las distribuciones correspondientes a situaciones reales y observar sus efectos [24, 25].

Este modelo se ha empleado desde enfoques aproximados basados en modelos de fluidos, es decir, con fuentes que transmiten de forma continuada durante el intervalo ON o con distribuciones simples para los tiempos entre paquetes en el mismo [16, 12, 18, 35], hasta simulaciones basadas en modelos de nodos de red con pilas de protocolos con funcionamiento análogo al real que permiten analizar los efectos de los protocolos empleados, principalmente los efectos de transmisión fiable y control de congestión [24, 25].

M/G/∞ Este modelo considera que las llegadas a un servidor con cola infinita se producen siguiendo un proceso de Poisson. El proceso X_t que interesa es el correspondiente a la cuenta de clientes en el sistema respecto al tiempo $t = 0, 1, 2....$ Para que el proceso X_t sea asintóticamente autosimilar de segundo orden hace falta que la distribución del tiempo de servicio en el servidor presente cola pesada con varianza infinita [14, 22].

fARIMA El modelo ARIMA fraccional es una extensión del ARIMA o AutoRegressive Integrated Moving Average [20]. Una de las características de estos procesos es su capacidad para capturar tanto características de dependencia a corto plazo o correlaciones para valores pequeños de desplazamiento como el carácter de dependencia a largo plazo. Los modelos autorregresivos se han empleado bastante para ajustar funciones empíricas de autocorrelación, pero en general son incapaces de ajustarse también a la distribución marginal del proceso [17].

Mapas caóticos Los sistemas caóticos se basan en la evolución del estado de un sistema mediante una transformación no lineal y determinista. Sin embargo, aunque los valores del proceso quedan perfectamente determinados con el estado inicial, estos sistemas presentan una elevada sensibilidad a las condiciones iniciales, de forma que dos estados iniciales prácticamente idénticos pueden dar lugar a trayectorias completamente diferentes. Aunque el concepto está en si muy desligado del concepto de fractales sucede que las trayectorias descritas en el espacio de estados de ciertos sistemas caóticos pueden ser descritas mediante geometría fractal. Así, se han empleado sistemas caóticos como generadores de procesos fractales, se han planteado nuevas técnicas para abordar los problemas de dimensionamiento de colas basadas en las características matemáticas de estos modelos [36] e incluso se han aplicado a la caracterización del funcionamiento de TCP ante control de congestión [37].

1.3 Objetivo

En este trabajo se va a abordar el estudio del tráfico en Internet desde el punto de vista de una suma de contribuciones. En el caso telefónico se podía considerar que cada usuario contribuía al tráfico total de forma idéntica al resto de entre una población mínimamente homogénea. Se pretende abordar el modelo del tráfico de datos de Internet desde un punto de vista similar, donde las contribuciones estarán formadas por conexiones TCP.

Así se va a orientar el desarrollo hacia un estudio de las conexiones TCP con el objetivo de distinguir las características de las mismas que permitirían modelar su funcionamiento y su aportación al tráfico en la red y a las prestaciones y calidad de servicio obtenidas por los usuarios.

El protocolo TCP es el más empleado dado que la mayoría de los servicios requieren la transferencia fiable de información. Con él se incorporan los efectos de control de flujo y control de congestión que como se verá van a tener una importante influencia en el funcionamiento de estos flujos o conexiones. Esta transferencia de información se realiza dentro de servicios basados en la resolución de transacciones, generalmente en la forma de peticiones de información que son contestadas con los resultados de la transacción. Tal es el caso de servicios como FTP o WWW y quedarían fuera de esta categoría los servicios de terminal remoto y otros servicios interactivos (IRC, Talk...).

Desde el punto de vista de la descomposición del tráfico en sus diferentes contribuciones se pretenden abordar los fenómenos presentes en el proceso multiplexado como es la dependencia a largo plazo o la fractalidad del mismo, con el objetivo de entender las causas, efectos y situaciones en las que resultan necesarios para completar modelos que puedan servir para estudios de prestaciones de protocolos, algoritmos o redes.

Con una visión más clara de estos efectos se realizará un estudio de la calidad de servicio en una red sin control como la Internet orientada a estos servicios transaccionales y al punto de vista del extremo del usuario. Para ello se definirá una métrica de QoS basada en la tasa de transferencia de información y se buscará un modelo estadístico para este parámetro. Al modelo más adecuado a servicios transaccionales se intentará añadir este parámetro para incluir la visión de calidad de los servicios dentro del conjunto de variables.

Con este nuevo modelo se podrá abordar el problema de dimensionamiento aproximado para cumplir condiciones de calidad de servicio al usuario o utilización de enlaces.

1.4 Herramienta de medida

Para la realización de los análisis propuestos se dispone de información detallada del tráfico de datos del acceso a Internet de la red del Campus de la Universidad Pública de Navarra (UPNA), obtenida mediante una nueva herramienta de monitorización desarrollada a tal efecto y que presenta avanzadas características de precisión y almacenaje de información.

1.4.1 Escenario de medida

Para la obtención de datos reales se pretende emplear el escenario de red disponible en la Universidad Pública de Navarra donde un enlace de IP sobre ATM conecta más de un millar de terminales de internet a la red académica nacional RedIRIS a través de su nodo principal en Madrid. Este enlace da servicio a la Universidad Pública de Navarra (a partir de ahora UPNA), a la Universidad de Navarra (UNAV) y a la Asociación de la Industria de Navarra (AIN). La tasa de transferencia a nivel ATM en este enlace es de 4Mbps en el sentido de RedIRIS a Navarra (a partir de ahora sentido entrante) y de 2Mbps en el opuesto (sentido saliente).

La topología de RedIRIS es una estrella de circuitos virtuales permanentes (PVCs) que interconectan universidades y centros de investigación de todo el país con el punto central en Madrid. RedIRIS a su vez participa en el Proyecto TEN-155 mediante un acceso a 45Mbps que la interconecta con las distintas redes académicas y de investigación europeas; igualmente posee un enlace ATM a 19.2Mbps para el tráfico con Estados Unidos y un enlace de aproximadamente 18Mbps con Ibernet que le permite el intercambio de tráfico con proveedores comerciales de Internet. Se considera este un escenario lo suficientemente significativo a nivel de usuario y de red para obtener conclusiones generalizables respecto al problema propuesto (figuras 1.1 y 1.2).

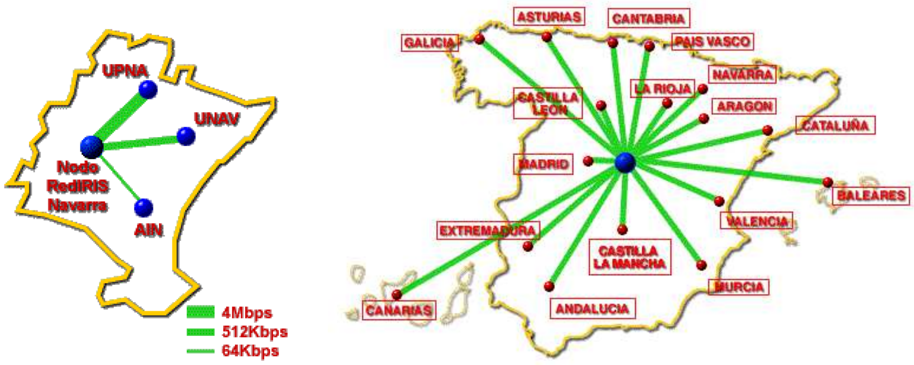


Figura 1.1: Enlaces nacionales en el entorno de medida

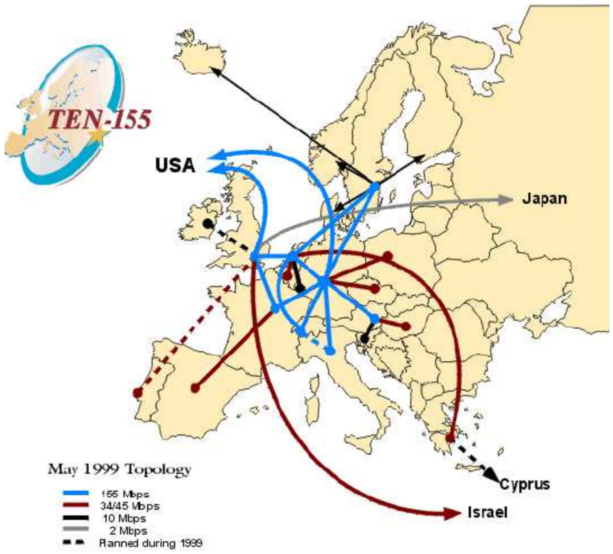


Figura 1.2: Enlaces internacionales en el entorno de medida

Por su parte la Universidad Pública de Navarra posee más de 900 máquinas de usuario conectadas a la red con acceso al exterior. Estas computadoras están compuestas tanto por maquinas compartidas de laboratorios como por ordenadores personales de despachos. Este número de usuarios reparte su actividad entre un gran número de servicios de Internet (FTP, NNTP, WWW, SMTP, POP3...) enmarcados dentro de lo que sería un escenario de tarifa plana. El conjunto de destinos accedidos representa una muestra real que comprende tanto destinos académicos dentro de la red nacional como la europea y un gran número de accesos a través los caminos internacionales con EE.UU. y el resto del mundo, dado que los usuarios no se encuentran restringidos a un número predeterminado de destinos.

El escenario analizado es un ejemplo representativo de un gran número de configuraciones de red muy comunes en la actualidad. Por ejemplo, la mayor parte de los proveedores de servicio de Internet (ISPs) contratan circuitos virtuales ATM con una operadora para dar servicio de acceso a Internet a sus clientes. Una situación análoga se da con redes corporativas o académicas, que emplean este mismo tipo de enlaces para conectar sus servidores y usuarios a Internet.

1.4.2 IPmiser

Durante los dos años previos a este trabajo se ha participado en el desarrollo de una novedosa herramienta para el análisis del tráfico en un enlace de alta capacidad de Internet, el sistema *IPmiser*. Este trabajo ha sido realizado por miembros del Grupo de Redes, Sistemas y Servicios Telemáticos de la Universidad Pública de Navarra en colaboración con Telefónica I+D.

En el sistema *IPmiser* un elemento hardware se encarga de la extracción de las celdas del enlace ATM a 155Mbps (STM-1), sin pérdidas, de manera que las cabeceras de los distintos protocolos se envían a la estación de proceso, que es el núcleo del sistema. Esta estación no solo procesa los datos del enlace sino que sirve la información en tiempo real a través de la Internet, utilizando encriptación y autentificación.

La estación de trabajo implementa un sistema de clasificación de paquetes que hace que no se pierda información de monitorización, pudiendo dar un gran número de parámetros tanto a nivel de Red como de conexiones TCP.

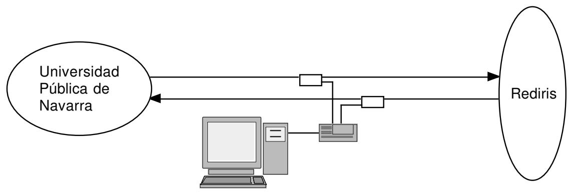


Figura 1.3: Sistema de medida

Como resultado de este trabajo se disponía de un sistema de monitorización de redes ATM de alta velocidad con elevada precisión en la captura de datos y altamente configurable, el cual ha sido empleado en varios estudios de monitorización de redes privadas así como de la mencionada universidad [38].

Así pues, con la capacidad de esta herramiente de monitorización se ha dispuesto de un inmejorable sistema para la obtención de una medida de red con la cual llevar a cabo el trabajo propuesto.

Capítulo 2

Análisis macroscópico de tráfico

2.1 Introducción

En este capítulo analizaremos el tráfico presente en el enlace IP sobre ATM de la Universidad Pública de Navarra. Presentaremos su composición según protocolos y servicios para identificar cuáles son las principales componentes de este tráfico. Como veremos, los servicios de Internet más empleados se basan en TCP como protocolo de transporte. Así pués, las características del tráfico de Internet dependerán del comportamiento del tráfico TCP. Una vez localizado el servicio principal, tanto en cuestión de tráfico generado como por popularidad entre los usuarios, presentaremos una caracterización macroscópica del mismo orientada a los flujos entre pares de nodos de red que lo componen para en el capítulo siguiente analizar estos flujos centrando el estudio en los descriptores de ATM para tamaño de ráfagas (MBS) y tasa sostenible (SCR). Mediante este análisis podremos determinar los recursos por SVC (Switched Virtual Circuit) necesarios en un hipotético enlace ATM con capacidades de conmutación de flujos. Al mismo tiempo estaremos determinando la calidad de servicio percibida por una población representativa de usuarios mediante la tasa o throughput obtenida en los flujos iniciados por ellos. Comenzaremos presentando una caracterización macroscópica del tráfico en un enlace de IP sobre ATM, a lo largo de un año, en términos de protocolos y servicios más empleados, porcentajes y medias de tamaño y duración de transacciones.

A continuación, en la sección 2.2 presentamos las trazas de tráfico que se han recogido y un estudio de los protocolos y servicios presentes en las mismas. A continuación, en la sección 2.3 estudiaremos el funcionamiento dinámico de los flujos del servicio principal.

2.2 Análisis macroscópico

En esta sección presentaremos un análisis de las trazas de tráfico disponibles a nivel de tráfico acumulado separando por protocolos, servicios o sentido del enlace.

Veremos que la componente principal del tráfico se debe al protocolo de transporte TCP y que de todos los servicios soportados es el WWW el que presenta mayor número de transacciones y utilización del enlace, lo cual llevará a centrar el estudio posterior sobre sus flujos.

2.2.1 Trazas de tráfico

La tabla 2.1 resume las características de las trazas disponibles para este estudio, recogidas con el equipo IPmiser [38]. La segunda columna indica el número de paquetes presentes en la traza, recogidos de ambos sentidos del enlace ATM. Las dos últimas columnas indican los días de comienzo y finalización de las medidas, los cuales no suelen estar por completo en la traza. Como se puede observar las trazas disponibles se reparten en un intervalo de más de un año.

Traza	Paquetes (días) en millones	Comienzo	Final
NAV98	511 (47)	26Nov1998	11Ene1999
AGO99	122 (15)	13Ago1999	28Ago1999
OCT99	335 (24)	13Oct1999	5Nov1999
FEB00	280 (11)	11Feb2000	23Feb2000
ABR00	429 (15)	24Abr2000	10May2000
14N00	77 (1)	13Nov2000	15Nov2000
DIC00	460 (10)	12Dic2000	22Dic2000
NAV00	160 (4)	22Dic2000	26Dic2000

Tabla 2.1: Características de las trazas

Presentamos un primer análisis macroscópico en la tabla 2.2 donde se desglosa cada una de las trazas según el tráfico producido en cada sentido (saliente o US y entrante o DS) y presentamos el total de paquetes que han circulado así como la cantidad de bytes a nivel de enlace, es decir, incluyendo cabeceras hasta las de las propias celdas ATM. Esta medida de bytes nos permitirá calcular la utilización del enlace, el cual trabaja a 4Mbps a nivel ATM.

De estas estadística globales podemos apreciar una clara simetría en el número de paquetes que circulan en cada sentido. Dado que la capacidad de proceso de un Router IP viene determinada por el número de paquetes por segundo que es capaz de reenviar, ambos sentidos del tráfico introducen una carga similar de procesado en el router de acceso. Sin embargo, la cantidad de bytes del sentido entrante es entre 3 y 4 veces superior a la del saliente. Como veremos, el tráfico ascendente va a deberse a peticiones de información y a confirmaciones (ACKs) debidas a los mecanismos del principal protocolo empleado (TCP) o a la información solicitada por usuarios extrenos.

Los usuarios de las subredes de universidades y centros de investigación servidos por este enlace actúan principalmente como consumidores de información en lugar de como proveedores. Así, el sentido descendente es el más interesante a la hora de buscar técnicas de dimensionamiento y al mismo tiempo será aquél por el que circule el tráfico producido por los usuarios y por tanto el determinante de la calidad del servicio dada.

A partir de aquí, salvo que se diga lo contrario, centraremos el estudio exclusivamente sobre el canal de tráfico entrante a los usuarios.

Traza	Pqs US	Bytes US	Pqs DS	Bytes DS
NAV98	226M	42GB	285M	193GB
AGO99	61M	10GB	61M	48GB
OCT99	167M	27GB	168M	141GB
FEB00	133M	29GB	147M	119GB
ABR00	207M	50GB	221M	176GB
14N00	36M	9.8GB	41M	34GB
DIC00	228M	64GB	232M	188GB
NAV00	79M	33GB	80M	61GB

Tabla 2.2: Asimetría del tráfico de datos

2.2.2 Protocolos empleados

El único circuito virtual presente en el enlace ATM en estudio está configurado para transportar paquetes IP. Todos los protocolos presentes van a ser de la familia TCP/IP. En la tabla 2.3 tenemos el reparto del tráfico de cada traza según estos protocolos. Es claro el predominio de TCP, con más del 95% del tráfico en todas las mediciones salvo en la más antigua (NAV98), en la que se mantuvo un único flujo UDP de 300Kbps durante una semana entera. TCP es el protocolo empleado por la mayor parte de los servicios de Internet. UDP suele destacar en accesos con elevado uso de aplicaciones de video y audio-conferencia, aunque algunas de las aplicaciones de este tipo también son capaces de trabajar sobre TCP. Hoy en día estos servicios sobre UDP aún no representan un elevado porcentaje del tráfico, aunque en el futuro es posible que la situación cambie.

Los porcentajes de tráfico de control (ICMP) son reducidos, tal como es de desear. El resto del tráfico está formado por protocolos como IGRP (enrutamiento), IGMP o IP (gestión y tráfico IP sobre IP debido al túnel de MBone, con escasa utilización).

Traza	TCP	UDP	ICMP	Otros
NAV98	80.94%	14.73%	0.13%	4.2%
AGO99	96.4%	1.8%	0.5%	1.3%
OCT99	98.74%	0.96%	0.21%	0.09%
FEB00	98.13%	1.67%	0.11%	0.09%
ABR00	98.76%	1.08%	0.095%	0.065%
14N00	97.88%	2.04%	0.07%	0.01%
DIC00	98.23%	1.62%	0.1%	0.05%
NAV00	97.03%	2.34%	0.61%	0.02%

Tabla 2.3: Composición del sentido entrante en % de tráfico

2.2.3 Servicios más populares

El predominio de TCP era previsible dado que la mayor parte de servicios de Internet precisan entrega de datos en orden y confirmación de los mismos, lo cual les lleva a emplearlo como nivel de transporte. Veremos ahora cuáles son exactamente estos servicios (reconocidos por el puerto conocido del extremo servidor) y cuál o cuáles contribuyen de forma más significativa al tráfico actual.

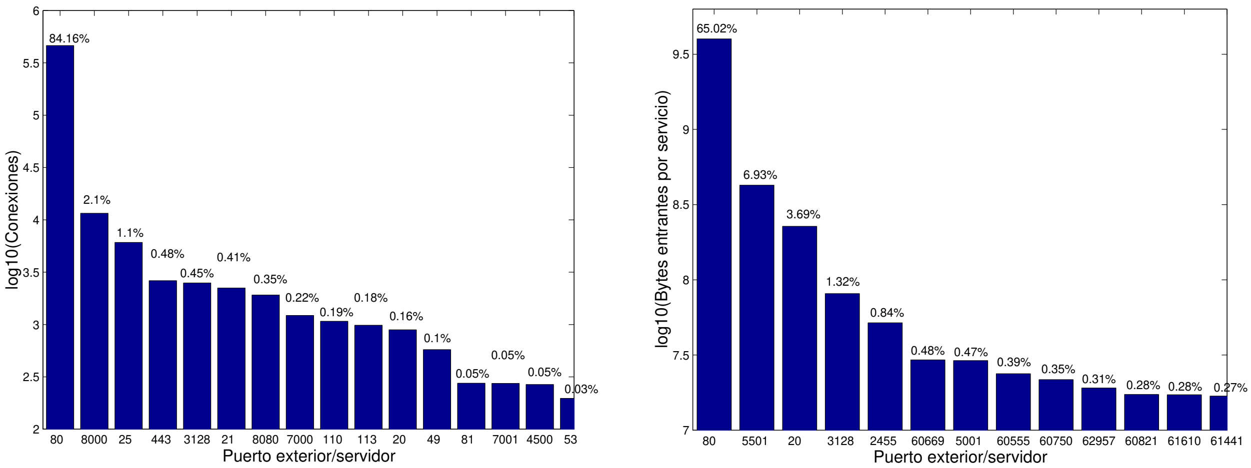
En la mayoría de los casos podemos reconocer los servicios por el puerto empleado por el servidor, que será la máquina que se encuentre en el lado externo del enlace. Aunque esperamos que la mayor parte de las conexiones TCP presentes en el enlace se deban a usuarios existirá un pequeño porcentaje debido a peticiones realizadas a los servidores presentes en las universidades. Este porcentaje es lo suficientemente pequeño en comparación con el debido a conexiones realizadas por los usuarios del enlace como para no desviar gravemente los resultados. Aún así ha de tenerse en cuenta su presencia para no contabilizar estas conexiones. Nos interesan las conexiones generadas por los usuarios internos, el problema radica en distinguirlas de las generadas por usuarios externos.

En principio podría optarse por tomar las conexiones que han sido iniciadas por direcciones internas. Es decir, aquellas conexiones en las que el primer paquete con la bandera para sincronización de números de secuencia (SYN) circuló hacia el exterior. Sin embargo, existen servicios en los que la conexión la inicia el propio servidor. Tal es el caso por ejemplo de la parte de FTP encargada de la transferencia de datos. En FTP el servidor suele emplear el puerto reservado a tal efecto (ftp-data) pero inicia él la conexión hacia el cliente tras recibir la petición por una conexión por otro puerto (ftp-ctrl). Sin embargo, existen situaciones en las que las transferencias de FTP se realizan sin emplear el puerto reservado a tal efecto, sino que se van utilizando puertos diferentes para cada conexión, los cuales son seleccionados entre los valores no reservados del rango de puertos TCP posibles (mayores que 1024). Esto último nos impide reconocer al servidor por el puerto empleado, dado que no siempre se cumplirá la regla de ser menor que 1024. Estas situaciones podrían repetirse con nuevos servicios.

Así pues tendremos un conjunto de conexiones en el enlace, debidas a FTP, iniciadas desde el exterior y empleando en ambos extremos puertos no reservados. FTP es un protocolo que tradicionalmente ha contribuido de forma importante al tráfico debido a las posibles transferencias de grandes ficheros [15], así que es necesario llevar a cabo todas estas consideraciones. Las características mencionadas harán difícil distinguir estos flujos FTP de las conexiones iniciadas por usuarios externos.

Seleccionamos a continuación un intervalo de un día de duración de una de las trazas, concretamente el 19/10/1999 contenido en la traza OCT99 (tabla 2.1) para tener una cantidad manejable de datos. En este día el 97.94% del tráfico entrante es debido al protocolo TCP y se registraron un total de 549592 conexiones que se establecieron correctamente (se completó el *three-way handshake*).

En la figura 2.1 podemos ver el número de conexiones y de bytes entrantes que se han dado según el puerto empleado por el extremo exterior (atención a la escala semilogarítmica). Los puertos que se observan en la figura 2.1 (b) engloban el 80.63% de los bytes entrantes. Los bytes servidos desde el puerto 80, el Web, representan el 65.02% de la carga total del sentido entrante y el 84.16% de las conexiones del enlace.



(a) (b)

Figura 2.1: a) Conexiones y b) bytes por servicio (19Oct1999)

Vemos que en cuestión de bytes por servicio aparecen una serie de puertos servidor anormalmente altos. Caben tres posibilidades. La primera, que esas conexiones sean iniciadas por usuarios externos, tal vez para introducir datos en servidores de las universidades (dado que tal cantidad de bytes, solo en confirmaciones, implicaría una cantidad ingente de datos en el otro sentido). La segunda posibilidad es que se trate de servidores de FTP que no emplean el puerto reservado ftp-data. La tercera posibilidad engloba cualquier otro servicio que emplee puertos no reservados (algún servicio nuevo, como ha sucedido con Hotline(puerto 5501) o Napster(6699)), aunque este caso es difícil por los enormes valores de puerto empleados. Para desvelar la hipótesis correcta empezamos por ver cuántas conexiones se dan de cada uno de esos puertos y analizamos los extremos de algunas de ellas. Aunque no se ve en la figura 2.1 es considerablemente grande la cantidad de puertos de valor alto (cercano a 65535) con estas características.

En el caso de los puertos 60669, 60555, 60750, 61610 y 61441 por ejemplo, hay menos de 3 conexiones para cada uno. Analizando las conexiones de forma individual vemos que para los cinco puertos ha sucedido practicamente lo mismo, varias conexiones generan una cantidad despreciable de bytes hacia servidores internos variados (Web, SMTP) y la mayor parte de los bytes es debida a una conexión dirigida al servidor de News interno, la cual introduce una actualización de las news que representa practicamente la totalidad de los bytes por ese puerto que se ven en la figura 2.1 (b).

El puerto 5001 corresponde a un servicio desconocido que realiza 13 conexiones transfiriendo una cantidad comparable de bytes en cada una. Todas las conexiones provienen del mismo usuario y se dirigen al mismo servidor externo pero no hemos logrado reconocer el servicio, aunque parece algún tipo de transferencia de ficheros. Otras, como el caso de los puertos 60957 y 60821 presentan las características típicas de transferencias de FTP que no emplean puertos reservados.

A continuación eliminamos todas las conexiones cuyo puerto interior sea menor que 1024, con lo que retiramos las conexiones a servidores internos (News, Web, ...) manteniendo conexiones entrantes que sí se deban a acciones de usuarios internos. Las conexiones restantes forman el 84.85% del tráfico entrante y podríamos considerarlas las conexiones *generadas* por los usuarios. Las conexiones eliminiadas son principalmente actualizaciones del servidor de News (NNTP, 6.67% del tráfico entrante), los mensajes de e-mail entrantes (SMTP, 4.33%) y peticiones web desde el exterior (HTTP 0.89%).

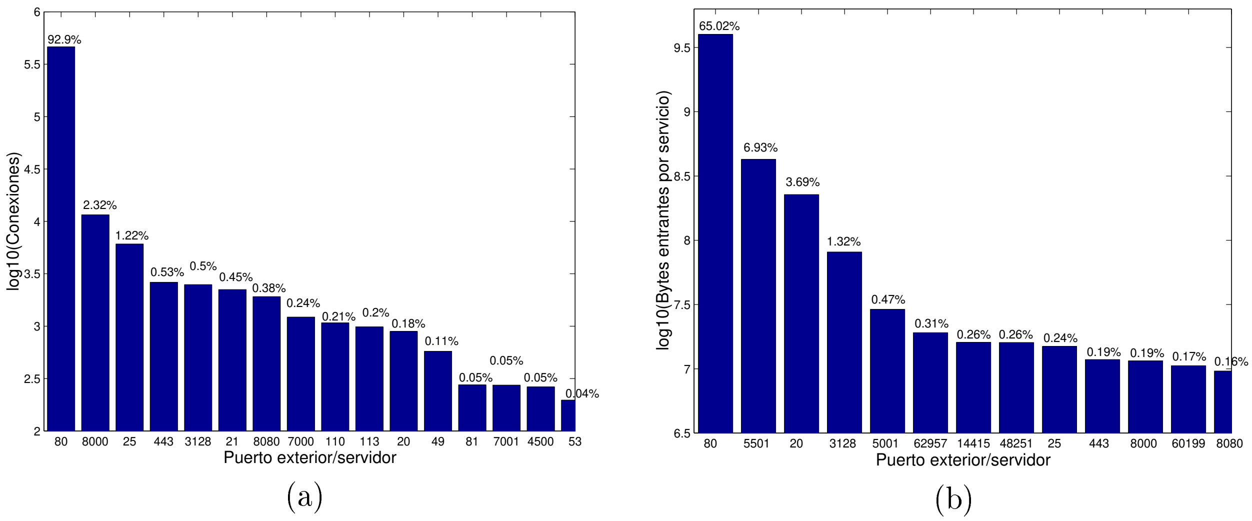


Figura 2.2: a) Conexiones de usuario y b) bytes por servicio (19Oct1999)

Vemos que no hay grandes cambios en los servicios principales. Aun así, si bien el tráfico debido al servicio Web representa el 65.02% del tráfico total del sentido entrante, respecto al tráfico entrante por conexiones generadas por los usuarios se convierte en un 76.6% (65.02/84.85). Es decir: *Tres cuartas partes del tráfico solicitado por los usuarios del enlace corresponde al servicio Web sobre el puerto 80.*

Otro detalle de la figura 2.2 digno de mención es que algunos servicios generan un gran número de conexiones con pocos bytes, como son FTP-ctrl(21), POP3(110), AUTH(113), DNS sobre TCP (53) o LOGIN(49). El puerto de FTP-ctrl se emplea para enviar comandos indicando las transferencias de FTP a realizar y tanto los listados de directorios como las transferencias se realizan empleando otro puerto, generalmente el FTP-data(20). El servicio AUTH suele emplearse de forma auxiliar al FTP para permitir la autentificación de los clientes. POP3 puede presentar muchas conexiones sin transferencias apreciables de bytes por comprobaciones de la presencia de nuevos mensajes de correo y el DNS sobre TCP se debe a actualizaciones entre servidores. Por otro lado, servicios como Hotline (5501) o FTP-data (20) consumen una parte considerable del ancho de banda del acceso con muy pocas conexiones (406 MBytes con solo 39 conexiones en el caso de Hotline). Hotline integra servicios de chat, noticias y transferencia de ficheros en la misma sesión; las transferencias de ficheros son la causa del gran número de bytes generado por las conexiones de este servicio.

La traza está dominada por el WWW con el 76.6% del tráfico y el 92.9% de las conexiones debidos a los usuarios. Como vemos en la figura 2.2 el WWW emplea el puerto 80 para conexiones directas pero valores de puerto servidor altos como 8000, 443, 3128 o 8080 también pueden atribuirse a este servicio, bien con servidores convencionales empleando otro puerto, bien a través de proxys o mediante transacciones seguras HTTPS(443). Tras el WWW se encuentran en generación de bytes los mencionados Hotline y FTP. Un pequeño porcentaje de los bytes se debe al envío de mail por SMTP(25) y existen aún una serie de puertos de valor elevado y difíciles de identificar. Tras llevar a cabo un análisis más detallado de las conexiones con puertos servidor 62957, 14415, 48251, o 60199 se concluye que, debido a los servidores a los que se conectan, parecen ser conexiones FTP que no han empleado el puerto ftp-data (20). En la medida habrá otras conexiones de FTP con características similares. Una característica típica de las conexiones de ftp-data es que en uno de los sentidos no se transfiere ningún byte de datos de aplicación (datos del nivel de transporte TCP), la transferencia es unidireccional. Se puede emplear esta característica para intentar reconocer las conexiones del servicio FTP-data, lo cual da un porcentaje de tráfico de "ftp-data" que asciende al 9.78% del tráfico total entrante.

Los resultados de este análisis macroscópico muestran que la mayor parte del tráfico del enlace IP sobre ATM se debe a conexiones TCP de Web. Se observan efectos parecidos en el resto de las trazas así como en otros estudios realizados sobre tráfico de puntos de Internet muy diferentes [13].

El Web es un servicio transaccional basado en el protocolo HTTP, que en su versión 1.0 es muy simple [39]. El navegador inicia la conexión con el servidor y envía la petición del fichero que desea recibir, la cual suele caber en un paquete. A continuación el servidor (si no hay error) envía por esta conexión el fichero y para indicar la finalización cierra la conexión. En la versión 1.1 [40] se permite la transmisión de varios ficheros en la misma conexión, aprovechando la localidad espacial y temporal en las peticiones y evitando así nuevos establecimientos y tener que repetir la etapa de slow-start en cada conexión.

Vemos que con el Web se engloba más del 90% de las transacciones solicitadas por los usuarios, las cuales producen más del 65/76% del tráfico el enlace (total/de usuarios). Debido a que este servicio es la componente principal centraremos el estudio sobre sus flujos, consiguiendo con ello la caracterización de una gran parte del tráfico del enlace y desde el punto de vista del usuario de prácticamente la totalidad de sus solicitudes de servicio.

2.3 Análisis de conexiones TCP

A continuación centraremos el análisis sobre los flujos del protocolo TCP. Describiremos de forma breve los flujos de los principales servicios y nos centraremos en el más popular, en la dinámica que sigue el protocolo y presentaremos las características de un estimador simple del retardo de ida y vuelta (RTT) entre cliente y servidor.

2.3.1 Estadísticas globales

En la tabla 2.4 se presentan estadísticas a nivel de conexiones de los servicios que más tráfico generan y los más populares en la traza del día 19/10/1999. Como se puede ver por los tamaños medios de las transferencias en cada sentido la mayoría de los servicios son claramente asimétricos, con predominio del sentido entrante. Por otro lado existen servicios con duración media por conexión acorde con los bytes transferidos (Web, Hotline...) mientras que otros (servicio de puerto 8000) destacan por mantener las conexiones durante largo tiempo entregando un reducido número de bytes a la red.

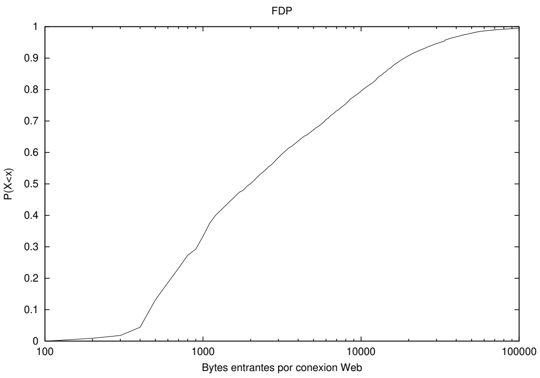
Respecto al servicio de uso mayoritario (el WWW) vemos que sus conexiones (que representan la mayor parte) son en media de pequeño tamaño (8.4 KB), con predominio del sentido entrante (descargas de información) y duraciones acordes a los tamaños.

Puerto servidor	E[Bytes in por cnx]	E[Bytes out por cnx]	E[duracion]
80 (http)	8.4 KB	1.4 KB	55 s
5501 (hotline)	10 MB	6 MB	2153 s
20 (ftp-data)	249 KB	8.2 KB	125.9 s
3128 (proxy-www)	32 KB	2 KB	20 s
5001 (desconocido)	2.1 MB	51 KB	1026 s
8000 (desconocido)	997 B	868 B	1671 s
25 (smtp)	2.4 KB	45 KB	48 s
443 (https)	4.4 KB	1.8 KB	28.4 s
21 (ftp-ctrl)	3.2 KB	1.6 KB	201.5 s
8080 (desconocido)	4.9 KB	1.1 KB	3.6 s
7000 (desconocido)	873 B	636	3.2 s
110 (pop3)	6.9 KB	1 KB	10.9 s
113 (auth)	346	455	3.7
49 (login)	463 B	542 B	0.6 s
53 (dns)	568	484	0.5 s

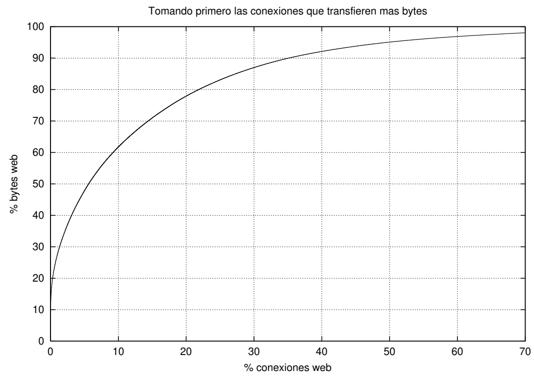
Tabla 2.4: Bytes y duración por conexión y servicio (19Oct1999)

En la figura 2.3 (a) vemos la función de distribución de los bytes entrantes de estas conexiones Web. Como se ve en la figura, el 99% de las conexiones transfieren menos de 70KB por el enlace entrante, el 95% menos de 32KB, el 88% menos de 16KB y el 76% menos de 8KB. Es decir, la mayor parte de las conexiones son cortas, pequeñas transferencias que por tanto pasarán la mayor parte de su tiempo en la fase de slow start.

Sin embargo, existe un conjunto de conexiones que transfieren una gran cantidad de bytes. Son pocas pero generan la mayor parte del tráfico. Así, el tráfico está dominado por un pequeño porcentaje de ráfagas muy grandes [41, 15]. El 90% de las conexiones Web (figura 2.3 (b)) generan el 38.25% de los bytes de conexiones Web, que se convierten en el 29.31% de los bytes en el sentido entrante del enlace debidos a conexiones de usuarios. El 10% restante de las conexiones, las mayores transferencias, generan 61.75% de los bytes web y por tanto el



(a)



(b)

Figura 2.3: a) Distribución de bytes por conexión Web. b) % del tráfico de las principales conexiones. (19Oct1999)

47.32% del tráfico de usuarios. Estas grandes transferencias se deben a descargas de software, videos o grandes imágenes que se dan con escasa frecuencia pero que aún así, debido a su tamaño, pueden dominar una traza. En el caso de un usuario residencial, conectado a través de un modem telefónico analógico, con tasas máximas de 56Kbps, estas transferencia serían mucho menos frecuentes debido a la enorme cantidad de tiempo necesario para completarlas. Sin embargo, los nuevos escenarios de tarifa plana y accesos de alta velocidad (HFC, xDSL...) pueden derivar en situaciones muy similares a la analizada aquí.

2.3.2 Dinámica de TCP

Dado que la intermitencia y la tasa de transferencia obtenidas en las conexiones va a depender del funcionamiento de TCP se hace necesario destacar las características de este protocolo que más interesantes puedan resultar para este análisis. Emplearemos un esquema en el que un extremo, el cliente, inicia la conexión y es el otro extremo, el servidor, quien envía datos al mismo.

Una conexión TCP comienza con el establecimiento, seguido por la fase de *slow-start*. El algoritmo de slow-start regula la transmisión desde el servidor mediante las confirmaciones (ACKs) enviados por el cliente. Cada paquete de ACK hace que se incremente en un segmento el tamaño de la ventana de congestión, lo cual permite enviar tantos segmentos como los que han sido confirmados más el aumento de la ventana, hasta que supere el tamaño negociado para la ventana de control de flujo. De esta forma la transmisión de datos sigue un esquema de *parar y esperar*, alcanzando el tamaño límite de ventana o el ancho de banda del cuello de botella en el camino de servidor a cliente. Las pérdidas de paquetes producirán fluctuaciones en este esquema alterando bruscamente el tamaño de la ventana y la progresión de incremento de la misma mediante técnicas de control de congestión. El tamaño final de la ventana depende de los buffers y velocidades en los enlaces que forman el camino entre ambos extremos, valor que se puede aproximar por el producto *retardo x ancho de banda*, es decir, el producto entre el ancho de banda del cuello de botella en el camino y el retardo de ida y vuelta (RTT o round-trip time) entre cliente y servidor [42].

El algoritmo de slow-start produce una progresión geométrica en el tamaño de las ráfagas transmitidas, con 2^i segmentos en el intervalo i-ésimo de duración igual al RTT, hasta que se alcanza un estado estacionario. La tasa de transferencia crecerá de forma análoga. Nos encontramos con que tanto la tasa de transferencia como la intermitencia dependerán del RTT, así como de la probabilidad de pérdida de paquete y del jitter que se introduzca en el camino entre cliente y servidor, el cual alterará la separación de los paquetes dentro de las ráfagas. Una estimación de la probabilidad de pérdida o del jitter requeriría una traza simultánea del extremo del cliente y del servidor, lo cual obligaría a restringirse a un conjunto de destinos que se pudieran controlar, situación en la cual se pierde la generalidad obtenida mediante accesos sin restricción. Con una traza en el extremo del cliente veremos que sí podemos obtener una estimación del RTT de cada conexión, que podremos emplear para determinar la separación típica entre paquetes de diferentes ráfagas.

2.3.3 Estimación del RTT

Dado que nos estamos centrando en las conexiones del servicio WWW y tomando la traza junto al extremo cliente, podemos realizar una estimación del RTT mediante el handshake inicial. Tomamos el tiempo que transcurre entre que se detecta el segmento de cliente a servidor con la bandera de SYN activa hasta el primer segmento proviniente del servidor (la confirmación al anterior) (figura 2.4). En el caso de servicios para los que el establecimiento se inicia desde el servidor externo esta estrategia no sería válida.

Mediante este método de medida el tiempo de respuesta del servidor no afecta, dado que el segmento SYN+ACK de respuesta es enviado por el nivel de protocolo TCP del kernel antes de notificar el establecimiento de la conexión a la aplicación. Por otro lado el primer paquete de la conexión (el primer SYN) puede sufrir un retardo mayor que el resto de paquetes debido a los efectos de las memorias cache de los routers intermedios, lo que se conoce como efecto de *ruta-fría* (cold-route). Sin embargo se ha comprobado que el aumento del retardo debido a este efecto es significativamente inferior al que introduciría el efecto de *servidor-frío* [43], en el cual se incluiría el tiempo necesario para *despertar* a la aplicación servidor, tiempo que no está incluido en este método de estimación.

La pérdida de cualquiera de los segmentos de SYN o el ACK del servidor al cliente produce una retransmisión a los 3 segundos [42]. Este tiempo es determinista dado que en la etapa de establecimiento de la conexión los agentes TCP de cliente y servidor no tienen ningún medio para realizar una estimación del RTT, así que deben basarse en un valor por defecto. Se puede detectar estos casos con alta fiabilidad debido a la baja probabilidad de que el RTT real sea superior a 3 segundos. Ademas podemos evitar en gran medida la desviación por retransmisión del SYN de cliente a servidor tomando el tiempo entre el último segmento de SYN enviado por el cliente y el primer SYN recibido del servidor.

Estimadores basados en la diferencia de tiempos entre segmentos de datos y sus confirmaciones no pueden emplearse en esta situación dado que la transferencia de datos se realiza en el sentido servidor a cliente y la traza está tomada en el extremo cliente. Se podría emplear el paquete con el comando GET que envía el cliente y medir el tiempo que transcurre hasta

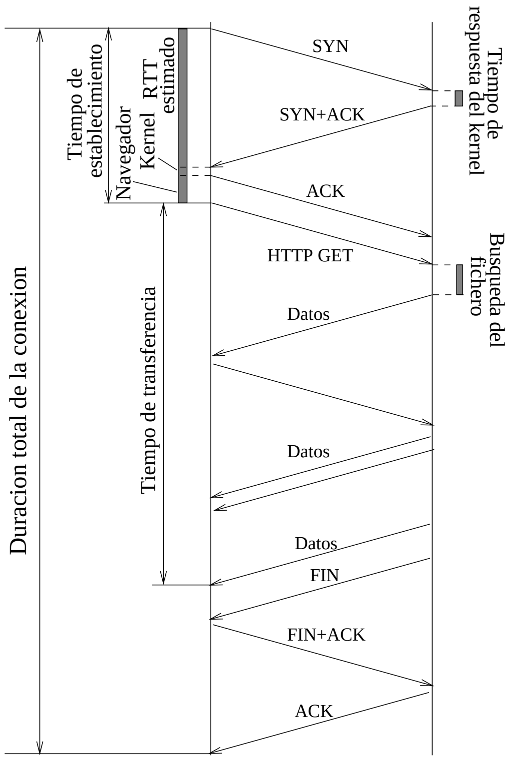


Figura 2.4: Fases de una conexión TCP (Web)

la confirmación por parte del servidor (evitando el efecto de ruta-fría), pero caso de que la implementación de TCP del servidor espere a tener datos que enviar para hacer piggybacking del ACK o emplee la común estrategia de *delayed ACK* esta medida se vería seriamente modificada por tiempos de búsqueda y acceso a disco para obtener el URL solicitado (cayendo en el efecto de servidor-frío) o esperas artificiales.

Este método nos da una estimación del RTT con un error reducido. Para reducir el error aún más se necesitaría por ejemplo emplear paquetes de ICMP (*echo request*) de cliente a servidor. Estos paquetes deberían enviarse simultáneamente a la conexión o justo después de ésta. Así los paquetes ICMP no encontrarían una ruta fría y la situación de congestión de la red sería comparable a la encontrada por la conexión, algo que no sería cierto en caso de realizar la medición mediante sondeo ICMP en otro intervalo de tiempo, dada la fuerte dependencia con la hora del día o la cantidad de tráfico. Dada la complejidad de implementación de esta medida no se ha considerado una técnica rentable.

En la figura 2.5 podemos ver la función de densidad de probabilidad del estimador del RTT entre las conexiones de Web de la medida del día 19/10/1999 (dentro de la traza OCT99). El rango comprendido en la figura engloba el 90% de las conexiones. Podemos definir al menos 3 zonas. La primera que llamaremos "RTT bajo" corresponde al primer lóbulo de la figura. Valores tan bajos de RTT se deben a accesos dentro de la propia RedIRIS. La segunda zona, que llamaremos de "RTT medio" comprende los siguientes lóbulos inferiores

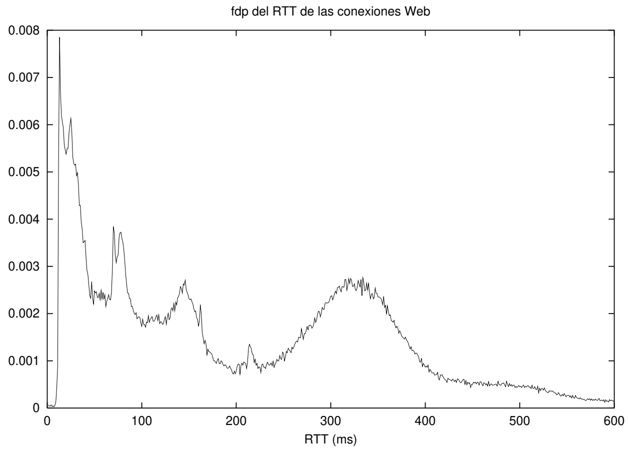


Figura 2.5: Densidad de probabilidad del RTT (19Oct1999)

a unos 200ms y corresponden generalmente a conexiones nacionales o dentro de la red del Ten155. La tercera zona, que llamaremos "RTT alto" engloba al último lóbulo, con forma de campana y que corresponde a las conexiones internacionales, generalmente con EE.UU. La forma de este lóbulo resulta muy interesante dado que sugiere que al atravesar los paquetes un gran número de saltos con retardos diferentes y aleatorios (tal vez independientes) se da un efecto de Teorema Central del Límite (agregación de un gran número de contribuciones independientes) con convergencia a una gaussiana.

2.3.4 Impacto de la etapa de slow-start

Hemos visto que una conexión TCP comienza en slow start hasta alcanzar un estado estacionario. Durante esta etapa de slow start la tasa de transferencia de la conexión es inferior a la máxima que se alcanzará para evitar una posible congestión de los enlaces. El tiempo que la conexión permanezca por debajo de la tasa de transferencia que se le permite alcanzar es el mínimo entre el necesario para exceder el ancho de banda disponible y el correspondiente para abrir la ventana de congestión hasta sobrepasar la ventana de control de flujo. Los tamaños más típicos de ventana de control de flujo se encuentran en 8, 16 y 32KB. Para determinar el impacto de esta fase inicial en la duración total de la conexión en la figura 2.6 podemos ver la distribución del número de paquetes por conexión TCP Web. Suponiendo una MTU de la red de 1500 bytes (correspondiente a Ethernet) una ventana de 8KB son menos de 6 paquetes de este tamaño y una de 32KB unos 22 paquetes.

Vemos que la mayor parte de las conexiones son cortas. El 80% de las conexiones tienen menos de 11 paquetes y como se ha visto (figura 2.3) menos de unos 16KB (lo cual corresponde al tamaño medio de paquete mencionado). Con tan pocos paquetes la mayor parte de la duración de la conexión va a transcurrir durante la etapa de slow-start.

Para analizar la evolución temporal de las conexiones TCP necesitamos examinar cómo se distribuyen estos paquetes en la duración de las mismas. Definimos para ello la variable X_i

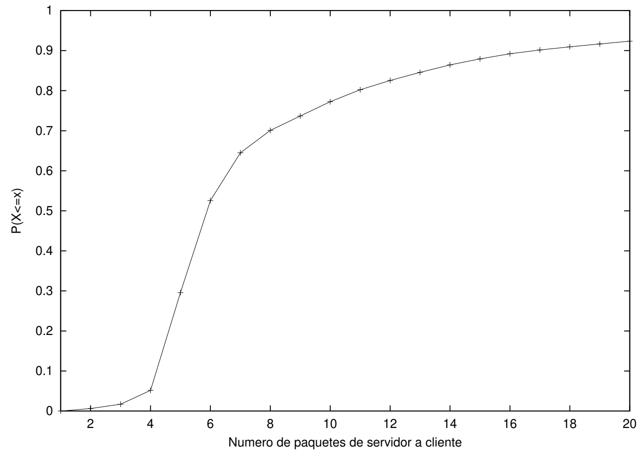
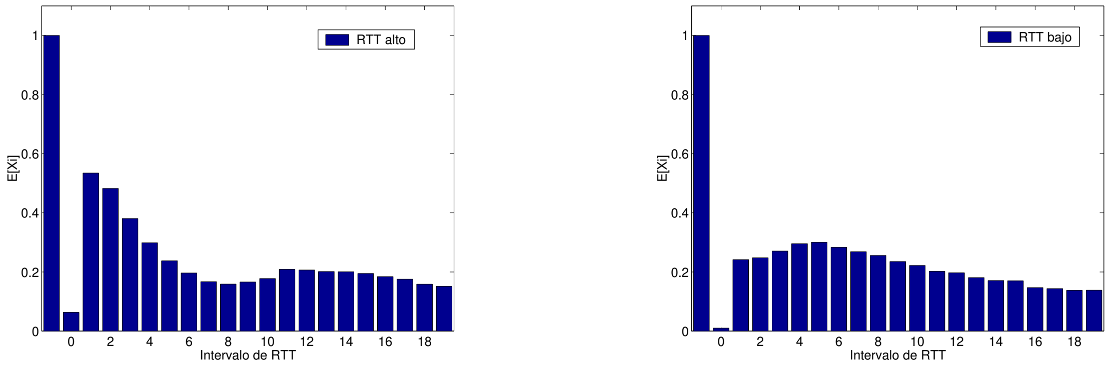


Figura 2.6: Paquetes por conexion

que representa el número de paquetes que se han recibido en el intervalor i-ésimo de anchura temporal igual al RTT de la conexión. Por ejemplo, para una conexión de 10 segmentos TCP, en el caso ideal sin pérdidas, con confirmación de cada paquete y ventana de control de flujo lo suficientemente grande la variable X_i tomaría los valores 1, 2, 4 en los primeros tres intervalos X_1, X_2, X_3 de anchura el RTT, lo cual corresponde a la fase de slow start y en el cuarto intervalo se transmitirán los 3 paquetes restantes.

Para una situación real en la figura 2.7 vemos los valores medios de estos X_i para las conexiones Web del día de tráfico considerado. En el intervalo -1 hay siempre un solo paquete por la forma en que está definido el RTT (el paquete de SYN). Encontramos una bajada considerable en el siguiente intervalo (intervalo 0), el cual corresponde a la petición (PDU de GET URL) de cliente a servidor. En este intervalo, viendo la figura 2.4, es lógico que no existan paquetes en el sentido entrante, tan solo debido a jitter o error de estimación del RTT pueden aparecer en algunas ocasiones paquetes en este intervalo, lo cual es más probable en el caso de RTT alto por dar los errores un mayor márgen de variación. A partir del intervalo de RTT número 1 vemos cómo va disminuyendo el número de paquetes que encontramos en cada intervalo, en vez de aumentar, como indica slow-start. Esto puede deberse a que muchas conexiones no llegan a tener suficientes paquetes para durar hasta un intervalo de RTT elevado. Sin embargo, dado que el 80% tienen al menos 11 paquetes, el 80% de las conexiones deberían llegar hasta el intervalo 4, habiendo en cada uno el doble de paquetes que en el anterior y dominando el comportamiento de la gráfica, pero esto no se observa en la gráfica. El descenso de $E[X_i]$ que se da demuestra que en situaciones reales las conexiones *no siguen la progresión geométrica impuesta por el slow start.*

Para el caso de conexiones con un RTT alto, el mayor número de saltos entre cliente y servidor introduce una variabilidad aún mayor en la llegada de paquetes. En general no se da una progresión geométrica como debería provocar el algoritmo de slow start sino que la dinámica de la red condiciona en mayor medida el proceso de llegadas. Debido a esta alteración del proceso de llegadas la formación de ráfagas en el enlace de acceso se va a ver alterada respecto a lo especificado en el algorítmo. Veremos en el siguiente capítulo cómo resultan los tamaños de las ráfagas una vez que han pasado por la red.



(a) (b)

Figura 2.7: Paquetes por RTT para a) RTT alto o b) RTT bajo

2.4 Conclusiones

Como hemos visto, el tráfico típico de un enlace de usuarios de Internet está dominado por el protocolo TCP y sobre él, el servicio que más tráfico genera en la actualidad es el transporte de páginas Web, las cuales incluyen desde pequeñas figuras a ficheros de aplicación y contenidos multimedia.

El funcionamiento cliente-servidor de TCP nos ha permitido definir un estimador del retardo extremo a extremo de las conexiones y en función de él estudiar el comportamiento dinámico de las mismas. Hemos visto que las conexiones de Web son en su mayoría cortas (si bien hay unas pocas transferencias de gran tamaño) con lo que la etapa de slow-start tiene una gran importancia en la duración de las conexiones. Las conexiones Web pasan casi toda su duración en la etapa de slow-start y por tanto su evolución y velocidad de transferencia depende del comportamiento de este algoritmo. Como slow-start está controlado por la realimentación de confirmaciones por parte del cliente y éstas se realizan de acuerdo a la escala de tiempos impuesta por el retardo extremo a extremo se concluye que este parámetro va a ser uno de los principales determinantes de la calidad de servicio obtenida en estas transferencias. Sin embargo, la progesión de los bytes entregados en cada uno de los intervalos de tiempo de duración igual al retardo extremo a extremo no coincide con lo que se espera del algoritmo de slow start. La red introduce un *jitter* a la llegada de paquetes que rompe el esquema exponencial de slow start.

Capítulo 3

Modelo estocástico de conexiones TCP

3.1 Introducción

La creciente demanda de nuevos servicios de Internet está provocando la sustitución de las antiguas troncales de redes de datos por otros más modernos basadas en tecnologías IP de alta velocidad. Sin embargo, la tecnología más adecuada a emplear sigue siendo una cuestión por resolver. Uno de los candidatos más fuertes en estos últimos años ha sido el Modo de Transferencia Asíncrono (MTA) o ATM (Asynchronous Transfer Mode). Una de las principales ventajas de las redes ATM es su capacidad para la reserva dinámica de recursos, lo cual convierte a las redes ATM en ideales para tráfico de alta variabilidad como flujos de video, voz o datos. Sin embargo, el estado del arte en equipos ATM revela que las capacidades de reserva dinámica que se ofrecen son limitadas. De hecho, la mayor parte de los operadores de equipos ATM emplean tan solo circuitos virtuales de tasa constante (CBR) o *best-effort* (UBR). Mientras que la clase de servicio CBR ofrece una reserva estática de recursos, la clase UBR sirve para el transporte *best-effort* de celdas ATM. Dado que se emplea comúnmente una configuración best-effort para enlaces IP sobre ATM están surgiendo una serie de alternativas a ATM, tales como POS (Packet-over-Sonet) o IP sobre WDM (Wavelength-division multiplex), las cuales ofrecen este servicio sin la sobrecarga de segmentación y reensamblado presente en ATM. Así pues, el éxito de ATM como la tecnología empleada en la troncal de la próxima generación de Internet depende de su capacidad para ofrecer no solo servicios best-effort sino también otros que aseguren una calidad de servicio extremo a extremo.

De hecho, a medida que la tecnología y los estándares sobre ATM han ido evolucionando han empezado a aparecer tecnologías que emplean circuitos virtuales conmutados (SVC) para dar calidad de servicio para cada flujo de datos. Estas técnicas permitirían asignar recursos a flujos de datos independientes o a conjuntos de flujos de Internet, multiplexándolos en un SVC de acuerdo a características como direcciones origen y destino, subredes origen y destino, protocolo o puertos que describan el servicio. La reserva de recursos para estos SVCs precisa la caracterización de estos flujos de tráfico. Esta caracterización debería hacerse en base a parámetros de descripción del tráfico de ATM, dado que generalmente será la tecnología que soporte la calidad de servicio. Los dos parámetros más típicos en ATM son el tamaño máximo de ráfaga (Maximun Burst Size, MBS) y la tasa de transferencia sostenible (Sustainable Cell

Rate, SCR), los cuales especifican la intermitencia y el throughput de un flujo de tasa variable (VBR). Una característica importante de estos descriptores de tráfico es que la tasa sostenida de transferencia es también un indicador de la calidad de servicio percibida por el usuario para el caso de servicios transaccionales tales como el Web o el FTP. Así pues, la intermitencia y la tasa de transferencia son útiles como modelo de conexión dado que no solo determinan las características del flujo en términos de tráfico ATM sino que también sirven para evaluar la calidad de servicio percibida.

Como hemos visto, los servicios de Internet más empleados se basan en TCP como protocolo de transporte. Así pues, las características del tráfico de Internet dependerán del comportamiento del tráfico TCP. Este comportamiento depende de un gran número de factores como pueden ser la presencia de congestión en el camino entre cliente y servidor, las ventanas de transmisión negociadas en el establecimiento, el tamaño máximo del segmento, el retardo extremo a extremo, la versión de TCP empleada (Reno, Tahoe, Vegas), con diferente implementación de políticas de *delayed ACK* o de mecanismos de recuperación de ventana. Debido a la enorme complejidad de TCP gran parte de los estudios que se realizan se centran en aspectos específicos de TCP y su efecto en escenarios de simulación, configuraciones específicas o rutas concretas de cliente a servidor. Sin embargo, las conexiones reales de Internet se encuentran seriamente afectadas por tráfico interferente, el cual es complicado de introducir en modelos de simulación. Además, los usuarios navegan por la red de forma aleatoria, por lo que no podemos derivar características de calidad de servicio del estudio de entornos experimentales con caminos seleccionados entre cliente y servidor. Para caracterizar la calidad de servicio percibida por el usuario necesitaremos medidas de tráfico real en entornos sin restricciones de destinos, rutas o servidores.

En los últimos años han aparecido gran cantidad de estudios sobre el protocolo TCP [44, 45, 46, 47, 48, 49, 50, 51, 52, 53]. La mayor parte de estos trabajos intentan describir el funcionamiento de aspectos concretos de TCP o sus prestaciones en ciertos entornos de simulación. En [54] se emplea una traza de tráfico para analizar la dinámica de TCP en una red real. Sin embargo, el estudio se limita a aspectos concretos de la relación entre la compresión de ACKs y la pérdida de segmentos, así como sus efectos sobre las prestaciones de TCP. Por otro lado, partiendo de trazas en [41] y [15] se describen estadísticas por flujo TCP como la asimetría, número de paquetes y tiempo entre paquetes, pero sin analizar parámetros de calidad de servicio y sin la presencia del servicio más popular en la actualidad: el WWW. Respecto a este servicio la mayor parte de los estudios sobre él están basados en registros de servidores o proxies [55, 56]. En [16] se emplean trazas del extremo del cliente para estudiar las características de las sesiones WWW tales como duración, tamaño y aspectos del comportamiento del usuario. Todos estos estudios sobre el WWW se centran en estadísticos del nivel de aplicación, debido en parte al tipo de registros en que se basan, con lo que no se estudia la dinámica de TCP. Además, los modelos propuestos alcanzan una considerable complejidad en su intento de predecir el comportamiento de TCP en el mayor número de situaciones, según el tamaño de la ventana, el tamaño del segmento, las pérdidas, tamaños de buffers, retardo extremo a extremo, etc.

Sin embargo, el análisis de conexiones TCP reales requiere un compromiso entre capturar el comportamiento de los flujos con un modelo estocástico o estudiar un aspecto concreto del funcionamiento de TCP en un escenario de simulación. La caracterización de los flujos de TCP en un escenario de red sin restricciones implica necesariamente el empleo de modelos estocásticos, dado que no es posible modelar las conexiones de manera determinista, debido al gran número de factores que influyen en su evolución. Tal modelo estocástico debe proveernos de una caracterización de los flujos TCP en función de parámetros de red, lo cual permitirá el control y dimensionamiento de la misma. Buscaremos además una caracterización simple, basada en el menor número de parámetros necesarios.

Con estos objetivos, a continuación presentamos un modelo estocástico para los flujos de tráfico TCP. Este modelo captura las características de intermitencia y tasa sostenida (throughput) de dichos flujos, parámetros muy relacionados con los descriptores de ATM Tamaño Máximo de Ráfaga (MBS) y Tasa Sostenible de Celdas (SCR, que generalmente mediremos en bits por segundo) [57]. Estudiaremos la posibilidad de emplear una estrategia de conmutación de flujos con reserva de ancho de banda y buffer en los nodos de la red.

Para trabajar con una traza de tamaño manejable y mantener coherencia con el capítulo anterior, si no se especifica lo contrario los análisis estarán realizados sobre la medida correspondiente al 19 de Octubre de 1999, perteneciente a la traza OCT99 (tabla 2.1).

A continuación, en la sección 3.2 presentamos el modelo propuesto y en las secciones 3.3, 3.4 y 3.5 sus parámetros, que son respectivamente el número de conexiones simultáneas, la tasa de transferencia y la intermitencia de las mismas. Finalmente, en la sección 3.6 presentaremos los mismos resultados para una traza más extensa como validación del modelo.

3.2 Parámetros del modelo

Partimos del proceso A_n, que representa el total acumulado de celdas que han llegado en el intervao $[0, n]$. Se dice que este proceso sigue una restricción (σ, ρ) si se cumple (3.1)

$$\sum_{n=k}^{m} A_n \leq \rho\,(m-k+1)+\sigma \tag{3.1}$$

para todo k, m tales que $k < m$. El parámetro σ caracteriza la intermitencia y ρ el throughput del flujo A_n. Se puede demostrar que el proceso de llegadas A_n, ante un sistema de servidor único de tasa ρ y capacidad del buffer de σ celdas, no presenta pérdidas.

Los procesos con restricciones (σ, ρ) están directamente relacionados con las técnicas de conformado de tráfico mediante *leaky bucket*. El proceso a la salida de un leaky bucket con tasa de tokens ρ, tamaño del buffer de tokens σ y encolado del tráfico no-conforme (lo que se llama un regulador (σ, ρ)) sigue una restricción (σ, ρ) [57].

Otra de las características interesantes de los flujos con restricción (σ, ρ) es la propiedad de multiplexación. Si tenemos k procesos $A_n^{(1)}, A_n^{(2)}, ..., A_n^{(k)}$ con restricciones $(\sigma^{(1)}, \rho^{(1)})$, $(\sigma^{(2)}, \rho^{(2)})$,... $(\sigma^{(k)}, \rho^{(k)})$ se demuestra que [58] el proceso $A_n = \sum_{i=1}^{k} A_n^{(k)}$ sigue una restricción $(\sigma\prime, \rho\prime)$ donde $\sigma\prime = \sum_{i=1}^{k} \sigma^{(i)}$ y $\rho\prime = \sum_{i=1}^{k} \rho^{(i)}$.

Estos dos parámetros no solo describen un modelo para flujos sino que en el caso de servicios transaccionales sobre TCP como el Web o el FTP el parámetro de throughput va a estar estrechamente relacionado con la calidad de servicio percibida por el usuario.

En el resto del capítulo trabajaremos exlusivamente con las conexiones Web debido a que son el 92.9% de las conexiones generadas por los usuarios y para ellas podemos emplear el estimador de RTT definido en el capítulo anterior.

Para este modelo, en primer lugar deberemos comprobar que un sistema real de conmutación pueda mantener la información de estado necesaria para el número de flujos simultáneos que se vayan a dar. Gracias al creciente esfuerzo volcado en soluciones basadas en conmutación de flujos IP, tag switching o MPLS, hoy en día puede plantearse el siguiente paso, orientado hacia la reserva de recursos por flujo independiente. Así pues, tras verificar esta condición estudiaremos el comportamiento estadístico de la tasa media de estos flujos TCP, así como la intermitencia que se da dentro de las mismas. Finalmente verificaremos los resultados obtenidos empleando otra traza aún más extensa.

3.3 Circuitos necesarios

En un hipotético escenario con conmutación de flujos es muy importante conocer el número de flujos simultáneos que se tienen que conmmutar. Este parámetro afectará por ejemplo a los tiempos empleados para determinar el flujo al que pertenece un paquete (dentro del módulo clasificador) y por tanto al tiempo que un paquete va a pasar en el nodo. Por otro lado la tecnología empleada para la conmutación de estos flujos es probable que tenga una limitación del número de flujos que puede distinguir. Por ejemplo, en el caso de una implementación basada en ATM el número máximo de Circuitos Virtuales, y por tanto de flujos individuales, que se puede distinguir es $2^{16} = 65536$.

La figura 3.1 muestra el número de conexiones Web generadas por los usuarios que se encontraban establecidas en un instante cualquiera de tiempo. Vemos que el valor más grande que se da es de unas 1500. Si consideramos las conexiones generadas por los usuarios de cualquier servicio vemos que el máximo asciende a 4500. El puerto servidor de las conexiones que en esas primeras horas elevan la cifra a 4500 es el de valor 8000, con lo que no es un valor *asignado* y tampoco se ha encontrado entre los puertos *registrados* [59]. Este servicio no ha podido ser reconocido debido a la desaparición del servidor pero por el valor del puerto (80XX) puede ser un servidor Web. De todas formas, estas conexiones no representan más del 2.36% de las conexiones de usuarios y el 0.19% del tráfico del enlace. Las conexiones de este servicio generan poco tráfico pero se mantienen establecidas durante un tiempo considerable. Este patrón hace necesario un criterio de selección de aquellas conexiones a conmutar, bien por servicio, número de paquetes, par de hosts o subredes [60, 61, 62, 63].

Vemos que las tecnologías actuales que trabajan con direcciones para los circuitos en torno a los 16 bits son más que suficientes para enlaces de acceso con unos centenares o miles de usuarios. El caso de enlaces WAN de alta capacidad puede dimensionarse en base a una aproximación de linealidad en la multiplexación del número de conexiones simultáneas.

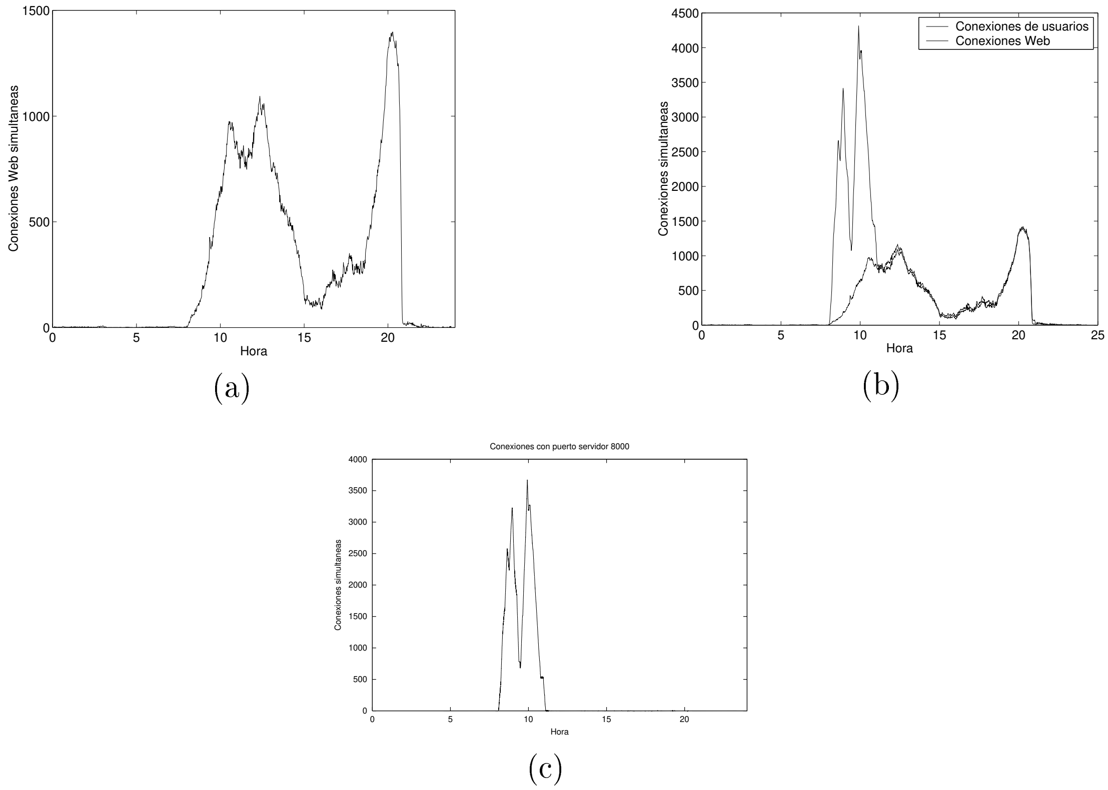


Figura 3.1: Conexiones simultáneas. a) Conexiones Web. b) Todas. c) Conexiones del puerto 8000.

3.4 Tasa de transferencia de las conexiones

Desde el punto de vista de la calidad de servicio percibida por el usuario los parámetros más típicos que se han empleado con servicios transaccionales son la duración de la transacción y el throughput o tasa de transferencia. De estos dos parámetros el último permite un modelado más simple dado que el primero depende de otro parámetro: la cantidad de bytes transferidos.

La tasa media de transferencia de una conexión puede calcularse simplemente dividiendo los bytes transferidos por la duración de la conexión. Podemos distinguir dos intervalos diferentes en una conexión Web TCP. En el primero, tras la solicitud por parte del usuario, se establece la conexión con el destino remoto, iniciada por dicho cliente, y se envía la petición. En el segundo intervalo se reciben los datos transmitidos por el servidor. Estos dos intervalos pueden verse en la figura 2.4. Así pues podremos definir dos mediciones de throughput: La primera medición relaciona los bytes transferidos a nivel de aplicación de usuario (datos del nivel TCP) con la duración total de la conexión. Esta será la tasa, y por tanto la Calidad de Servicio (QoS), percibida por el usuario. La segunda medición del throughput relaciona los bytes transferidos a nivel ATM con la duración de la fase de transferencia de la conexión. Esta segunda medida es más cercana a la tasa media de transferencia de tráfico que va a afectar a la utilización del enlace ATM.

$$SCR_{usuario} = \frac{\text{Bytes de datos TCP}}{\text{Duración total}}, \quad SCR_{red} = \frac{\text{Bytes ATM}}{\text{Duración de la transferencia}} \tag{3.2}$$

La diferencia entre estas dos mediciones de throughput se va a deber principalmente a dos aspectos. Por un lado la cantidad de bytes manejada es diferente en cada definición. Los bytes a nivel ATM incluirán las cabeceras de los niveles IP y TCP, la cola y el padding del nivel AAL5 así como las cabeceras de todas las celdas ATM. Así, por ejemplo, un paquete con 1460 bytes de datos TCP tendrá 1696 bytes a nivel ATM, un 16% más. Debido a la influencia de cabeceras de protocolos esperamos un decremento en torno a un 16% del throughput a nivel de usuario frente al throughput a nivel de red.

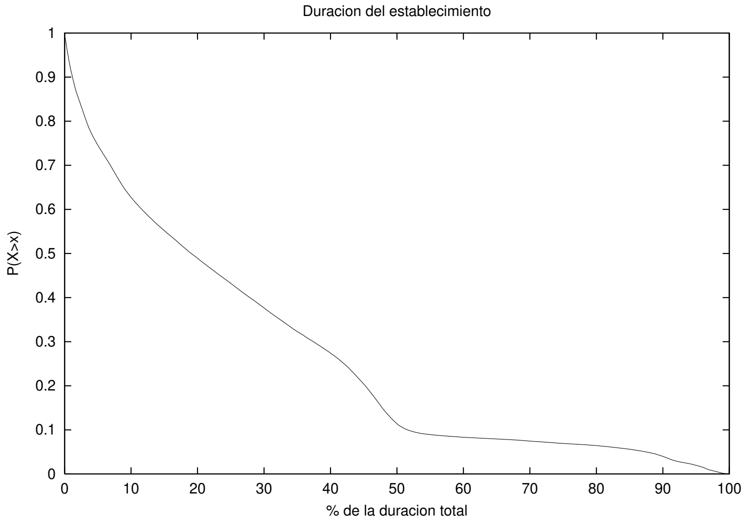


Figura 3.2: Tiempo de establecimiento

Por otro lado, el tiempo empleado en el establecimiento puede ser apreciable y ésta será la principal diferencia entre la duración del flujo empleada en cada medida. En la figura 3.2 se muestra la función de supervivencia del porcentaje de tiempo de la duración de la conexión que representa el establecimiento de la misma. Como se puede observar en el 27% de las conexiones el establecimiento representa más del 40% de la duración de la conexión. En media el 25% de la duración de las conexiones se pierde en la fase de establecimiento, en la cual no hay transferencia de datos útiles al usuario.

Así pues la tasa a nivel de red va a ser superior a la percibida por el usuario debido a dos factores: por un lado se transmiten más bytes de los que realmente recibe el usuario y por otro la transferencia se realiza en menos tiempo que el que aprecia el cliente.

Otro factor importante en la duración de las conexiones es el tiempo empleado en la finalización de las mismas. Al igual que el tiempo de establecimiento durante este intervalo no hay transferencia de datos, así pues no tiene sentido incluirlo en el cálculo del SCR como carga añadida al enlace. Por otro lado, llegado este punto se ha completado la transferencia del fichero, de forma que se dispone de todos los datos solicitados por el usuario y le pueden ser entregados. En general las implementaciones de TCP no hacen esperar a la aplicación a que se finalice el cierre de la conexión sino que le notifican el comienzo del mismo y es el agente del protocolo quien se encarga de llevarlo a cabo, sin informar a la aplicación de cuándo se concluye. Así pues tomaremos tanto la duración desde el punto de vista del usuario como desde el punto de vista de la red sin incluir este tiempo de finalización.

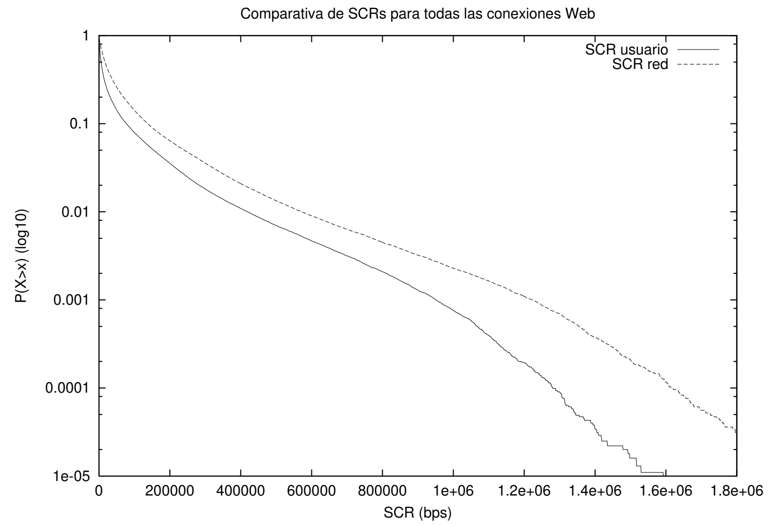


Figura 3.3: Diferencia entre el SCR de usuario y de red

En la figura 3.3 podemos ver finalmente la forma de la función de supervivencia de ambos throughputs. Podemos apreciar la diferencia entre los dos, como era de esperar la tasa a nivel de red alcanza valores superiores con una mayor probabilidad. Teniendo en cuenta la escala semilogarítmica de la gráfica destaca la semejanza de ambas con la distribución de una variable aleatoria de Weibull. Si bien una variable exponencia sigue una recta en escala semilogarítmica una variable aleatoria de Weibull está deformada mediante una potencia diferente de la unidad en la variable tal como se ve en la ecuación 3.3.

$$P(X < x) = 1 - e^{-\alpha x^{\beta}}, \quad x \geq 0, \quad \alpha, \beta > 0 \tag{3.3}$$

Una variable aleatoria de Weibull tiene una cola más pesada que una exponencial pero aún así todos sus momentos van a ser finitos, a diferencia de las distribuciones con cola de tipo Pareto.

Como comprobación gráfica, si representamos $log(-log(P(X > x)))$ frente a $log(x)$ deberíamos encontrar una recta en caso de que la distribución siga esta distribución (3.4). En la figura 3.4 vemos cómo estas distribuciones se podrían aproximar bastante bien con variables aleatorias de Weibull, obteniendo un modelo más complejo que en el caso de distribuciones exponenciales. Para la estimación de los parámetros de la variable emplearemos regresión lineal en este tipo de representaciones tomando puntos equiespaciados en el eje x. En este caso obtenemos unos parámetros de $\alpha = 9.56e - 3$ y $\beta = 0.4847$ para el SCR de usuario y $\alpha = 1.99e - 3$ y $\beta = 0.59$ para el SCR de red.

$$ln(-ln(P(X > x))) = -\alpha ln(x) \quad (3.4)$$

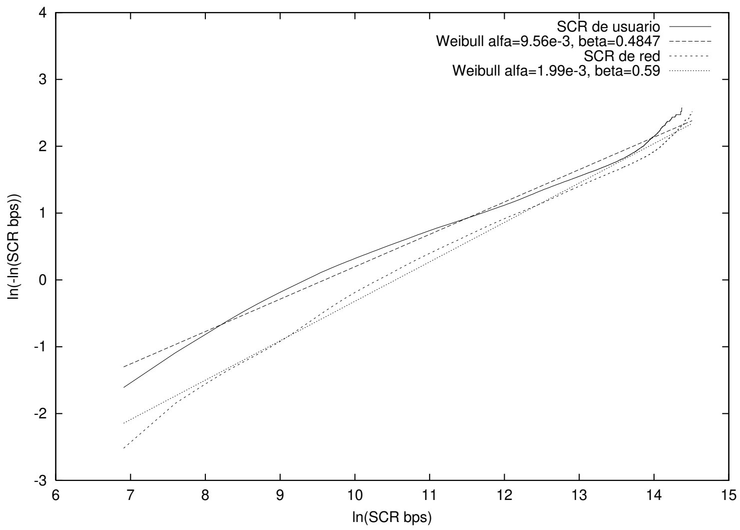


Figura 3.4: Representación lineal para distribuciones de Weibull

3.4.1 Calidad de servicio según el RTT

Hemos visto que el tiempo de establecimiento tiene un efecto apreciable en la tasa percibida por el usuario. Por otro lado sabemos que cuanto mayor sea el RTT de una conexión ésta evolucionará más lentamente y el valor final de tasa de transferencia que pueda alcanzar será menor. Con esto esperamos encontrar peores tasas de transferencia en las conexiones con mayor RTT siempre que los tamaños de los ficheros transferidos sean comparables. Es razonable que el tamaño del fichero transferido sea independiente del RTT dado que el primero viene determinado por el interés del usuario y la oferta en la red y el segundo por el estado de esta última. Sin embargo, también es posible que en aquellas situaciones en las que la tasa de transferencia sea menor los usuarios tiendan a evitar los ficheros grandes, en busca de servidores con mejor acceso. Así pues es conveniente comprobar primero esta hipótesis. En la figura 3.5 vemos la distribución de los tamaños de fichero transferidos según el rango de RTT. Vemos que la forma de la distribución es similar en el rango donde está comprendido el 99% de las conexiones.

Dado que las distribuciones de los ficheros son similares esperamos que según el RTT las distribuciones de los SCRs se reescalen teniendo en cuenta el aumento de la duración de las mismas. En general, según el tamaño del fichero podríamos aproximar el número de intervalos de RTT que tardará en transferirse y con estos calcular la duración real y la tasa de

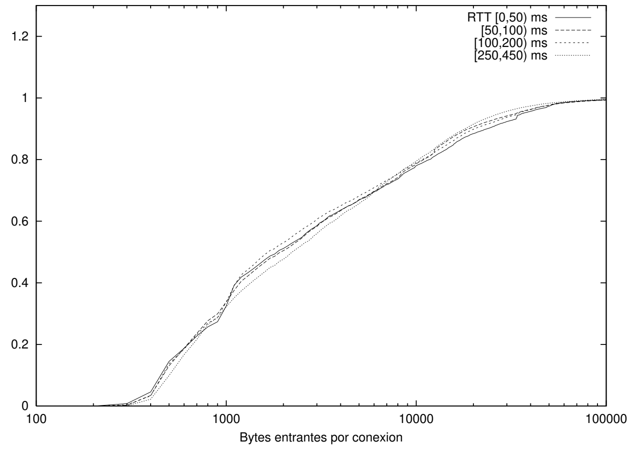


Figura 3.5: Bytes transferidos según el RTT de las conexiones

transferencia según el RTT de la conexión. Sin embargo, dado que como hemos comprobado en el capítulo anterior la conexión no va a seguir una evolución exponencial como marca el algoritmo de slow-start, debemos recurrir al análisis de los experimentos para dar con este valor. En la figura 3.6 vemos cómo tiene lugar el escalado debido al aumento del RTT. El efecto es similar tanto en el SCR de usuario como en el de red. Al reducir el número de experimentos para cada clasificación reducimos la calidad de la aproximación a la distribución de probabilidad, sin embargo, en general se mantiene la forma de la distribución que hemos visto para el conjunto total de conexiones (Weibull) dado que la curvatura de las funciones de supervivencia en escala semilogarítmicas las alejan considerablemente de una aproximación exponencial.

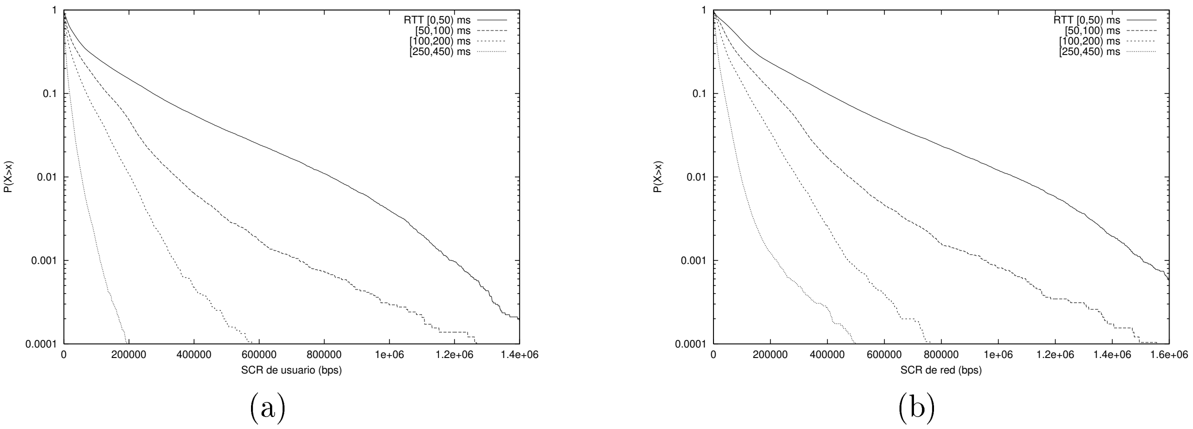


Figura 3.6: SCRs según el RTT de las conexiones. a)SCRs de usuario b) SCRs de red.

En las gráficas de la figura 3.7 podemos ver la diferencia entre la medición de throughput a nivel de usuario y a nivel de red, así como el efecto que tiene el RTT sobre las mismas. Resulta visible la diferencia entre el throughput de usuario y el de red. Además, para RTTs grandes aumenta el efecto del tiempo de establecimiento, lo cual lleva a que el SCR de usuario se reduzca mucho más que el de red. En la figura 3.8 podemos ver, equivalentemente a la figura 3.2, el porcentaje de tiempo que supone la fase de establecimiento, pero en este caso en función del RTT de la conexión. Se espera que con distribuciones de tamaños de ficheros similares el porcentaje de la duración que representa el establecimiento sea independiente del RTT de la misma. Sin embargo, como hemos dicho las conexiones no siguen la progresión esperada, bien por pérdidas, retardos variables, etc, lo cual se observa en este caso en que el establecimiento tiende a ser una parte cada vez más importante de la duración de la conexión a medida que el RTT de las mismas aumenta. Cuando el RTT alcanza el rango de valores de los enlaces internacionales de Internet ([250,450) ms) el gran número de efectos debidos a los diferentes caminos y situaciones de cuello de botella a los que se enfrentan las conexiones hacen que aún cambie más el comportamiento de este parámetro.

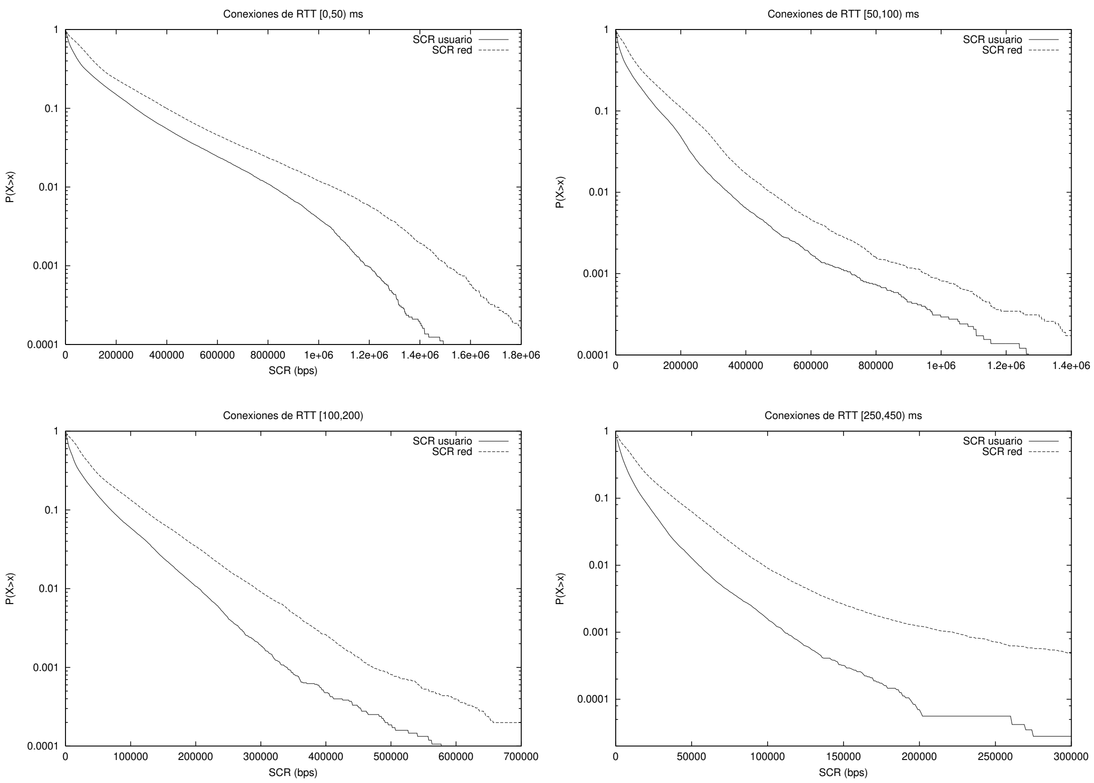


Figura 3.7: SCRs según el RTT de las conexiones. SCRs para RTT [0,50)ms d) [50,100) ms e) [100,200) ms f) [250,450) ms

Si ajustamos una variable aleatoria de Weibull para estas distribuciones de probabilidad nos encontramos con los parámetros que aparecen en la tabla 3.1. El exponente de la variable independiente (β) determina la curvatura de la función de supervivencia y vale 1 para una variable aleatoria exponencial. Vemos que en función del RTT el exponente se mantiene dentro de un margen de variación de ± 0.02 para sl SCR de usuario y de ± 0.06 para el de red.

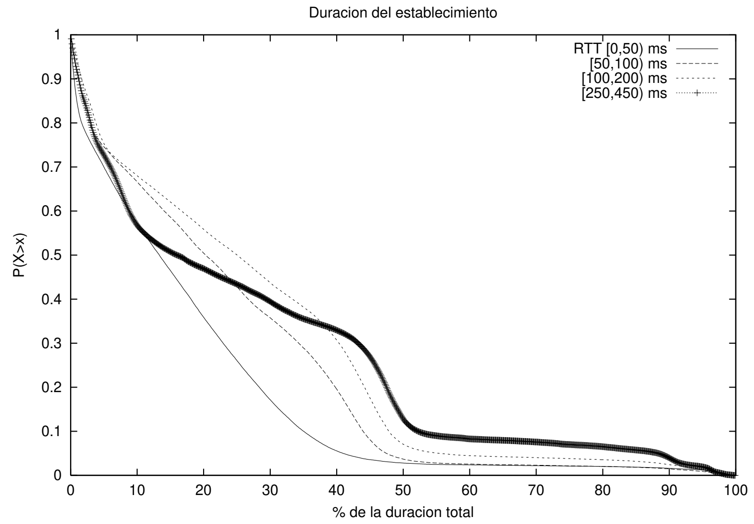


Figura 3.8: Tiempo de establecimiento según el RTT.

La forma de la Weibull se mantiene incluso en el caso del RTT para accesos internacionales, especialmente desde el punto de vista del usuario. El cambio principal se da en el parámetro α de la distribución, el cual determina la *rapidez* con que cae la función de supervivencia. Así, en situaciones de elevado RTT la probabilidad de obtener un cierto SCR (tanto a nivel de usuario como de red) disminuye muy rápidamente con el SCR deseado, si bien la curvatura característica de la función de supervivencia se mantiene aproximadamente igual. Si las variables se pudieran aproximar por exponenciales puras sería equivalente a decir que cambia la tasa λ pero sin dejar de ser exponencial.

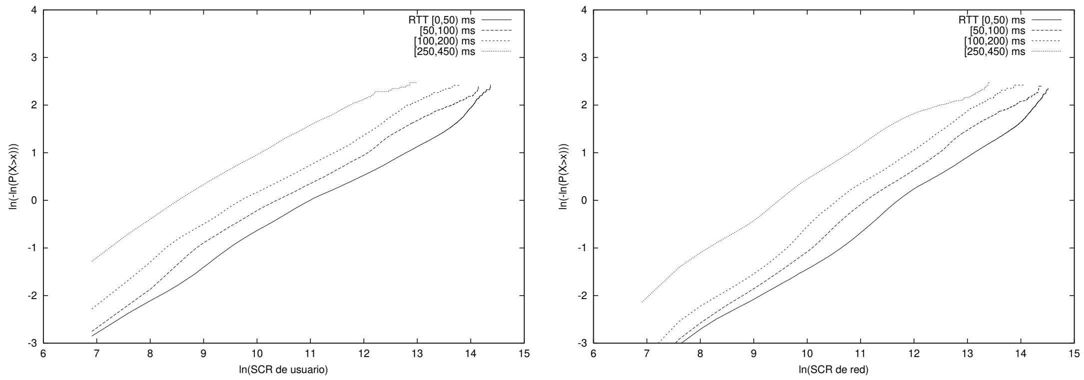


Figura 3.9: Representación loglog-log del SCR

Finalmente veremos cómo afecta la hora del día al valor medio de estas tasas de transferencia.

	SCR de usuario		SCR de red	
RTT (ms)	α	β	α	β
$[0, 50)$	0.651e-3	0.660	0.173e-3	0.734
$[50, 100)$	0.752e-3	0.680	0.126e-3	0.798
$[100, 200)$	1.215e-3	0.678	0.182e-3	0.801
$[250, 450)$	4.000e-3	0.643	1.396e-3	0.689

Tabla 3.1: Parámetros estimados para una distribución de Weibull

En la figura 3.10 tenemos el número de conexiones en cada intervalo de una hora. Vemos que las horas nocturnas tienen un escaso número de muestras, insuficientes para calcular un valor medio con un cierto nivel de confianza. Así pues nos centraremos en las horas para las que tenemos suficientes muestras.

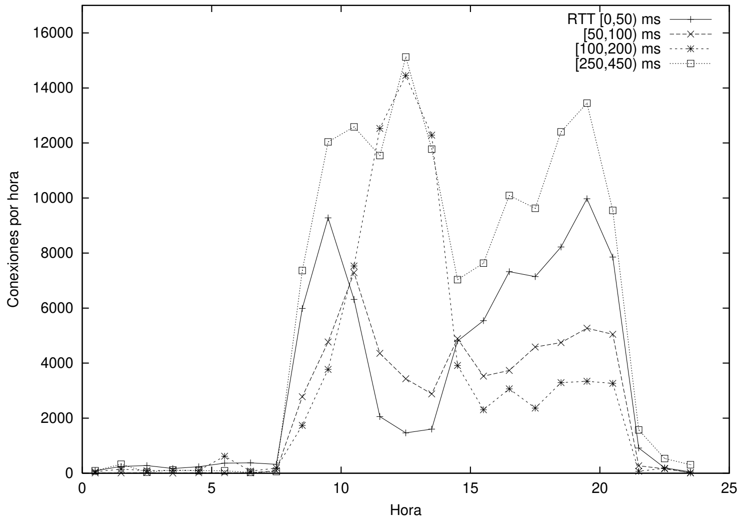


Figura 3.10: Conexiones por hora (19Oct1999)

En la figura 3.11 se presentan los valores medios por hora de la tasa de transferencia tanto a nivel de usuario como a nivel de red según el RTT de las conexiones. Vemos que los valores de ambas y por tanto su diferencia se mantiene razonablemente constante a lo largo del día (figura 3.12). Las zonas sombreadas corresponden a las horas nocturnas en las que como se ve en la figura 3.10 existe un número insuficiente de conexiones. Así pues, el parámetro principal para determinar el SCR que alcancen las conexiones va a ser el retardo entre cliente y servidor.

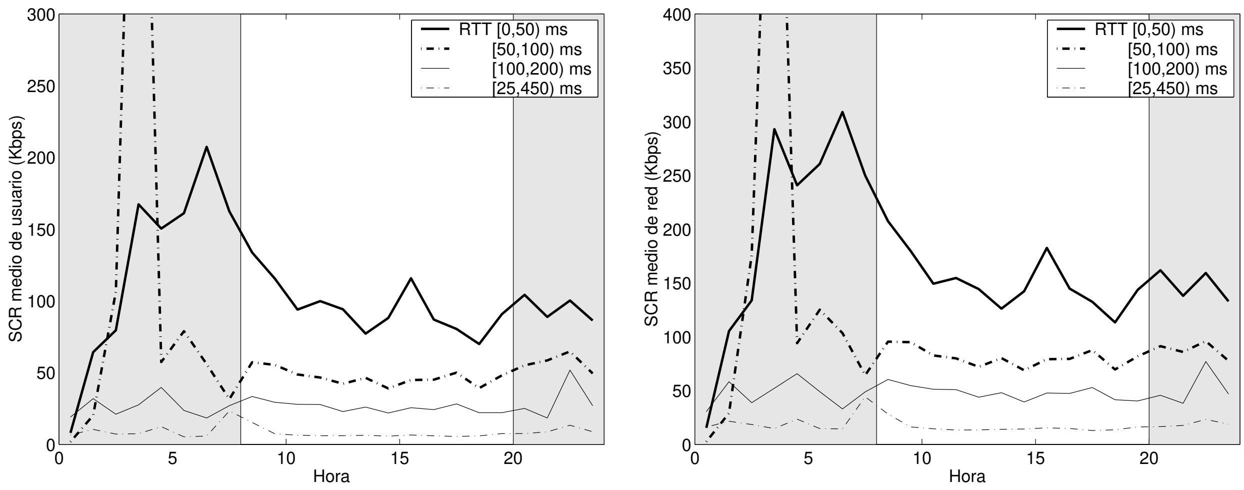


Figura 3.11: SCRs medios por hora (19Oct1999)

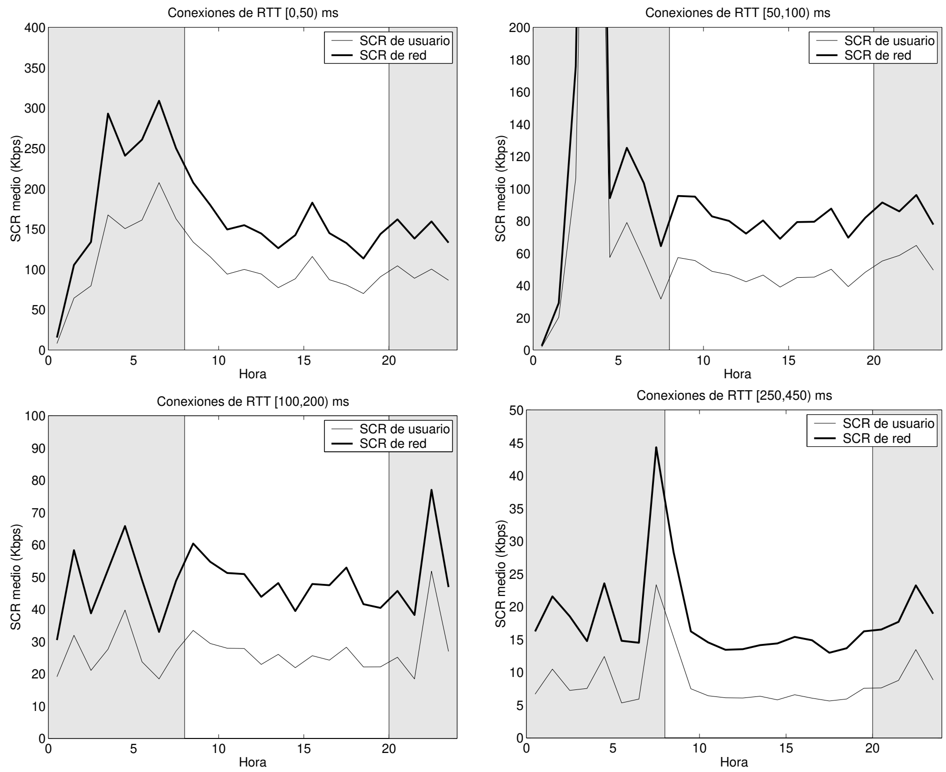


Figura 3.12: SCRs por hora y RTT (19Oct1999)

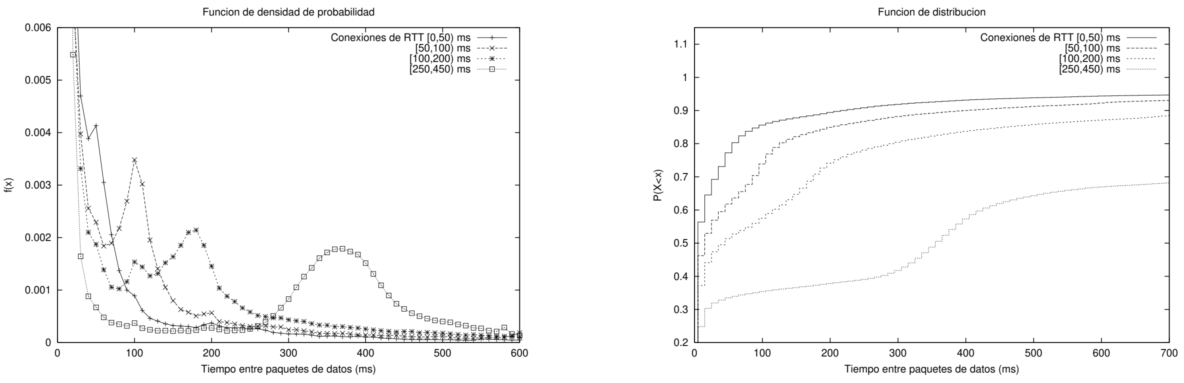
3.5 Intermitencia

Determinar el Tamaño Máximo de Ráfaga requiere un umbral para el tiempo entre paquetes para decidir si una serie de paquetes consecutivos pertenece o no a la misma ráfaga. Este umbral viene determinado por la tasa máxima del enlace o PCR. Las ráfagas debidas a acumulación de paquetes en el buffer se transmitirán a esta tasa de pico por lo que podremos emplear esta condición para determinar si una serie de paquetes pertenece a la misma ráfaga.

Comprobamos en la traza de tráfico que *no se producen ráfagas de paquetes pertenecientes a la misma conexión a la tasa máxima del enlace en todo el periodo de medida.*

Esto se debe a que las ráfagas que se producen dentro de las conexiones TCP dependen del comportamiento de toda la ruta entre servidor y cliente y de la dinámica de funcionamiento de TCP. Por un lado la red introduce un jitter o desviación en los tiempos entre paquetes que puede separar los paquetes. Por otro lado, dado que las conexiones son cortas y pasan la mayor parte de su vida en la fase de slow start, las ráfagas transmitidas desde el origen serán cortas, por lo que pequeñas desviaciones romperán las posibles ráfagas en origen.

Para estudiar con mayor detalle la formación de ráfagas dentro de las conexiones TCP analizamos los procesos de tiempo entre paquetes de servidor a cliente de la misma conexión. En la figura 3.13 se ha representado la función de densidad de este tiempo entre paquetes para conexiones separadas en diferentes conjuntos según el RTT estimado de las mismas. Existe una gran dispersión de los valores posibles de tiempo entre paquetes pero en todas las gráficas destacan dos zonas principales. La primera se encuentra localizada en valores pequeños de tiempo entre llegadas y corresponde a los paquetes que formarían ráfagas al ser transmitidos desde el servidor y que se han separado ligeramente debido el jitter introducido por la red. La segunda zona destacable se reconoce por ser un lóbulo con forma de campana gaussiana en torno al RTT de esas conexiones. Este segundo grupo de valores en torno al RTT corresponde a los tiempos entre los paquetes cuando el servidor tuvo que detener su transmisión a la espera de confirmaciones (ACKs) que le permitieran transmitir más paquetes, es decir, son la prueba del funcionamiento mediante técnica de *parar y esperar* de TCP al colocar en la red el máximo permitido por las ventanas de control de flujo y de congestión.



(a) (b)

Figura 3.13: Tiempo entre paquetes de la misma conexión de servidor a cliente a) fdp b) FDP (19Oct1999)

Sin embargo, con la figura 3.13 no podemos saber cuál es el tamaño típico de las ráfagas de paquetes o su distribución. Para ello necesitamos información sobre la estructura de correlación de los tiempos entre paquetes. Aunque, por ejemplo, el 30% de los tiempos entre paquetes sean pequeños, la formación de ráfagas más o menos largas dependerá de que se dé un gran número de ellos consecutivos o no.

La figura 3.13 nos sugiere un umbral para el tiempo máximo entre paquetes pertenecientes a la misma ráfaga por debajo de 50ms. Tomaremos un valor de 20 milisegundos, siendo correcto principalmente para las conexiones cuyo RTT sea superior a este valor, para aquellas con RTT cercano a los 20 ms se podrá confundir paquetes separados por el RTT con paquetes en la misma ráfaga. El tamaño de ráfaga obtenido con este umbral será siempre mayor que el que encontraríamos a nivel ATM empleando el PCR, con lo que nos da una cota superior que puede resultar útil con propósitos de dimensionamiento. Hay que tener en cuenta que la separación de dos paquetes de la MTU de Ethernet (1696 a nivel ATM) transmitidos al PCR del enlace en estudio (4Mbps) es de unos 3.4 milisegundos, un orden de magnitud inferior al valor que estamos tomando.

Dado que estamos monitorizando un solo circuito virtual no encontraremos multiplexación estadística con lo que las celdas de un paquete se transmitirán una tras otra sin ninguna otra intercalada. Además, dado que los extremos del enlace son routers IP, éstos tendrán que realizar un reensamblado de la PDU de AAL5 para comprobar la integridad de la misma y obtener la cabecera IP para poder reenviar el paquete. Así pues, cuando el router intenta enviar el paquete posee todas las celdas por lo que éstas saldrán al PCR del enlace y sin entremezclarse con celdas de otra PDU. Debido a esto podemos pensar en los paquetes como unidades de información transmitidas al PCR del enlace, con lo que podemos relacionar fácilmente el tamaño de ráfaga en número de paquetes (o mejor en bytes) con el equivalente en celdas.

En las figuras 3.14 y 3.15 presentamos la probabilidad de los tamaños de ráfaga más comunes en intervalos de una hora y según el RTT de las conexiones. Vemos que la intermitencia de las conexiones es muy pequeña. El 80% de las ráfagas son de solo dos paquetes. Esto sugiere el empleo tamaños pequeños de buffer por conexión, los cuales bastarían para evitar pérdidas. La mayor parte de las conexiones no alcanzan un estado estacionario y producen ráfagas cortas de paquetes, las cuales se enfrentan al jitter de la red que, aumentando la separación entre paquetes, suele romperlas en ráfagas más pequeñas.

3.6 Validación del modelo

Los resultados presentados hasta el momento se basan en una medición de un día de duración (19/10/1999). Gracias a la disponibilidad de varios registros de tráfico (tabla 2.1) vamos a comprobar que los resultados son análogos y reproducibles independientemente del intervalo de tiempo seleccionado. Para ello tomamos la semana del 14/12/1998 al 20/12/1998, englobada en la traza NAV98, y presentamos los resultados principales.

En esta medida (una semana) el tráfico TCP representa el 98.74% y de él el 91.71% de los flujos (72.91% de los bytes) se deben al servicio Web. Así pues vemos que de nuevo podemos modelar el comportamiento del tráfico en base al servicio más empleado. Para este servicio estudiaremos el número de conexiones simultáneas, la tasa de transferencia media y la intermitencia de las mismas.

En la figura 3.16 vemos el número de conexiones Web simultáneas presentes en el inter-

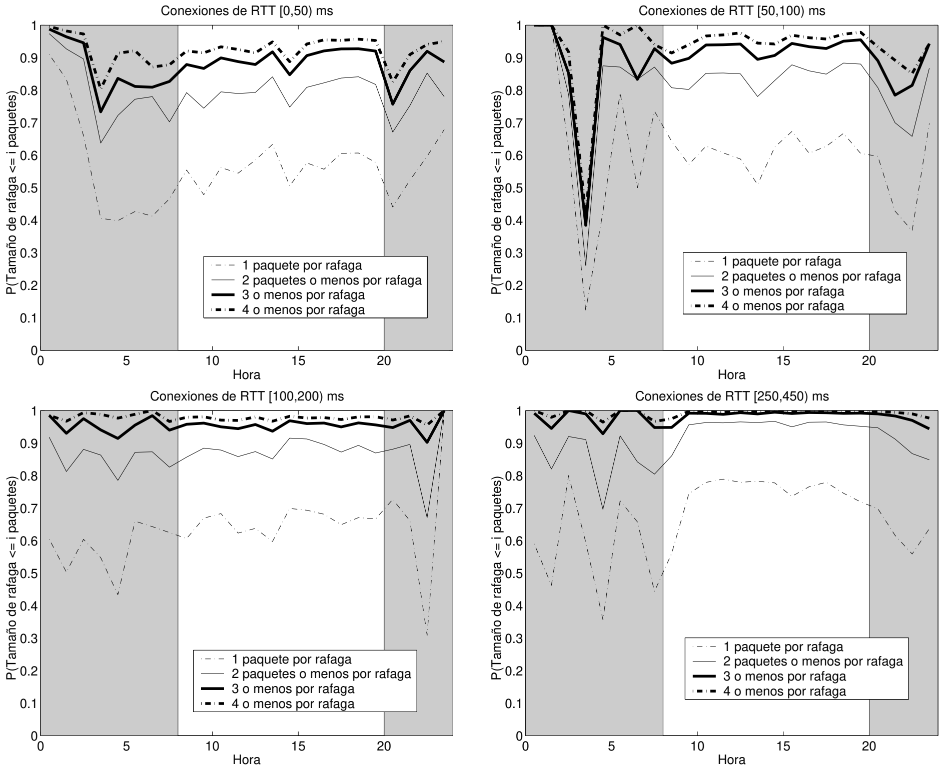


Figura 3.14: Tamaño de ráfagas en paquetes

valo mencionado. Se observa la misma periodicidad debida a la menor actividad durante el mediodía y la noche, así como el diferente patrón, generalmente con cotas más reducidas, correspondiente al fin de semana. En estos fines de semana el estado más descargado de la red podría producir que aún con un número de usuarios real más reducido se puedan generar cifras de conexiones comparables. El principal parámetro destacable al cambiar el periodo de medida es el menor número de conexiones simultáneas en Diciembre de 1998. En el transcurso de 10 meses desde Diciembre de 1998 hasta Octubre de 1999 se produjo un aumento en un factor de cinco a uno.

Para el estudio de la tasa de transferencia y la intermitencia de las conexiones procederemos a agruparlas según el RTT de las mismas. A la vista de los RTTs presentes en la medida (figura 3.17) definimos dos grupos según su RTT sea “Bajo” ([5,100) ms) o “Alto” ([450,650) ms). Comprobamos mediante la figura 3.18 que tanto el SCR de usuario como el de red en función de la hora del día mantienen el mismo patrón expuesto para la traza de Octubre.

El último parámetro del modelo simple propuesto se basa en la probabilidad de llegada de ráfagas de paquetes de una misma conexión. De nuevo separamos los flujos en dos intervalos de RTT y con un umbral de tiempo entre paquetes de 20ms presentamos dicha probabilidad en la figura 3.19. Vemos que de nuevo las ráfagas son en general de pequeño tamaño, a penas de 1 ó 2 paquetes.

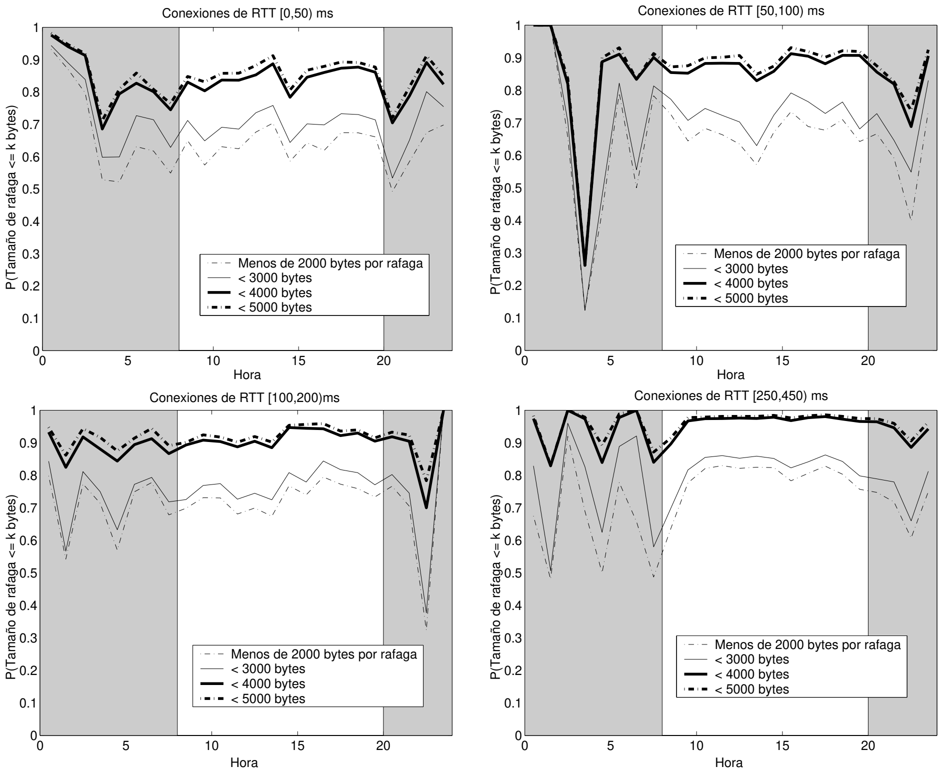


Figura 3.15: Tamaño de ráfagas en bytes

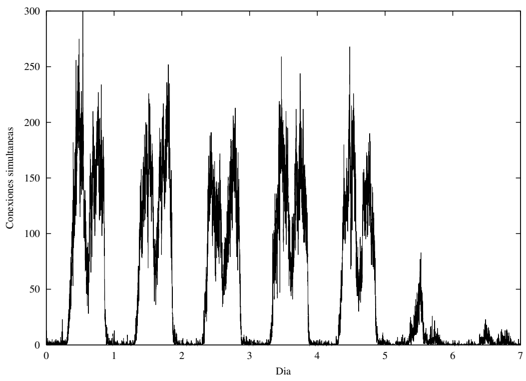


Figura 3.16: Conexiones simultáneas para una semana (14Dic1998-20Dic1998)

3.7 Conclusiones

Como hemos visto, al contrario de los supuesto en la mayoría de modelos de tráfico las conexiones TCP no siguen un patrón exponencial de crecimiento. Debido al reducido tamaño de las transferencias y a la separación entre paquetes introducida por la red los paquetes

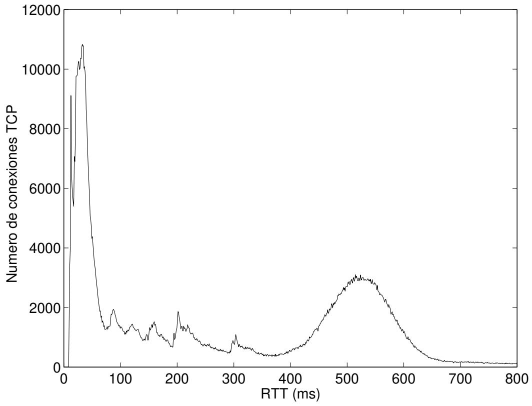


Figura 3.17: RTTs de conexiones (14Dic1998-20Dic1998)

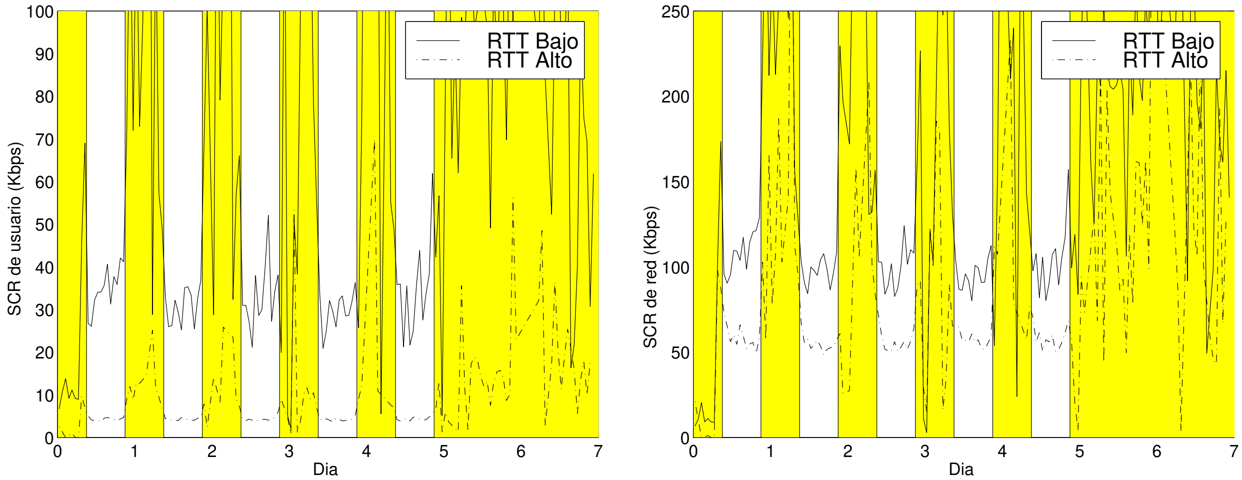


Figura 3.18: SCRs para una semana (14Dic1998-20Dic1998)

llegan muy separados, formando ráfagas de 2 ó 3 paquetes aún con umbrales de medición muy pesimistas. Por otro lado la tasa media de transferencia está fuertemente condicionada por el retardo extremo a extremode la conexión. Así pues podemos representar el flujo de datos en estas conexiones mediante un proceso de tasa media constante según el RTT y con una intermitencia en torno al tamaño de ráfaga presente en el tráfico real (1 ó 2 paquetes).

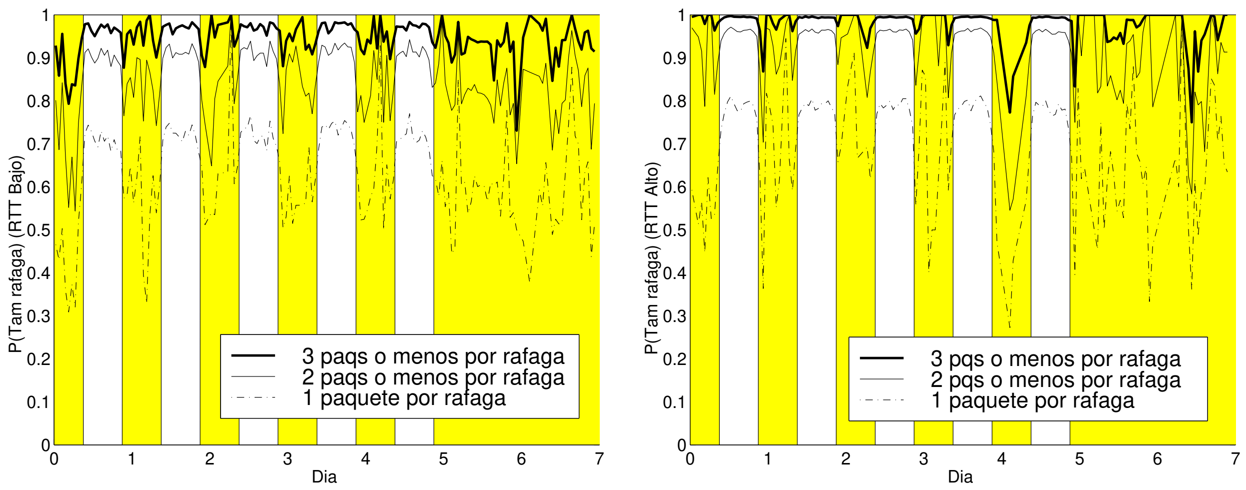


Figura 3.19: Ráfagas para una semana (14Dic1998-20Dic1998)

Capítulo 4

Modelado de tráfico de Internet como multiplex de conexiones TCP no regulares

4.1 Introdución

Numerosos estudios recientes muestran que la carga de Internet se debe principalmente a las conexiones TCP del servicio WWW. En el capítulo 2 hemos visto que en el enlace de la UPNA este servicio representa el 65.02% del tráfico del sentido entrante, el 76.6% si solo consideramos las conexiones generadas por los usuarios, y el 92.9% de las conexiones de los mismos. Encontramos resultados similares en otros trabajos realizados sobre medidas tomadas de una gran variedad de entornos, tanto académicos como industriales [13, 64, 65].

Dado que las conexiones TCP determinan la mayor parte de la carga de los enlaces de Internet hay un creciente interés por el desarrollo de modelos de tráfico que permitan explicar las características principales del tráfico de Internet en función de estas conexiones TCP multiplexadas.

Sabemos que el resultado de multiplexar fuentes ON/OFF con intervalos ON u OFF según una variable aleatoria con cola pesada es un proceso autosimilar [27, 19]. De hecho, aun con un proceso de llegadas de conexiones de Poisson, con tal de que los intervalos ON sigan una distribución de ese tipo el resultado es asintóticamente autosimilar [22]. Estos procesos autosimilares resultan muy interesantes dado que capturan de forma simple características importantes del tráfico de Internet como la presencia de dependencia a largo plazo.

Estos modelos asumen que las diferentes fuentes transmiten a igual tasa, es decir, la intensidad del tráfico en un intervalo ON es la misma para todos los posibles flujos, lo cual significa que se considera constante la tasa de transferencia a lo largo de toda la duración del flujo y que todos los flujos tienen esa misma tasa. Esta tasa constante no tiene un significado físico en la construcción del modelo. Sin embargo, se ha empleado este tipo de modelos haciendo una equivalencia entre los intervalos ON y las conexiones TCP. En [64] se adopta este modelo para explicar la presencia de autosimilaridad en el tráfico de WWW, basándose en que tanto los tamaños de los ficheros transferidos como las duraciones se pueden modelar con variables aleatorias de cola pesada. Como simplificación del modelo se asume que todas las conexiones TCP presentan la misma tasa constante. Estas mismas hipótesis se han empleado repetidamente en trabajos de simulación y dimensionado de red [27, 66], con lo que es un modelo comúnmente aceptado para explicar la presencia de dependencia a largo plazo en el tráfico de Internet.

Sin embargo, como hemos comentado en el capítulo anterior, tanto la hipótesis de tasa constante durante toda la duración de la conexión como entre diferentes conexiones distan de cumplirse. Dentro de la duración de una conexión la etapa de slow start hará que la tasa máxima alcanzable varíe al evolucionar ésta y el efecto de los retardos en los enlaces de la red aún introducirá una mayor variabilidad. Por otro lado, se ha comprobado que la tasa media de las conexiones TCP no es constante sino que varía según la situación específica del enlace entre cliente y servidor. El protocolo TCP ajusta la tasa de transferencia a los parámetros de este enlace, tales como tasa del cuello de botella o retardo extremo a extremo, resultando en diferentes tasas de transferencia media por conexión.

En este capítulo abordaremos el modelado del tráfico de Internet como un multiplexado de conexiones TCP. Llamaremos *modelo de tasa constante* o *modelo regular* a aquellos que se basan en la misma tasa de transferencia para todos los flujos. El modelo que propondremos tendrá en cuenta las diferentes tasas de transferencia de las conexiones, lo cual nos llevará a un tráfico resultante con mayor variabilidad que con el modelo de tasa constante. Así pues, el modelo de tasa constante llevará a estimaciones optimistas. Esto es debido a que las prestaciones de los sistemas con cola no son debidas tan solo a la dependencia a largo plazo, la cual capturan ambos modelos, sino que dependen también de la variabilidad o coeficiente de variación de la distribución marginal del proceso. Mientras que la dependencia a largo plazo caracteriza la evolución temporal del proceso de tráfico la distribución de la marginal viene determinada exclusivamente por el comportamiento instantáneo del proceso.

Finalmente veremos cómo el empleo de nuevos modos de transferencia diferentes de TCP y orientados a las nuevas redes de alta velocidad con conmutación IP y técnicas de conmutación óptica de ráfagas (OBS Optical Burst Switching) puede llevar a un aumento en la variabilidad de la distribución marginal a la vez que desaparece la dependecia a largo plazo al hacerse las conexiones cada vez más cortas. En este tipo de escenarios el factor determinante del funcionamiento de la red será la distribución marginal del proceso en vez de la dependencia a largo plazo.

Para este estudio hemos escogido una medición más actual, en este caso del Lunes día 14 de Febrero del 2000, englobado dentro de la traza FEB00 (tabla 2.1). Dado que los modelos estudiados se basan en procesos estacionarios deberemos restringirnos a intervalos de tiempo en los que esta hipótesis sea razonable, así pues, una traza más larga no aportará más información. Durante este día se recogieron más de 1 millón de conexiones TCP, de las cuales el 96.9% se debieron al servicio Web. Estas conexiones de Web supusieron el 82.8% del tráfico TCP, por lo que de nuevo centraremos el estudio sobre los flujos debidos a este servicio.

A continuación, en el apartado 4.2 modelaremos el proceso de llegadas de conexiones así como las duraciones, tamaños y tasas de transferencia de las mismas. En el apartado 4.3 describiremos los parámetros de los modelos estudiados y en el 4.4 presentaremos una comparativa de las prestaciones de cada modelo. Finalmente, como veremos en el apartado

4.5, las conclusiones extraídas del estudio de estos modelos aportarán un nuevo punto de vista para la evolución del tráfico de datos en las futuras redes ópticas de alta velocidad.

4.2 Caracterización de conexiones TCP

Para construir un modelo de tráfico en base al multiplexado de conexiones TCP necesitamos caracterizar el proceso de llegada de conexiones, los tamaños de las mismas tanto en bytes transferidos como en duración y la tasa de transferencia (relación entre bytes y duración). Emplearemos el tiempo entre comienzos de conexiones como caracterización del proceso de llegadas. Posteriormente analizaremos la duración y tamaño de las conexiones, comparándolas con los resultados de otros estudios [13, 64]. Finalmente introduciremos la característica de no regularidad de las conexiones mediante un análisis de la tasa de transferencia de las mismas.

4.2.1 Proceso de llegadas

Al igual que la mayoría de modelos, el de tasa constante presentado en [27, 22] asume que el proceso de llegadas es estacionario. Los modelos estacionarios que sean capaces de capturar las características del tráfico serán mucho más sencillos de tratar analíticamente que los modelos no estacionarios, por lo que son la primera solución a buscar. En este capítulo nos centraremos en modelos que mantengan esta hipótesis de estacionariedad. Para ello deberemos escoger intervalos de tiempo en los que esta hipótesis sea razonable.

En la figura 4.1 vemos la tasa media de llegadas en intervalos de una hora de duración. Claramente el proceso de llegadas de conexiones no es un proceso homogéneo. Existe un gran número de usuarios activos en las horas de mayor utilización, produciendo un mayor número de conexiones TCP, y podemos observar bajones bruscos en el periodo del mediodía o de la noche.

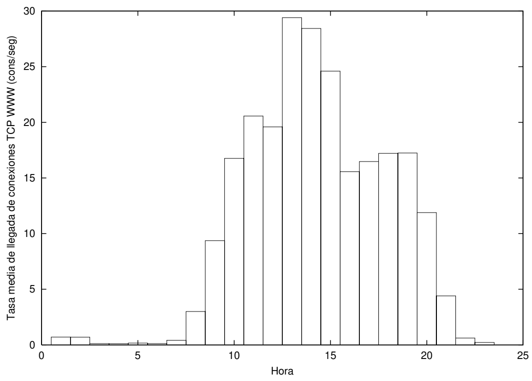


Figura 4.1: Tasa media de llegadas de conexiones Web

Para mantener la condición de estacionariedad nos restringiremos a intervalos de una o dos horas de duración. Para estos intervalos calculamos la función de supervivencia $P(X > x)$

del tiempo entre llegadas de conexiones (figura 4.2). Incluso por simple inspección visual se comprueba que la marginal del tiempo entre llegadas puede modelarse con una variable aleatoria exponencial, la cual daría una recta en esta representación semilogarítmica. Por otro lado, tal y como se ha hecho en otros trabajos [66], podemos suponer independientes a unos usuarios de otros y por lo tanto independientes las conexiones generadas por unos de las de otros. Un proceso de llegadas con tiempos entre llegadas exponenciales, independientes e idénticamente distribuidas es un proceso de Poisson.

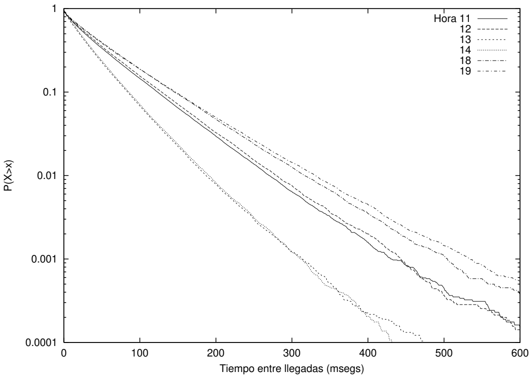


Figura 4.2: Función de supervivencia del tiempo entre llegadas

4.2.2 Duración y tamaño de las conexiones

La figura 4.3 muestra las funciones de supervivencia de los tamaños de las transferencias (bytes) y las duraciones de las conexiones (segundos) en escala logarítmica. Sabemos que la cola de la función de supervivencia de una variable aleatoria de cola pesada (1.10) da una línea recta de pendiente $-\alpha$ cuando se representa en escalas logarítmicas. Así pues representamos también la recta de regresión por mínimos cuadrados para la cola en ambas figuras y estimamos unos valores del α para la misma de 1.15 y 1.2 (tomando puntos de la cola equiespaciados en escala logarítimica) para la duración y el tamaño respectivamente. Estos valores están de acuerdo con los resultados de otros estudios [64].

Encontramos que la distribución de los tamaños de las transferencias (en general atribuibles a los tamaños de los ficheros) en Internet posee una cola pesada, debida a la naturaleza tan diversa de la información disponible en la red, desde pequeños ficheros de texto hasta videos de larga duración y alta calidad [64].

Respecto a la duración de las conexiones vemos que la variabilidad es incluso mayor que en el caso de los tamaños (tiene un α menor). Posiblemente esto se deba a la dinámica de TCP en presencia de congesión, la cual hace que las duraciones de las conexiones crezcan si se dan pérdidas.

El modelo regular tiene en cuenta los parámetros estudiados hasta el momento. Dado que tanto el valor de α para la cola de la distribución de los tamaños como de las duraciones está en el rango $1 < \alpha < 2$ el parámetro de Hurst del tráfico resultante se encontrará en el rango

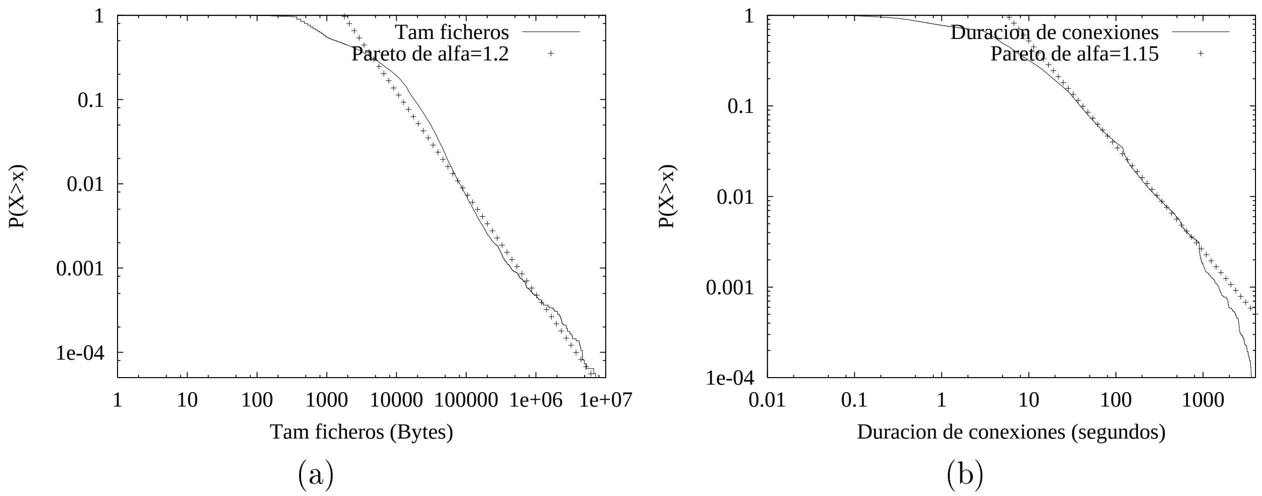


Figura 4.3: Función de supervivencia de los tamaños (a) y las duraciones (b) de las conexiones

$0.5 < H < 1$, presentando dependencia a largo plazo. Esta LRD se debe a las grandes ráfagas de tráfico indicadas por las funciones de supervivencia de la figura 4.3. Sin embargo, este modelo no captura la presencia de diferentes tasas de transferencia en las conexiones TCP.

4.2.3 Tasa de transferencia de las conexiones

La figura 4.4 muestra la función de supervivencia de la tasa media de transferencia de las conexiones TCP. Calculamos esta tasa como el cociente entre los bytes transmitidos a nivel ATM entre la duración de la conexión. Observamos que la tasa de transferencia no es constante, sino que al igual que en el capítulo anterior, se puede modelar con una variable aleatoria de Weibull.

Una característica destacable de la distribución de Weibull es que no tiene una cola tan pesada como para que su segundo momento no exista (aunque en este caso sí es más pesada que la de una variable aleatoria exponencial). Las variables aleatorias de duración de la conexión y tamaño de la transferencia tienen una cola de Pareto. En caso de ser independientes, el cociente entre ambas, que es la tasa de transferencia, debería tener también una cola de Pareto. Sin embargo estas dos variables se encuentran correladas, como era de esperar, produciendo una distribución con una cola menos pesada (de Weibull).

Puede comprobarse esta correlación con la figura 4.5, que muestra el valor medio de la tasa de transferencia en función del tamaño en bytes de la transferencia y del RTT estimado de la conexión. Podemos ver que estas variables (tasa y tamaño) no son independientes y por lo tanto tampoco lo son la duración y el tamaño. Al representar la tasa media frente al tamaño del fichero vemos el efecto de la etapa de slow-start en la tasa de transferencia, para ficheros pequeños se alcanzan menores tasas, lo cual en la variable duración significa flujos más largos.

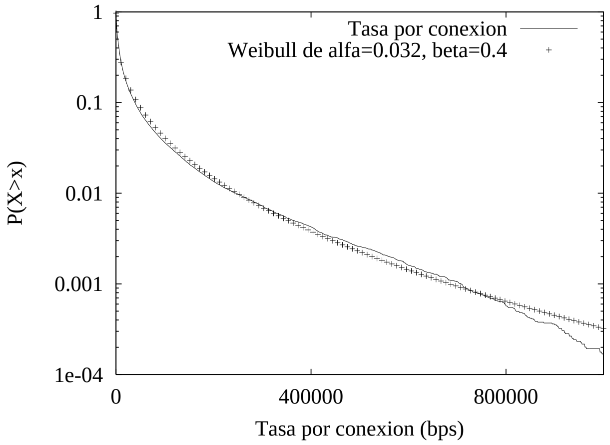


Figura 4.4: Función de supervivencia de la tasa

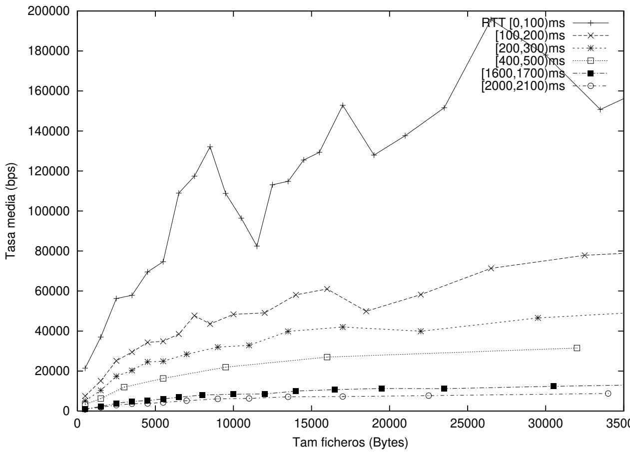


Figura 4.5: Tasa media por conexion frente al tamaño de los ficheros para varios RTTs

4.3 Características del múltiplex de conexiones TCP

En esta sección compararemos el modelo de tráfico *regular* y el *no regular*, definidos de la siguiente forma:

- *Modelo regular*: El tráfico de Internet es el resultado de multiplexar llegadas según un proceso de Poisson con duración según una variable aleatoria de cola pesada y todas con la misma tasa de transferencia (modelo $M/G/\infty$). Sintetizaremos un proceso de tráfico con los mismos instantes de llegada que en la traza real (Poisson) y las mismas duraciones de las conexiones (de cola pesada), asumiendo que los paquetes se transmiten a tasa constante e igual para todos los flujos.

- *Modelo no regular*: El tráfico de Internet es el resultado de multiplexar llegadas según un proceso de Poisson y duración con cola pesada pero con diferentes tasas de transferencia. Sintetizaremos un proceso de tráfico con los mismos instantes de llegada, el mismo número de bytes y la misma tasa de transferencia media por conexión que en la traza real.

Para relizar una comparación justa seleccionamos el valor de la tasa de transferencia para el caso regular de forma que resulte en la misma carga (bytes) que en el caso no regular.

La figura 4.6 muestra el número de MBytes por minuto en un intervalo de dos horas de duración para la traza real de tráfico, el modelo no regular y el regular correspondientes a esas llegadas. Se observa claramente que el modelo regular presenta una intermitencia inferior a la ofrecida por el caso no regular. Por otro lado, el modelo no regular se ajusta mediante simple inspección visual mucho mejor a la traza real .

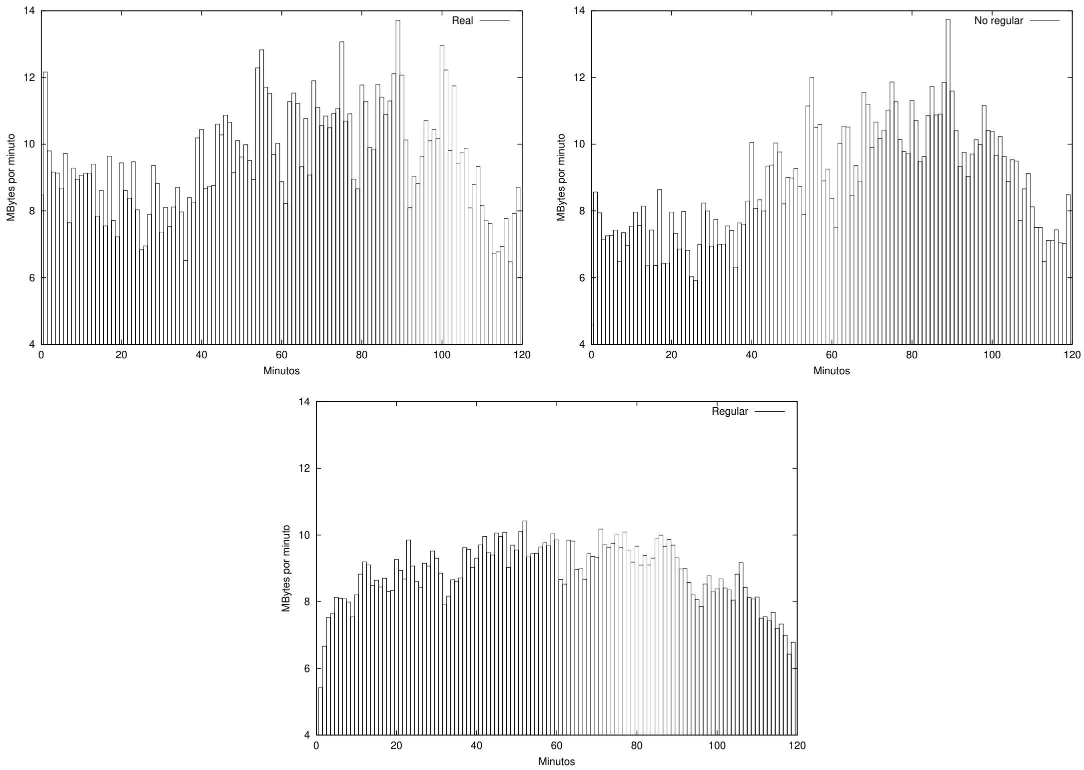


Figura 4.6: Modelos de tráfico real, no regular y regular

Aunque la figura 4.6 muestra evidencias visuales del mejor ajuste a las características del tráfico real por parte del modelo no regular llevamos a cabo simulaciones con estos procesos (a nivel de paquete) siendo la entrada de un sistema compuesto por un servidor con cola infinita, para así poder evaluar los efectos sobre un elemento de red. Los resultados se pueden ver en la figura 4.7, que presenta el retardo medio de paquete frente al factor de utilización del servidor. Se observa una pérdida de prestaciones con el modelo no regular, similar a la obtenida con el tráfico real. Sin embargo el modelo regular da unos resultados poco realistas y muy optimistas. Como veremos a continuación esto se debe al peor modelado que hace de la distribución marginal del proceso de tráfico.

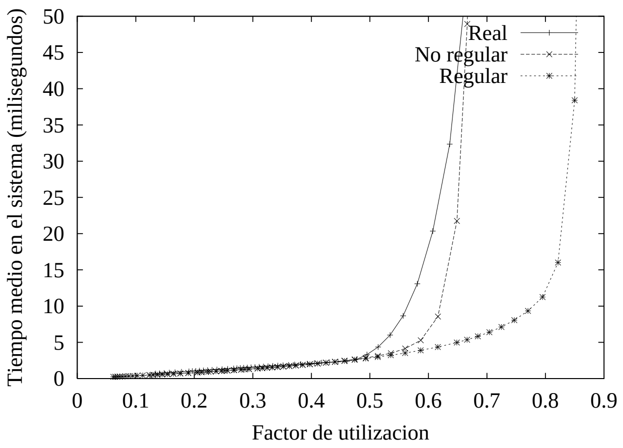


Figura 4.7: Evaluación de prestaciones de un sistema de un servidor frente al tráfico real, al de tasa variable y al de tasa constante

4.4 Evaluacion de prestaciones de ambos modelos

En [32] se pueden encontrar resultados analíticos para las prestaciones de un servidor único con cola infinita ante un proceso de entrada autosimilar (FGN[1]). La función de supervivencia del retardo por paquete en el sistema viene dada por:

$$P(Q > x) \sim exp\left(-\frac{(C-m)^{2H}}{2k(H)^2 c_v^2 m^2} x^{2-2H}\right) \tag{4.1}$$

donde C es la capacidad del enlace, m el tasa media del tráfico de entrada, $c_v = \sigma/m$ la desviación estándar dividida por la media del tráfico de entrada (coeficiente de variación de la marginal del tráfico de entrada) y $k(H) = H^H(1-H)^{(1-H)}$, donde H es el parámetro de Hurst. Esta ecuación muestra que las prestaciones de una cola dependen de las siguientes dos variables (sin contar el factor de utilización): en primer lugar, de la dependencia a largo plazo del tráfico de entrada (parámetro H) y en segundo lugar de la variabilidad de la distribución marginal (parámetro c_v). El modelo regular captura correctamente la característica de dependencia a largo plazo del tráfico (por converger a un proceso autosimilar [19]), sin embargo, no es capaz de hacer lo mismo con el coeficiente de variación de la distribución marginal. Para comprobar esta hipótesis estimamos los valores del parámetro H y del coeficiente de variación c_v para el modelo regular, el no regular y la traza real. Realizamos las estimaciones para intervalos de una hora de duración donde podremos considerar el tráfico estacionario (salvo en las horas de

[1]Fractional Gaussian Noise

comienzo de la actividad, 8AM, o de finalización, 9PM, donde se producen fuertes subidas o desvanecimientos de la actividad en el enlace). El coeficiente de variación se obtiene de forma inmediata a partir de la desviación y la media de la marginal. Para estimar el parámetro de Hurst llevamos a cabo regresiones lineales en representaciones logarítmicas de la varianza frente al nivel de agregación. Los resultados se muestran en las tablas 4.1 y 4.2.

Tabla 4.1: Estimación de H frente al tiempo

Hora	Real	No reg	Regular
8	0.97	0.98	0.98
9	0.87	0.78	0.95
10	0.78	0.81	0.99
11	0.85	0.90	0.95
12	0.85	0.88	0.97
13	0.81	0.87	0.97
14	0.91	0.95	0.99
15	0.92	0.95	0.97
16	0.88	0.80	0.96
17	0.80	0.87	0.99
18	0.86	0.87	0.91
19	0.88	0.94	0.95
20	0.83	0.86	0.99
21	0.94	0.97	0.99

Vemos que el parámetro de Hurst toma valores parecidos en todos los casos, mientras que el modelo regular presenta un coeficiente de variación inferior que el modelo no regular (en algunos casos hasta diez veces menor que el del tráfico real). Así pues, el modelo regular presenta una aproximación optimista de la variabilidad de la distribución marginal. Para mayor claridad, la figura 4.8 muestra la función de supervivencia de la distribución marginal (en escala semilogarítmica) para el tráfico real y los modelos regular y no regular para un intervalo de dos horas de duración (17:30 a 19:30). Vemos que el modelo no regular sigue una distribución con una cola más pesada que en el caso regular, dando una mejor aproximación del tráfico real.

4.5 Evolución del tráfico de Internet

Los resultados de la sección anterior muestran que las prestaciones de sistemas con cola ante tráfico de Internet no están determinadas solo por la dependencia a largo plazo de este tráfico sino también por la variabilidad de su distribución marginal.

Si consideramos el caso de un enlace de Internet que concentra a una gran población de usuarios podemos considerar al proceso de llegada de conexiones como un proceso de Poisson, debido a la demanda inherentemente aleatoria que presentan tales poblaciones. Por otro lado,

Tabla 4.2: Estimación de c_v frente al tiempo

Hora	Real	No reg	Regular
8	1.13	1.00	0.59
9	0.19	0.15	0.07
10	0.12	0.06	0.15
11	0.10	0.05	0.01
12	0.08	0.03	0.01
13	0.07	0.03	0.01
14	0.07	0.03	0.05
15	0.06	0.05	0.02
16	0.07	0.05	0.02
17	0.09	0.04	0.01
18	0.09	0.04	0.003
19	0.09	0.04	0.01
20	0.13	0.06	0.05
21	0.21	0.18	0.27

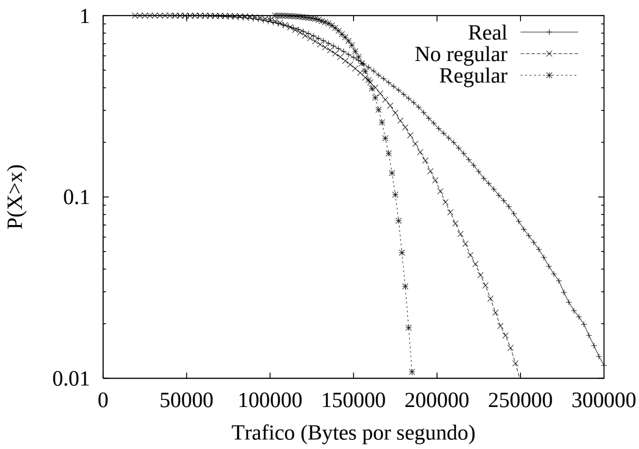


Figura 4.8: Función de supervivencia de la marginal del tráfico para ambos modelos

los bytes transferidos por conexión pueden modelarse con una variable aleatoria de cola pesada (con varianza infinita) debida a la enorme variedad de ficheros de información presentes en Internet, desde ficheros de texto a videos. Sin embargo, las características de dependencia a largo plazo del tráfico de Internet se deben a que los tiempos de servicio de estos flujos presentan también una cola pesada. Esto no solo se debe a la naturaleza de los tamaños de los ficheros (de cola pesada), sino que también se debe al limitado ancho de banda disponible en los enlaces, que hace que el tiempo de transmisión no pueda ser despreciable, así como a la dinámica de TCP, que contribuye a incrementar aún más la duración de las conexiones.

Nos planteamos ahora qué sucedería si se produjera un aumento considerable en el ancho de banda de la red. Este fenómeno ya está teniendo lugar ante el desarrollo de nuevas tecnologías ópticas que ofrecen tasas de transferencia de 1.6 Tbps en una única fibra empleando tecnología DWDM[2]. En la actualidad TCP es un protocolo que se encontraría muy limitado para sacar provecho a estas tecnologías debido al enorme producto *ancho de banda x retardo* presente en el segmento óptico. Hay que tener en cuenta que TCP se desarrolló para situaciones en las que los enlaces WAN en general tienen menor capacidad que las redes LAN de los extremos y donde las pérdidas de datos suelen deberse a congestión en las colas de los routers. TCP es capaz de adaptarse a grandes productos *retardo x ancho de banda* y de hecho lo hace exponencialmente, sin embargo, para aprovechar mejor estas tecnologías ópticas estas redes ofrecerán modos de transferencia por conmutación de ráfagas [67], los cuales permitirán la transferencia de ficheros a través de la red en un solo pulso óptico. En estas situaciones será la variabilidad de la distribución marginal del tráfico el factor limitante de las prestaciones.

Consideremos por ejemplo el caso de un enlace de capacidad 1 Gbps (Gigabit Ethernet punto a punto) en vez de los 4 Mbps del enlace de IP sobre ATM del cual se han tomado las trazas. Creamos una traza sintética con los mismos instantes de llegada de conexiones y el mismo tamaño de ficheros que en la traza real, pero transmitiéndolos a una velocidad de 1Gbps. Considerando una población grande es razonable mantener la suposición de un proceso de llegadas de Poisson y los ficheros transferidos como mucho aumentarían de tamaño con el incremento del ancho de banda disponible, dado que se deben principalmente a la oferta de ficheros existente y a cuáles solicitan los usuarios. En la figura 4.9 se ve el proceso de bytes por segundo resultante. Vemos que aparecen grandes picos debidos a la transmisión de grandes ficheros en un tiempo muy corto. Estos picos se encuentran muy alejados de la media del tráfico y son un claro síntoma de varianza infinita en la marginal del proceso.

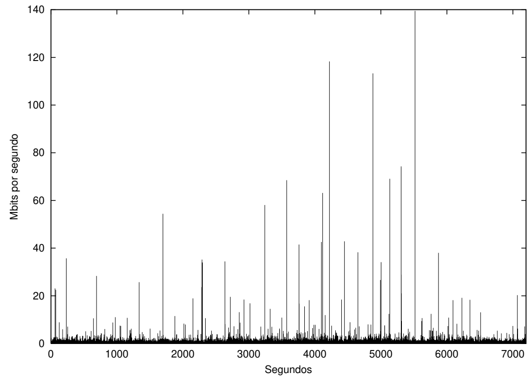


Figura 4.9: Tráfico empleando un enlace a 1Gbps

En la figura 4.10 tenemos la función de densidad de probabilidad de la marginal. Vemos que el tráfico resultante puede ser ajustado a una variable aleatoria α-stable (también representada). Una variable de este tipo en general no tiene expresión cerrada para la función de

[2]Dense Wavelength Division Multiplexing

densidad de probabildad, pero la transformada de Fourier de la misma (función característica) es del tipo:

$$E[e^{i\theta X}] = \begin{cases} e^{-\sigma|\theta|^{\alpha}(1-i\beta(sign\theta)tan\frac{\pi\alpha}{2})+i\mu\theta} & , \quad \alpha\neq 1 \\ e^{-\sigma|\theta|(1+i\beta\frac{2}{\pi}(sign\theta)ln|\theta|)+i\mu\theta} & , \quad \alpha = 1 \end{cases} \tag{4.2}$$

Donde para $\alpha = 2$ se reduce a una variable aleatoria Gaussiana. Sin embargo las variables aleatorias α-stables no gaussianas ($\alpha < 2$) presentan una mayor variabilidad, dado que no tienen segundo momento, es decir, tienen varianza infinita. El tráfico obtenido tras el enlace de 1 Gbps puede modelarse con una α-stable con parámetros $\alpha = 1.39$, $\beta = 1$, $\sigma = 43297$ y $\mu = 149147$ (Bytes por segundo).

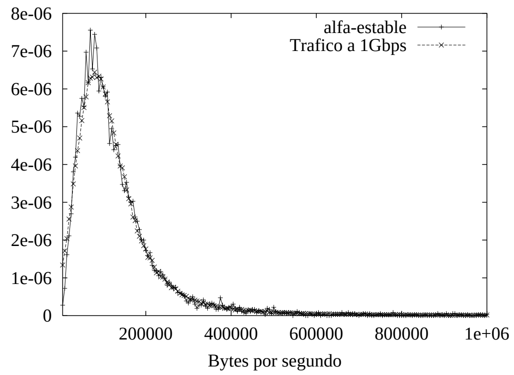


Figura 4.10: Función de densidad de probabilidad del tráfico por el enlace a 1Gbps

Ahora los ficheros con varianza infinita no se transmiten mediante conexiones de larga duración sino que viajan en cortos pulsos pero con varianza infinita en su amplitud. Así, la estructura de correlación de este nuevo proceso de tráfico degenera hacia la del proceso de llegadas de conexiones. Estimamos ahora un parámetro de Hurst de $H \approx 0.55$, un claro indicador de independencia, pero a costa de una mayor variabilidad en la marginal. Es decir, la naturaleza de cola pesada de los tamaños de los ficheros pasa de generar duraciones con varianza infinita y por tanto dependencia a largo plazo a producir una marginal con varianza infinita y por tanto una enorme variabilidad. Esta distribución de los tamaños de los ficheros es una característica inherente de la información disponible y no de Internet, por lo que en el futuro es de suponer que se mantendrá.

Para comprobar los efectos de este cambio de las características del proceso de tráfico realizamos nuevas simulaciones de servidor único con cola infinita empleando este tráfico de marginal α-stable (la traza a 1 Gbps) así como una versión del mismo de incrementos independientes (mediante mezclado aleatorio). Los resultados se ven en la figura 4.11, comparados con el tráfico de la traza original (a 4 Mbps).

Vemos en primer lugar que la traza α-stable conlleva un decremento dramático de las prestaciones del sistema en comparación con la versión a baja velocidad. Por otro lado, tanto la versión original de la traza α-stable como la de incrementos independientes dan resultados similares, lo cual significa que la estructura de correlación del tráfico no es el origen de la

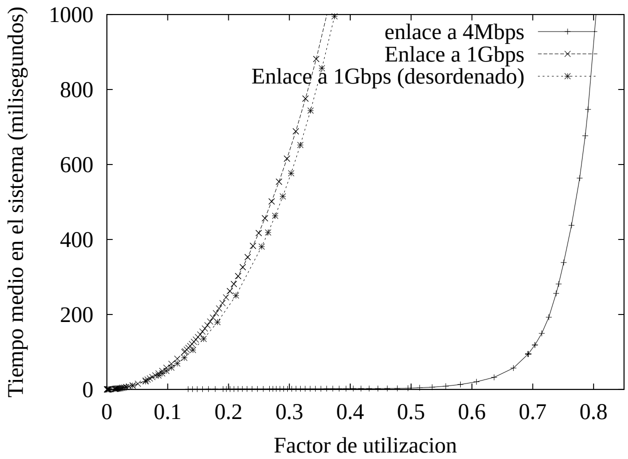


Figura 4.11: Evaluación de prestaciones de un sistema de un servidor ante tráfico a baja velocidad o α-estable (original y desordenado)

limitación de prestaciones, por lo que el factor determinante ha de ser el aumento de la variabilidad de la distribución marginal.

4.6 Conclusiones

Hemos visto que el tráfico de Internet presenta una distribución marginal con una variabilidad mayor que la resultante de modelos de flujos a tasa constante [27, 22] y que esto se debe a la variabiliad en las tasas de estos flujos. Así pues, hemos propuesto un modelo que se ajusta mejor al tráfico real, tomando como punto de partida el modelo de fuentes on-off pero cambiando la tasa constante por una tasa variable, la cual se puede modelar con una variable aleatoria de Weibull.

La variabilidad del tráfico resultante, más próximo al real, afecta a las prestaciones de elementos con encolado y puede representar una papel más importante que la dependencia a largo plazo a medida que los enlaces de red evolucionen hacia técnicas con conmutación de alta velocidad de flujos IP.

Capítulo 5

Estimación a-priori de ancho de banda para la reserva dinámica en enlaces de acceso de ISPs

5.1 Introducción

Tanto los accesos a Internet como los enlaces troncales de la misma pasan por continuas actualizaciones para conseguir dar a los usuarios unas velocidades de transferencia más elevadas. En el segmento de acceso, tecnologías como ADSL (Asymmetric Digital Subscriber Line) o HFC (Hybrid Fiber Coaxial) están permitiendo una mejora dramática de la velocidad ofrecida a los usuarios. De forma paralela las troncales van incorporando tecnologías ópticas en el rango de los Gbps. La unión entre la red de acceso y la troncal corre a cargo de los Proveedores de Acceso a Internet o ISPs (Internet Service Providers), los cuales ofrecen a los usuaarios residenciales la conexión con Internet a través de *enlaces de acceso*, como se muestra en la figura 5.1.

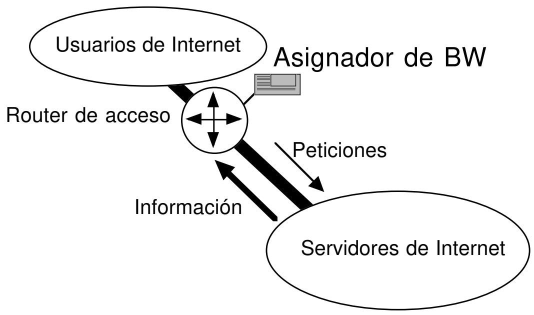


Figura 5.1: Configuración del enlace de acceso de un ISP

Un router de acceso realiza las funciones de unidad de interconexión entre la subred de acceso y la troncal de Internet, ofreciendo funcionalidades de traducción de direcciones y otras como control de acceso y contabilidad para gestionar el tráfico de usuario que circula por el enlace de acceso. De hecho, debido al veloz aumento del ancho de banda tanto en la red de acceso como en las troncales el ancho de banda del enlace de acceso se ha convertido en un recurso valioso, con un gran impacto sobre los beneficios obtenidos por los ISPs. Estos enlaces de acceso suelen ser alquilados a un gran operador o, en caso de que el ISP sea en si mismo un operador, alquilados a un segundo operador mediante acuerdos.

Generalmente para el ancho de banda del acceso se emplea una tarifa plana para la tasa configurada. Es decir, independientemente del volumen del tráfico el ISP debe pagar por una cantidad fija de ancho de banda sin posibilidad de variación según la hora o el día. La cantidad de ancho de banda contratado suele aproximarse con la ayuda de mediciones de tráfico de red como por ejemplo las estadísticas de uso de ancho de banda ofrecidas por el agente SNMP del router de acceso. Dado que el tráfico de Internet presenta un elevado grado de intermitencia una reserva constante de ancho de banda resulta claramente ineficiente. Por otro lado, en la mayoría de las situaciones, la determinación del ancho de banda para la reserva estática se realiza en base a técnicas heurísticas. Estas técnicas buscan maximizar la ganancia obtenida por multiplexación estadística mientras se mentiene una buena calidad de servicio ofrecida a los usuarios finales. Se basan normalmente en la experiencia del ISP en vez de en consideraciones analíticas. Esto se debe a la falta de reglas de dimensionamiento para enlaces de Internet, cuyo tráfico presenta no solo carácter auto-similar [14, 18] sino también una fuerte no estacionariedad [68].

Así pues, debido a la extremada variabilidad del tráfico de entrada, la cual complica el dimensionado del enlace mediante mediciones y técnicas *off-line*, parece más adecuado el empleo de técnicas dinámicas de reserva de ancho de banda para la optimización del uso del ancho de banda del enlace de acceso. Estas técnicas de reserva dinámica de ancho de banda pueden clasificarse en proactivas y reactivas. Ambos métodos con complementarios dado que las prestaciones de las ténicas reactivas, tales como por ejemplo la clase de servicio ABR (Available Bit Rate) de ATM, dependen fuertemente de la intermitencia y el retardo de propagación del tráfico en el enlace. Una ténica proactiva, que adapta los recursos de la red mediante *predicción del tráfico* puede servir para reducir la incertidumbre del asignador del ancho de banda, que podría predecir el uso de recursos antes de que se produjera el tráfico.

En el caso concreto que se plantea en este capítulo, con un enlace de acceso que ofrece conectividad con Internet a una población grande de usuarios, vemos que una coordinación entre los diferentes saltos en la red no es una hipótesis realista. De hecho, debido a la heterogeneidad y al gran número de posibles destinos a los que los usuarios pueden tener acceso, la reserva de recursos extremo a extremo está en la práctica lejos de ser realizable. Así pues, el router de acceso puede realizar reparto de ancho de banda solo en el enlace de acceso, que se puede considerar un subsistema aislado con un tráfico de entrada que es generado por un gran número de servidores en una Internet con disciplina de servicio *best-effort*. Estos servidores de Internet van a generar un tráfico compuesto principalmente por flujos (conexiones) TCP del servicio WWW, tal y como se ha visto en el estudio del enlace de acceso de la Universidad Pública de Navarra. En este escenario, las técnicas de reserva de ancho de banda por cada flujo presente son mecanismos proactivos que ofrecen ancho de banda bajo demanda, bien por flujo o por conjunto de flojos. Una vez que se conoce la necesidad de ancho de banda de un flujo se incrementa la capacidad del enlace para alcanzar los requisitos de calidad de servicio deseados. Estas técnicas de reserva por flujo están en consonancia con las nuevas técnicas de conmutación por etiquetas (label switching) tales como la propuesta del IETF conocida como Multiprotocol Label Switching (MPLS) [69, 70], en la cual se consideran los flujos de tráfico como las unidades respecto a las cuales hacer asignación de ancho de banda, en vez de serlo los paquetes. Por otra parte, un gran número de fabricantes están desarrolando soluciones de conmutación de *capa 4*, las cuales ofrecen reserva de recursos por flujo, de forma que se pueda cubrir la creciente demanda de calidad de servicio diferenciada.

Para poder ofrecer una solución de conmutación de flujos se precisa en primer lugar un conocimiento detallado de las características estadísticas de estos flujos. Más concretamente, una técnica de reserva de recursos por flujo requiere un conocimiento a-priori de los recursos del enlace consumidos por cada flujo. En los capítulos anteriores se han realizados extensivos análisis de trazas de tráfico del enlace IP sobre ATM de la UPNA. En estos análisis se ha visto que el protocolo TCP supone más del 95% de los bytes transmitidos hacia los usuarios, de estos más del 76% se deben a transacciones Web solicitadas por los usuarios, que en número de transacciones (conexiones TCP) son más del 90% de las existentes. Cifras similares se pueden encontrar en gran número de estudios, tanto en entornos académicos como industriales [13, 64, 65].

En consecuencia, para llevar a cabo reservas de ancho de banda proactivas para flujos de Internet de forma individual el sistema encargado de la asignación debe conocer la tasa de transferencia requerida por cada conexión TCP. Más concretamente, si llamamos r_i a la tasa de transferencia de la conexión TCP i-ésima, buscamos una función f tal que

$$r_i = f(\alpha_1, \ldots, \alpha_n) \tag{5.1}$$

donde $\alpha_1, \ldots \alpha_n$ son un número de parámetros de la conexión conocidos por el asignador de ancho de banda *a-priori*, es decir, después de que se establezca la conexión pero *antes* de que tenga lugar la transferencia de datos. Por ejemplo, uno de estos parámetros podría ser el tamaño de la transferencia a realizar en la conexión, que para el caso en estudio (Web) viene indicado por el protocolo HTTP, o puede ser la dirección IP del servidor o el tamaño de la ventana anunciada por el cliente. Hay que resaltar que la estimación de forma exacta de r_i es claramente irrealizable dado que el enlace de acceso se conecta a una Internet con disciplina best-effort. Así pues, debido al gran número de destinos posibles, cada uno de los cuales puede presentar un cuello de botella diferente en el camino de cliente a servidor, existe una cierta incertidumbre en la tasa de transferencia que alcanzará una conexión TCP dada.

En este capítulo nos centraremos en la estimación a-priori de la tasa de transferencia (r_i) de las conexiones TCP (Web) en un enlace de acceso que sirve a una población genérica de usuarios conectados a Internet. Aplicaremos conceptos de información mutua entre variables aleatorias para mostrar que el estimador simple de retardo extremo a extremo (RTT), basado en el intercambio de paquetes con el flag de SYN y presentado en el capítulo 2, reduce considerablemente la incertidumbre asociada a r_i. El tamaño de la conexión o la hora del día de su comienzo aportan una cantidad mucho más reducida de información respecto a la tasa de transferencia alcanzable. Basándonos en estos datos tomaremos la densidad de probabilidad de la tasa de transferencia condicionada al RTT, la cual será obtenida a-priori. Un asignador estático que reserva la media de dicha variable aleatoria condicionada al RTT más n veces la varianza de la misma, con diferentes valores de n, ofrece un ahorro considerable en el ancho de banda en comparación con una asignación basada en el cociente entre el tamaño de la ventana y el RTT.

La medida empleda en este capítulo corresponde al día completo 14/02/2000 (lunes). En esta traza se registraron más de 1 millón de conexiones TCP (16 millones de paquetes IP).

El resto del capítulo se encuentra organizado de la siguiente forma: en la sección 5.2 analizamos los posibles parámetros de la conexión que puedan afectar a la tasa de la misma y se puedan conocer a-priori. En la sección 5.3 presentamos las características estadísticas del estimador del RTT y las densidades de probabilidad condicionada de la tasa r_i. La sección 5.4 se dedica a la comparación de prestaciones de un asignador de tasa de pico frente a un asignador basado en el RTT. Finalmente la sección 5.5 presenta las conclusiones.

5.2 Estimación de la tasa de transferencia de conexiones en base a parámetros de conexión conocidos a-priori

Vamos a considerar el caso de un enlace de acceso para el cual los clientes están dirigiéndose a destinos aleatorios en la Internet. Deseamos estimar la tasa de transferencia de las conexiones establecidas por estos usuarios. Mientras que los clientes se conectan al enlace a través del router de acceso el servidor se encontrará localizado en otro lugar de Internet. Así pues, el asignador de ancho de banda del ISP es un subcomponente del router de acceso. Para esta configuración de red en particular ya se ha visto que el tráfico dominante es el dirigido de servidores a clientes (capítulo 2), mientras que el sentido ascendente se emplea principalmente para el transporte de paquetes de confirmación de datos (ACK). Así pues, solo vamos a considerar el sentido descendente (de servidores a usuarios) como el segmento con ancho de banda restringido que requiere una asignación proactiva del mismo.

Vamos a estudiar los parámetros de conexión accesibles a-priori y que vayan a influir en su tasa de transferencia. Al decir *accesibles a-priori* nos referimos a aquellos parámetros que se conocen justo después de la fase de establecimiento de la conexión TCP. Presentaremos en primer lugar un análisis cualitativo para eliminar a algunos de los candidatos. Se puede comenzar considerando parámetros independientes de la conexión como pueden ser la hora de inicio de la misma o la carga del enlace en ese momento. Ambos parámetros se encuentran relacionados dado que la carga del enlace fluctúa según la hora del día. Para un enlace con una carga reducida como el que se está estudiando el cuello de botella se encontrará siempre en otro lugar de la red y por lo tanto podemos limitarnos a considerar la hora de inicio de la conexión. Dado que la mayor parte de los usuarios acceden a servidores localizador en EE.UU. encontraremos mejores condiciones de congestión en las primeras horas de la mañana, las cuales coinciden con las horas de noche en EE.UU.

Por otra parte, existen otros parámetros dependientes de la conexión que pueden influir en sus prestaciones, como pueden ser la dirección IP origen y destino, el RTT y el tamaño de la transferencia. La dirección IP origen añade poca información respecto a la tasa de transferencia, dado que nuestros usuarios se encuentra conectados a un a LAN a 10/100 Mbps dentro del campus de la universidad. Estas LANs se encuentran conectadas principalmente a través de puentes transparentes y tienen condiciones de carga similares lo cual da como resultado que no haya diferencias apreciables entre los diferentes puntos de la red correspondientes a las diferentes direcciones IP. Tanto la dirección IP del servidor como la estimación del RTT dan información sobre la localización del mismo, la cual tendrá una influencia clara en las prestaciones. Sin embargo, no es factible deducir la localización del servidor a partir de su dirección IP por dos razones: en primar lugar la dirección IP no está ligada necesariamente al número de saltos o de enlaces congestionados que el paquete tenga que atravesar para recorrer el camino de cliente a servidor; en segundo lugar la "distancia" del servidor en términos de ancho de banda y retardo depende de la situación de congestión instantánea de la red, junto con los procedimientos dinámicos de enrutamiento en la Internet, mientras que la dirección IP es fija. Finalmente, dado que el número de direcciones IP posibles es muy grande la implementación de un asignador de ancho de banda basado en direcciones IP destino es de difícil realización. Así pues, la estimación del RTT es más útil para la caracterización de las condiciones del camino entre cliente y servidor. Por otra parte, el tamaño de la transferencia añade información sobre el comportamiento dinámico de la conexión TCP. Mientras que conexiones TCP largas son mejores candidatas para alcanzar un estado estacionario las conexiones cortas nunca abandonarán la fase de slow start, dando así una tasa de transferencia menor en iguales condiciones de RTT.

Otros parámetros como la probabilidad de pérdida de paquetes en todo el camino o el ancho de banda del cuello de botella no pueden ser estimados a-priori, aunque añadirían mucha información sobre la tasa de transferencia que pueda alcanzar la conexión [45]. Así pues, el conjunto de parámetros utilizables se reduce a tres: el tamaño de la transferencia, que se puede obtener fácilmente de la respuesta HTTP del servidor, la hora del día en la que se da la conexión y la estimación del RTT.

Respecto al estimador del RTT, éste se basa en el intercambio inicial de paquetes de establecimiento con la bandera de SYN. Hemos visto en el capítulo 2 las principales características de este estimador (figura 2.4).

Para evaluar la dependencia de la tasa de transferencia frente al tamaño, la hora de comienzo y RTT estimado emplearemos el concepto de información media mutua entre variables aleatorias [71].

Definición: Sea la variable aleatoria X, que puede tomar los valores $x_1, \ldots, x_n$ con probabilidades $p(x_1), \ldots, p(x_n)$ respectivamente. Llamamos entropía de la variable X al valor $H(X)$ definido como sigue:

$$H(X) = \sum_{i=1}^{n} p(x_i) log(1/p(x_i)) \tag{5.2}$$

Definición: Sea la variable aleatoria Y que puede tomar los valores $y_1, \ldots, y_m$ con probabilidades $p(y_1), \ldots, p(y_m)$ respectivamente. $p(x_i, y_j)$ es la probabilidad de que tengan lugar el evento x_i y el evento y_j. $p(x_i/y_j)$ es la probabilidad del evento x_i condicionada al evento y_j. Se define la entropía (condicional) de X condicionado a Y como sigue:

$$H(X/Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j) log(1/p(x_i/y_j)) \tag{5.3}$$

Definición: Se llama información media mutua entre las variables aleatorias X e Y a:

$$I(X;Y) = H(X) - H(X/Y) \tag{5.4}$$

Si interpretamos $H(X/Y)$ como la cantidad media de incertidumbre en X tras ver el resultado de Y y $H(X)$ como la cantidad media de incertidumbre antes de dicha realización entonces $I(X;Y)$ es la información media obtenida por cada valor de X por vada valor conocido de Y. Así pues, la información media mutua ofrece un medio para evaluar la dependencia entre variables aleatorias. Si X e Y son independientes se concluye directamente que $H(X/Y) = H(X)$ y la información media mutua entre X e Y es cero.

Para medir la información media mutua entre la tasa de transferencia y el RTT estimado, la hora del día del comienzo de la conexión y el tamaño de la transferencia comenzaremos tomandos versiones discretas de dichas variables aleatorias. Para ello, tomaremos intervalos de 1000 bps para la tasa, 100 ms para el RTT, 10 KBytes para el tamaño y una hora para el instante de comienzo de la conexión. La tabla 5.1 presenta la información media mutua y la entropía condicional de la tasa de transferencia r_i.

$H(r_i/RTT)$	$H(r_i/hora)$	$H(r_i/tam)$
4.941	5.346	5.341
$I(r_i, RTT)$	$I(r_i, hora)$	$I(r_i, tam)$
0.535	0.13	0.135

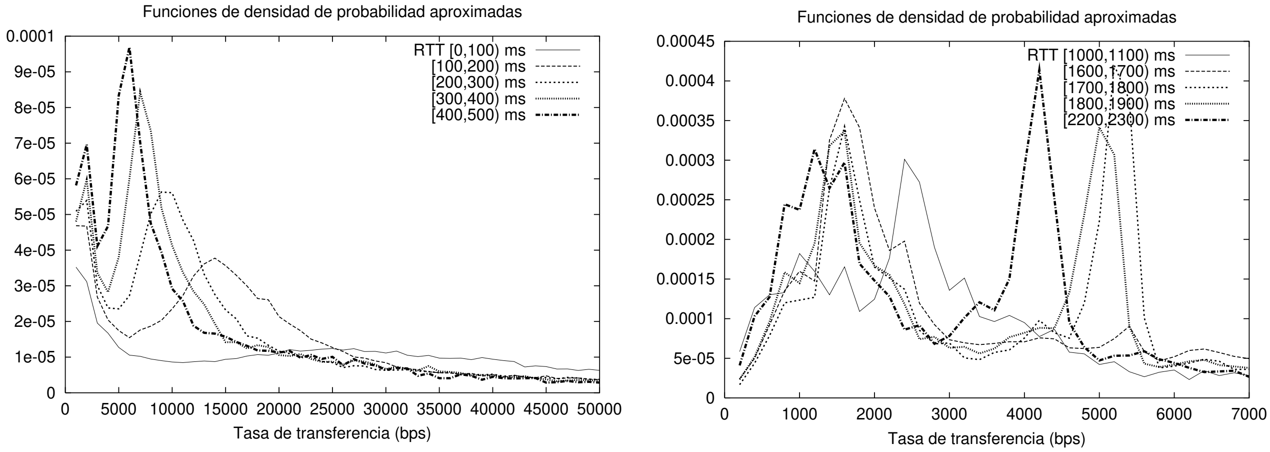
Tabla 5.1: Información y entropía condicional de r_i

La entropía de la variable aleatoria r_i, $H(r_i)$, es 5.476 bits. Destaca el RTT estimado como la variable con una mayor contribución a la hora de reducir la incertidumbre sobre la tasa r_i, seguido por el tamaño de la transferencia y la hora del establecimiento. La información media mutua entre la tasa y la hora del día toma las 24 horas que componen la traza de paquetes TCP. Un examen más detallado de la misma revela que, como era previsible, la mayor parte de las conexiones tienen lugar en el intervalo diurno (de 9AM a 9PM). Si restringimos el análisis a dicho intervalo (con lo que H(SCR) pasa a ser H(SCR/Dia)=5.324 bits) entonces la información aportada por la hora en que se da la conexión cae a 0.082 bits. Dado que durante la noche el enlace de acceso está muy descargado no tiene sentido emplear un asignador de ancho de banda en tal situación.

En lo que respecta a la información media entre la tasa y el tamaño vemos que las conexiones largas no alcanzan necesariamente una tasa superior, aunque son mejores candidatas para haber completado la fase de slow start y entrar en un régimen estacionario. Como conclusión, vemos que la estimación realizada del RTT es la variable que aporta la mayor cantidad de información respecto a la tasa de la conexión.

5.3 Estimación del RTT y distribución de la tasa de conexión

En esta sección nos centramos en el análisis de la estimación del RTT, como variable conocida a-priori y que sirve para reducir la incertidumbre de la tasa de transferencia alcanzable por conexiones TCP. Para ello dividimos los valores posibles de RTT en intervalos de 100 ms de la forma $(j100, (j+1)100]$ donde el índice j toma valores en el rango $0, \ldots, 22$ ms. Denotaremos al intervalo j como RTT_j. A continuación consideramos la densidad de probabilidad (condicional) de la tasa r_i respecto al intervalo de RTT RTT_j, lo cual se presenta en la figura 5.2.



(a) (b)

Figura 5.2: Distribución condicional de la tasa de transferencia para (a) RTT pequeño, (b) RTT grande

La figura 5.2 presenta varias regiones con forma de campana, así como una fuerte dependencia de la tasa con el RTT para valores pequeños de éste. Para valores grandes de RTT la dependencia con el mismo no es tan grande. Valores altos de RTT se corresponden con enlaces congestionados de Internet, en los cuales las conexiones probablemente sufran grandes pérdidas y elevados tiempos de respuesta del servidor, lo cual afecta seriamente a la tasa de la conexión.

Denotaremos a partir de ahora con μ_j y σ_j a la media y la desviación de la distribución (condicional) de la tasa respecto al RTT_j, respectivamente. En un caso ideal, sin pérdidas ni variaciones de reardo y con un tiempo de respuesta del servidor despreciable, la red debería comportarse como un sistema determinista con una correspondencia uno a uno entre la tasa que alcanza un fichero y el RTT. Sin embargo, el valor de tasa en el caso real es una variable aleatoria. Si se dispone de una caracterización a-priori de la media y desviación de la tasa de transferencia condicional podemos plantearnos una predicción de la tasa basada en el conocimiento del RTT de la conexión, el cual puede determinarse *antes* del comienzo de la fase de transferencia de datos (figura 2.4). De hecho, en la siguiente sección presentamos un estimador para un asignador de ancho de banda de conexiones TCP basado en el RTT.

5.4 Evaluación de prestaciones de un sistema de reserva de ancho de banda basado en el RTT

Las técnicas de conmutación de flujos se presentan como una solución prometedora que permita maximizar la utilización del ancho de banda a la vez que ofrecer una calidad de servicio diferenciada en la Internet. En esta sección presentamos un asignador de ancho de banda basado en la estimación del RTT presentada con anterioridad. Para comparar las prestaciones del asignador propuesto medimos el porcentaje de ancho de banda desperdiciado y no satisfecho para tres asignadores diferentes:

- Asignador de ancho de banda basado en el RTT: toma la estimación del RTT (RTT_i) como parámetro de entrada y devuelve un valor de ancho de banda para el flujo TCP i, basado en la densidad de probabilidad (condicionada) de la tasa respecto al RTT.

- Asignador de tasa pico: toma el tamaño de la ventana anunciada por el cliente para el flujo i (W_i) y el valor estimado del RTT (RTT_i) como datos de entrada y devuelve una asignación de ancho debanda igual a W_i/RTT_i. Este es el valor teórico máximo de tasa que puede alcanzar una conexión TCP con una ventana anunciada de valor W_i.

- Asignador estático: Consideramos el grupo de asignadores de ancho de banda que reservan la misma cantidad para cada flujo.

El asignador basado en el RTT basa su regla de funcionamiento en las curvas de la figura 5.2. Concretamente tiene en cuenta la división mencionada del RTT en regiones correspondientes a los intervalos $(j100,(j+1)100]$, donde j toma valor en el rango $0,\ldots,22$ ms (RTT_j), con parámetros de la distribución de la tasa: media μ_j y desviación σ_j. Como primera aproximación tomamos la siguiente regla de asignación de ancho de banda:

Para el RTT del flujo i seleccionamos el intervalo RTT_j tal que $RTT \in RTT_j$. La asignación de ancho de banda para el flujo i es $r_i = \mu_j + n\sigma_j$ donde n es una constante.

Mediante el ajuste del parámetro n podemos conseguir un resultado de compromiso entre maximizar la ganancia por multiplexacion estadística y la calidad de servicio asignada al flujo. Valores pequeños de n darán una mayor eficiencia del ancho de banda, a costa de buffers extra. Por otro lado, este asignador se basa en un algoritmo muy simple, lo cual permite su empleo en enlaces de gran velocidad.

Para medir sus prestaciones definimos dos variables. Sea r'_i la tasa alcanzada por la conexión *a-posteriori*, es decir, la tasa observada sin asignador de ancho de banda, valor que se obtiene directamente de la traza de tráfico. Dado que el enlace de acceso analizado no está muy cargado, y por lo tanto no es el cuello de botella, esta variable indica la tasa máxima que puede alcanzar un flujo TCP en particular en el enlace de acceso. Sea r_i la signación de ancho de banda realizada por el algoritmo. Este valor de r_i es determinado *a-priori*, es decir, antes de que la transferencia tenga lugar.

Definición: Para los flujos i que cumplen que $r'_i > r_i$ definimos el *porcentaje del ancho de banda no satisfecho* Δ_i como:

$$\Delta_i = \frac{r'_i - r_i}{r'_i} \tag{5.5}$$

Para aquellas conexiones con $r'_i \leq r_i$ la variable Δ_i vale por definición cero. El valor de Δ_i depende del escogido para n, cuanto mayor sea n menor será el porcentaje de ancho de banda no satisfecho para un flujo. La figura 5.3 presenta el porcentaje *medio* de ancho de banda no satisfecho frente a n, promediando Δ_i en el conjunto completo de flujos presentes en la traza.

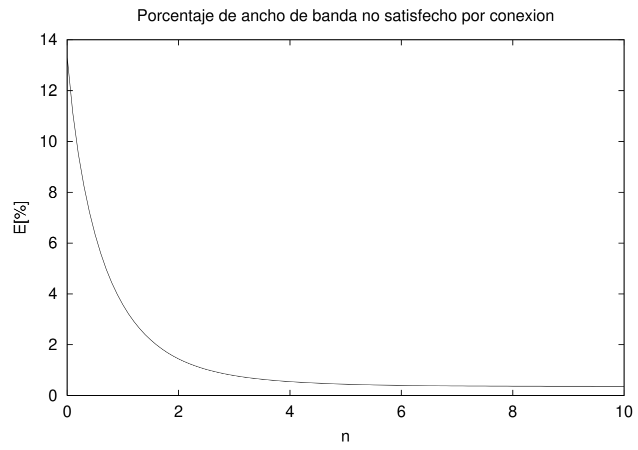


Figura 5.3: Porcentaje medio de ancho de banda no satisfecho

Vemos que para un valor de $n = 2$ se alcanza un ancho de banda no satisfecho próximo al 1.7%.

Definición Para los flujos i que cumplen que $r_i > r'_i$ definimos el *porcentaje de ancho de banda desperdiciado* o_i como:

$$o_i = \frac{r_i - r'_i}{r_i} \tag{5.6}$$

Para aquellas conexiones con $r_i \leq r'_i$ la variable o_i vale por definición cero. La figura 5.4 presenta el porcentaje *medio* del ancho de banda desperdiciado frente a n. Un valor de n en el intervalo $(0, 2)$ corresponde a un desperdicio entre el 40% y el 73% y un porcentaje no satisfecho entre el 13% y el 1.7%.

En comparación, para el asignador de tasa pico el porcentaje medio de desperdicio sube al 94.36%. Por otra parte, el porcentaje de ancho de banda no satisfecho es del 0.11%. El asignador basado en el RTT permite un mayor grado de flexibilidad a partir del parámetro n, que se traduce en un rango de puntos de operación que se pueden seleccionar en función de la

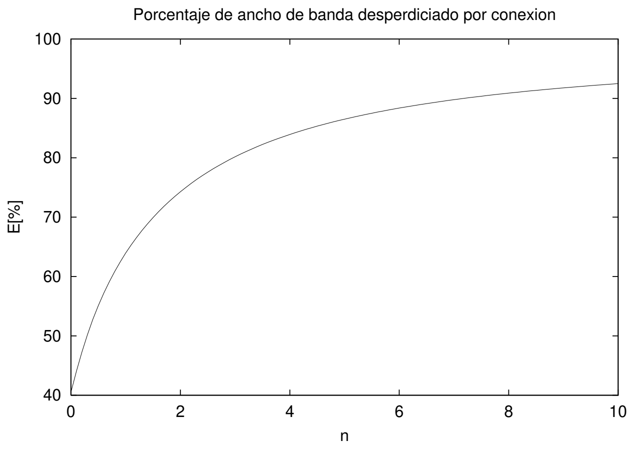


Figura 5.4: Porcentaje medio de ancho de banda desperdiciado

utilización del enlace y la calidad de servicio deseados, dando un ancho de banda desperdiciado inferior al de la asignación de r tasa pico. Por último, tomamos el grupo de asignadores estáticos (el mismo ancho de banda para todos los flujos) variando el ancho de banda asignado y los comparamos con el asignador propuesto basado en el RTT. La figura 5.5 muestra el porcentaje medio de ancho de banda no satisfecho frente al porcentaje medio de desperdicio de ancho de banda, con escala logarítmica en el eje y. Los resultados muestran que el asignador basado en el RTT mejora las prestaciones de cualquier asignador estático posible, dado que para el mismo porcentaje de ancho de banda desperdiciado el porcentaje de ancho de banda no satisfecho en menor.

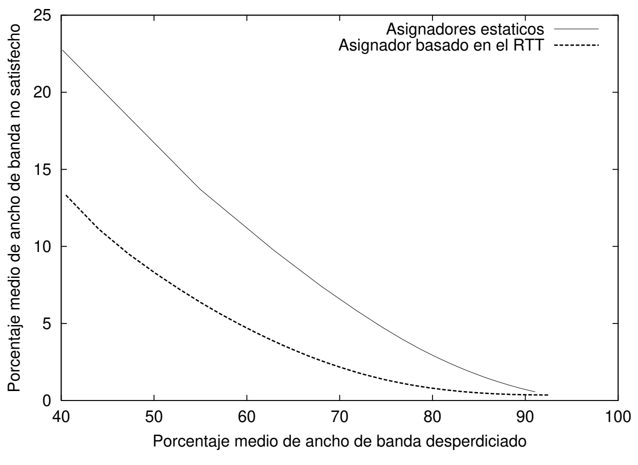


Figura 5.5: Comparación del asignador basado en el RTT con los asignadores estáticos

5.5 Conclusiones y trabajo futuro

En este capítulo hemos comprobado que es posible realizar estimaciones a-priori de la tasa de transferencia de conexiones TCP orientadas al dimensionado y reserva de ancho de banda. Se ha propuesto y estudiado un asignador basado en una estimación del RTT, el cual mejora las prestaciones de un estimador basado en la tasa de pico y de asignadores estáticos.

Esta asignación basada en el RTT se realiza mediante un conjunto simple y eficiente de reglas que parten de la media y la desviación de las distribuciones de la tasa de transferencia condicionada al RTT. Estas reglas se obtienen mediante el análisis de una traza de tráfico de gran tamaño y son implementadas *manualmente* dentro del asignador. Así pues, los parámetros están optimizados para un enlace en concreto. Sin embargo, dado que como hemos visto es posible realizar un estimador simple de la tasa de los flujos nos planteamos la posibilidad de obtener las reglas de dimensionamiento de forma automática. Para ello se plantea la posibilidad de emplear otros estimadores de la forma:

$$r_i = f(\alpha_1, \ldots, \alpha_n) \tag{5.7}$$

donde f es la función que provee la estimación, que se obtendría mediante análisis *on-line* del tráfico, y no a partir de una traza como se ha realizado en este trabajo. Los resultados de este estudio indican que existen ciertos parámetros que aportan información sobre el valor de r_i y que por tanto pueden emplearse como descriptores de la conexión.

Capítulo 6

Conclusiones y trabajos futuros

En este trabajo se ha obtenido una caracterización y modelado del tráfico de Internet. En el proceso se han realizado las siguientes contribuciones:

- Un análisis macroscópico del tráfico ha mostrado que la mayor parte del mismo se debe a conexiones TCP.
- El análisis de estas conexiones TCP indica que la tasa de transferencia de las mismas puede modelarse mediante una variable aleatoria de Weibull.
- Hemos visto que la tasa de transferencia depende fuertemente del retardo extremo a extremo (RTT) de las conexiones y que las ráfagas que se reciben en el cliente no siguen una evolución exponencial. Ante estas características de la tasa y la intermitencia hemos visto que se puede aplicar un modelo de proceso estocástico con restricción (σ, ρ) con σ en torno a un par de paquetes y ρ dependiente del retardo extremo a extremo de la conexión.
- Hemos realizado una propuesta de modelado de tráfico que mejora el modelo $M/G/\infty$ y que se ajusta mejor al tráfico real. Este modelo añade una variable aleatoria de Weibull a la tasa de transferencia del estado ON, permitiendo una mayor variabilidad en el tráfico resultante.
- El modelo propuesto captura mejor las características de variabilidad, que se han mostrado como un factor determinante en futuras redes de alta velocidad, donde se reducirá el efecto de dependencia a largo plazo del tráfico a costa de una mayor variabilidad.
- Finalmente, la caracterización del tráfico obtenida ha permitido definir una clase de asignadores de ancho de banda para enlaces de acceso en base a los parámetros de las conexiones que más información aportan sobre las mismas y que se pueden conocer antes de que tenga lugar la transferencia de información del flujo.

A partir de aquí el trabajo se continuará a través de dos aspectos: Por un lado se pretende extender el estudio de algoritmos de asignación de ancho de banda basados en predicción, bien a nivel de flujos, usuarios o tráfico total agregado. Por otro lado se están estudiando los efectos que puede tener la combinación de procesos de tráfico de entrada fractales con las nuevas propuestas de técnicas de transmisión a ráfagas (Optical Burst Switching) en redes WDM [72], así como la posibilidad de empleo de algoritmos de estimación de tráfico en la mejora de su eficiencia.

Apéndice A

Herramienta IPmiser

Las redes de alta velocidad que actualmente constituyen el núcleo de la Internet transportan un tráfico muy poco predecible, por lo que el análisis y dimensionamiento de estas redes se complica. Así, el gestor de red realiza una primera aproximación de diseño asignando un ancho de banda determinado a los enlaces que pone en servicio. Una vez se acomete este primer dimensionamiento es preciso disponer de un equipo que mida las prestaciones reales del enlace, mediante el análisis del tráfico que circula.

Ante esta necesidad surgió el proyecto IPmiser, realizado conjuntamente por investigadores del Grupo de Redes, Sistemas y Servicios de la Universidad Pública de Navarra y de Telefónica I+D S.A. El equipo resultante ha permitido la captura y análisis de las trazas que se han presentado en este trabajo. A continuación se presentarán las características que se exigieron para este equipo y de forma breve cuál fue la estructura de diseño que se tomó.

A.1 Características exigidas al sistema

A.1.1 Monitorización continuada

En la actualidad se utilizan sistemas basados en sondas de monitorización (sniffer) que en mayor o menor medida son capaces de recoger estadísticas de tráfico de la red durante un tiempo limitado. Estos sistemas se construyen con un hardware dedicado que permite la captura de paquetes con una resolución muy alta. Sin embargo, una de sus principales limitaciones es la imposibilidad de realizar una captura continua, en el rango de varias horas, días o incluso semanas. Este aspecto es especialmente importante en redes IP, donde el tráfico es claramente no homogéneo. De hecho, el tráfico de la Internet presenta características autosimilares. En este tipo de procesos estocásticos podemos encontrar ráfagas muy grandes en intervalos de tiempo considerables. Estas ráfagas no solo provocan la sobrecarga de las colas de los routers sino que además sobrepasan la escala de tiempo de captura de sondas tradicionales, haciendo que la estimación de parámetros del proceso de tráfico con trazas cortas sea errónea.

La figura A.1 muestra el tráfico del enlace de la UPNA durante una semana, donde se observa claramente la existencia de ráfagas en escalas de tiempo muy grandes. A la hora de monitorizar este enlace es necesaria la monitorización continua y no limitada en el tiempo, ya

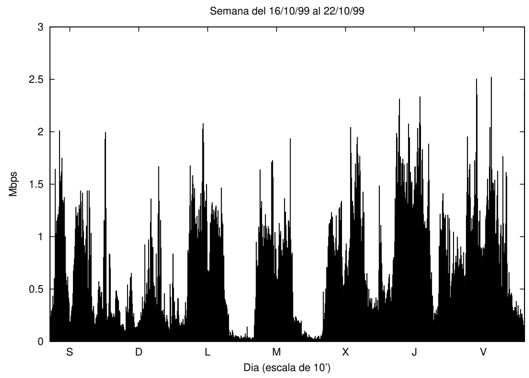


Figura A.1: Tráfico del enlace de la UPNA en una semana

que de otra manera perderíamos los picos de tráfico, que es precisamente lo más interesante a efectos de dimensionamiento de red.

A.1.2 Monitorización distribuida

Por otro lado, la creciente complejidad de las redes de comunicaciones impulsa la idea de monitorización distribuída en contraposición con el carácter centralizado de los equipos de medida actualmente existentes. De esta manera es posible tener un número de supervisores de red que se responsabilizan del correcto funcionamiento de cada una de las subredes. Para hacer esto posible, el sistema de monitorización debe permitir la obtención de vistas de los datos que almacena. La idea consiste en que cada gestor de red tenga acceso únicamente al tráfico que genera su subred pero no al de las otras subredes. En el entorno de RedIRIS, por ejemplo, se podría asignar a cada gestor de subred (esto es, los responsables de los centros de cálculo de cada una de las universidades conectadas) la vista del tráfico que genera su universidad, para que así tome las medidas oportunas si detecta generación excesiva de tráfico o servidores no deseados, intrusos, etc.

Finalmente, los parámetros de monitorización que los gestores de red exigen conocer son cada vez más variados. No solo bastan lo típicos parámetros de número de bytes por protocolo y servicio, número de conexiones, etc. Hoy en día debemos estar preparados para calcular en tiempo real parámetros como el jitter de los paquetes de una determinada conexión de Voz sobre IP, realizando además las funciones de tarificación y llevando cuenta de miles de conexiones concurrentemente. En un entorno de alta velocidad esto complica en gran manera el diseño de estas herramientas de monitorización, verdaderos sistemas operativos en tiempo real para propósito específico de monitorización.

A.2 Arquitectura de diseño del sistema IPmiser

En resumen, tres factores impulsan el diseño de un nuevo equipo de monitorización, el sistema IPmiser:

- La falta de *capacidad de almacenamiento* en los actuales sistemas de monitorización

- La falta de *flexibilidad* a la hora de permitir que distintos gestores tengan acceso de forma segura y fiable a la información de monitorización, por ejemplo, distintos proveedores de Internet pueden querer información del tráfico que generan, pero sin que sus competidores accedan a esta información.

- La falta de programabilidad de las herramientas de monitorización, para las que añadir nuevos parámetros resulta imposible, quedándose obsoletas en muy breve plazo.

El sistema *IPmiser* responde a esta triple necesidad: mayor capacidad de proceso y almacenamiento, distribución de la información de forma fiable y segura entre varios gestores y capacidades de reconfiguración de parámetros. La innovación tecnológica del IPmiser se centra en dos aspectos:

- Arquitectura optimizada para la monitorización de un enlace ATM de alta velocidad (155.5 megabits por segundo): El sistema consiste en un hardware dedicado conectado a una estación de trabajo con software de captura de datos y unidad de almacenamiento masivo. Esta unidad de almacenamiento masivo permite el almcenamiento de todas las cabeceras de protocolo IP, TCP, UDP e ICMP aparte de las marcas de tiempo de cada una de las celdas ATM que circulan por el enlace. Así, es posible realizar una auditoría perfecta del enlace a monitorizar, aparte de los estadísticos que, en tiempo real o en diferido, sirve el sistema IPmiser.

- Servidor de información de monitorización basado en tecnología de navegadores de Internet: El propio sistema IPmiser se convierte en servidor de información de monitorización en tiempo real y de días pasados. El sistema permite una jerarquía de supervisores de subredes en lugar de un supervisor centralizado como hasta la fecha. Por otro lado, la información de monitorización se distribuye de forma segura, utilizando mecanismos de autentificación y confidencialidad de datos. Además, la consola de monitorización no está ligada a una máquina en concreto sino que está distribuida en la Internet en forma de navegador estándar de Web (con soporte de Java).

A.3 Esquema del sistema IPmiser

El diagrama de bloques del sistema IPmiser se muestra en la figura A.2.

El sistema consta de un elemento hardware dedicado basado en computación paralela que monitoriza los dos sentidos del tráfico de un enlace ATM a 155Mbps (OC-3) sin pérdidas de celdas mediante la extracción de una fracción de la potencia óptica de la señal que circula por las fibras con dos divisores de potencia, de forma que si las pérdidas introducidas por los mismos no incumplen los requisitos de potencia del enlace nos encontramos ante una monitorización no intrusiva. Esta arquitectura evita los errores de medida que podrían darse con el empleo de los habituales "packet filters" [73].

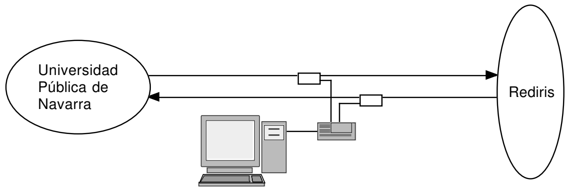


Figura A.2: Diagrama de bloques de *IPmiser*

Este elemento hardware extrae información de los niveles ATM, IP y transporte y construye para cada trama AAL5 una trama de datos con información referente a las direcciones IP origen y destino, tamaño de paquete y otra información del nivel de red y de transporte (puertos y code bits) así como timestamps muy precisos (resolución de 14/12 de μseg) de la llegada de las celdas que componen la trama, y envía esta trama de información a una estación de trabajo por un enlace ATM dedicado a 155Mbps. Esta técnica permite analizar grandes cantidades de datos sin interrupción dado que las mediciones no tienen elevados requisitos de almacenamiento.

Esta estación de trabajo dispone de dos procesadores UltraSPARC-II a 200MHz dedicados a la tarea de procesar con el mínimo retardo posible la información obtenida del elemento hardware y crear un registro global de tráfico, así como una representación online del estado de la red, siendo capaz de distinguir y configurar diferentes dominios de monitorización y presentar al usuario un interfaz gráfico multiplataforma para el acceso a la información. El software de procesado de la estación se basa en técnicas multi-hilo con asignación de clases de scheduling de tiempo real para obtener un procesado que optimiza la arquitectura paralela del harware y el sistema operativo de la estación.

Las grandes ventajas del empleo de hilos en lugar de procesos concurrentes son la comunicación entre hilos y la planificación de los mismos. La comunicación entre procesos requiere copias en espacio del kernel mientras que los hilos comparten todo el espacio de memoria por lo que la comunicación es a través de variables compartidas. En temas de planificación, los hilos representan entidades de proceso más "ligeras" que en el caso de procesos tradicionales de UNIX, lo cual permite que el paso de un hilo a otro en la CPU se realice a mayor velocidad dado que no han de realizarse costosos cambios de contexto, los hilos pertenecen a un mismo proceso por lo que el mapa de memoria se mantiene el mismo, así como otras características. Solo variables pertenecientes a cada hilo han de ser intercambiadas, como son la pila, los registros o el contador de programa.

Estas técnicas permiten un desarrollo más simple mediante tareas individuales que se ejecutan en hilos independientes y que aprovechan la arquitectura paralela del sistema con dos microprocesadores. El paralelismo de la computadora se controla mediante clases de prioridad entre hilos para dar servicio siempre en primer lugar a los hilos que requieren la mayor atención para poder realizar las tareas básicas de procesado, dejando el tiempo libre de CPU para hilos de usuario o hilos encargados de actualización del interfaz gráfico de información.

Los hilos principales, con prioridad de tiempo real, implementan un sistema de clasificación de paquetes que hace que no se pierda información de monitorización. Precisamente, este algoritmo de clasificación es uno de los principales méritos del sistema. En una red de alta velocidad puede ser necesario calcular miles de parámetros en tiempo real. Por ejemplo puede interesar conocer en detalle el tráfico generado por cada uno de los usuarios, que pueden ser muchos. También puede ser necesario conocer el estado de cada una de las conexiones TCP que circulan por el enlace, para saber exactamente el número de bytes o duración, o para realizar funciones de tarificación y control de la calidad de servicio.

En los trabajos [38, 74, 65] se pueden encontrar resultados del empleo que ha hecho el Grupo de Redes, Sistemas y Servicios de las capacidades del *IPmiser*.

Apéndice B

Publicaciones obtenidas de la Tesis Doctoral

- J. Aracil, D. Morató y M. Izal. Analysis of Internet Services in IP over ATM networks, *Proceedings of ICATM'99*. Junio 1999.
- J. Aracil, D. Morató y M. Izal. IPmiser: Herramienta de medida y análisis de redes de alta velocidad, *Proceedings of SPAST-I*, Pamplona, España, Diciembre 1999.
- J. Aracil, D. Morató y M. Izal. Analysis of Internet Services in IP over ATM Networks, *IEEE Communications Magazine*, Diciembre 1999.
- J. Aracil y D. Morató. Characterizing Internet Load as a Non-regular Multiplex of TCP Streams, *Proceedings of IEEE ICCCN 2000*, Las Vegas, Nevada, USA, Octubre 2000.
- J. Aracil, D. Morató y M.Izal. Analisys and stochastic characterization of TCP flows, *Annals of Telecommunications*, Noviembre-Diciembre 2000, n 55, 11-12, pags 559-566
- J. Aracil y D. Morató. A-priori flow bandwidth estimates for dynamic bandwidth allocation in ISP access links, *Proceedings of 14th ITC Specialists Seminar on Access Networks and Systems 2001*, Girona, España, Abril 2001.

Apéndice C

Curriculum

C.1 Situación profesional actual

Organismo: Universidad Pública de Navarra
Facultad, Escuela o Instituto: Escuela Superior de Ingenieros Industriales y de Telecomunicación
Depto./Secc./Unidad estr.: Automática y Computación
Dirección postal: Campus Arrosadía s/n C.P. 31006 Pamplona
Telefono: +34 948269854
Fax: +34 948168924
Correo electrónico: daniel.morato@unavarra.es
Especialización (Códigod UNESCO): 3305
Categoría profesional: Ayudante de Escuela Universitaria
Fecha de inicio: 4/10/1997
Situación administrativa: Contratado
Dedicación: A tiempo completo

C.2 Líneas de investigación

Modelos, análisis y caracterización de tráfico. Prestaciones de protocolos sobre redes de alta velocidad.

C.3 Formación académica

Ingeniero de Telecomunicación por la Universidad Pública de Navarra en Octubre de 1997.

C.4 Participación en Proyectos de I+D financiados en Convocatorias públicas

Título del contrato/proyecto: Migración de redes HFC a redes multiservicio
Tipo de contrato: Plan Nacional de I+D/FEDER
Empresa/Entidad Financiadora: Comisión Interministerial de Ciencia y Tecnología y Fondos FEDER Unión Europea
Duración desde: Mayo 1999 **hasta:** Mayo 2001
Investigador responsable: Javier Aracil
Número de investigadores participantes: 8
Presupuesto Total del Proyecto: 25.000.000 pts

C.5 Publicaciones o Documentos Científico-Técnicos

Autores (p.o. de firma): J. Aracil, D. Morató, M. Izal
Título: On the combined effect of self-similarity and flow control in quality of service for transactional Internet services.
Ref: Libro: Performance of Information and Communication Systems, Korner and Nilsson Eds, ISBN 0412837307
Clave: Capítulo de libro **Páginas, inicial:** 111 **final:** 123 **Fecha:** 1998
Editorial: Chapman&Hall

Autores (p.o. de firma): J. Aracil, D. Morató, M. Izal
Título: Analysis of Internet Services in IP over ATM Networks
Ref: Revista: IEEE Communications Magazine
Clave: Artículo **Volumen:** 37 **Páginas, inicial:** 92 **final:** 97 **Fecha:** Diciembre 1999

Autores (p.o. de firma): J. Aracil, D. Morató, M. Izal
Título: Analisys and stochastic characterization of TCP flows
Ref: Revista: Annals of Telecommunications
Clave: Artículo **Volumen:** 55 **Páginas, inicial:** 559 **final:** 566 **Fecha:** Diciembre 2000

Autores (p.o. de firma): J. Aracil, M. Izal, D. Morató
Título: Internet Traffic Shaping for IP over WDM Links with Source Output Buffering or Multiple Parallel Wavelengths
Ref: Revista: Optical Networks Magazine
Clave: Artículo **Volumen:** **Páginas, inicial:** 69 **final:** 77 **Fecha:** Enero/Febrero 2001

C.6 Participación en contratos de I+D de especial relevancia con Empresas y/o Administraciones

Título del contrato/proyecto: Proyecto Ipmiser, analizador de tráfico IP sobre ATM
Tipo de contrato: Contrato universidad-empresa
Empresa/Administración financiadora: Telefónica I+D
Entidades participantes: Telefónica I+D, Universidad Pública de Navarra
Duración, desde: Marzo 1997 **hasta:** Marzo 1998
Investigador responsable: Javier Aracil
Número de investigadores participantes: 8
Presupuesto Total del Proyecto: 16.000.000 pts

Título del contrato/proyecto: Realización de medidas ATM/IP del tráfico cursado en el Piloto ADSL
Tipo de contrato: Contrato universidad-empresa
Empresa/Administración financiadora: Telefónica I+D S.A.
Entidades participantes: Telefónica I+D, Universidad Pública de Navarra
Duración, desde: Marzo 1998 **hasta:** Agosto 1998
Investigador responsable: Javier Aracil
Número de investigadores participantes: 3
Presupuesto Total del Proyecto: 3.982.000 pts

Título del contrato/proyecto: Migración de la red HFC de Retena a red multiservicio
Tipo de contrato: Contrato universidad-empresa
Empresa/Entidad Financiadora: Redes de Telecomunicaciones de Navarra S.A.
Duración, desde: Mayo 1999 **hasta:** Mayo 2001
Investigador responsable: Javier Aracil, Jesús Villadangos, Univ. Pública de Navarra
Número de investigadores participantes: 8
Presupuesto Total del Proyecto: 8.200.000 pts

Título del contrato/proyecto: Análisis de la Red ADSL de Telefónica Data
Tipo de contrato: Contrato universidad-empresa
Empresa/Administración financiadora: Telefónica Data S.A.
Entidades participantes: Telefónica I+D, Universidad Pública de Navarra
Duración, desde: Julio 2000 **hasta:** Septiembre 2000
Investigador responsable: Javier Aracil
Número de investigadores participantes: 3
Presupuesto Total del Proyecto: 1.335.400 pts

Título del contrato/proyecto: Sistema de gestión del estado de máquinas expendedoras distribuidas geográficamente
Tipo de contrato: Contrato universidad-empresa
Empresa/Administración financiadora: GM Vending
Entidades participantes: GM Vending, Universidad Pública de Navarra
Duración, desde: hasta:
Investigador responsable: Jesús Villadangos
Número de investigadores participantes: 3
Presupuesto Total del Proyecto: 2.500.000 pts

C.7 Contribuciones a Congresos

Autores: J. Aracil, D. Morató, M. Izal
Título: Use of CBR for IP over ATM
Tipo de participación: Ponencia
Congreso: SPIE Voice, Video and Data Communications Conference
Publicaciń: Proceedings of the SPIE Voice, Video and Data Communications Conference
Lugar celebración: Dallas, TX, USA
Fecha: Noviembre 1997

Autores: J. Aracil, D. Morató, M. Izal
Título: Analysis of Internet Services in IP over ATM Networks
Tipo de participación: Artículo en Proceedings
Congreso: International Conference on ATM (ICATM'99)
Publicación: Proceedings of ICATM'99
Lugar celebración: Colmar, Francia
Fecha: Junio 1999

Autores: J. Aracil, D. Morató, M. Izal
Título: IPmiser: Herramienta de medida y análisis de redes de alta velocidad
Tipo de participación: Artículo en Proceedings
Congreso: I Seminario del Programa Nacional de Aplicaciones y Servicios Telemáticos (SPAST-I)
Publicación: Proceedings of SPAST-I
Lugar celebración: Pamplona, Navarra
Fecha: Diciembre 1999

Autores: J. Aracil, D. Morató
Título: Characterizing Internet Load as a Non-regular Multiplex of TCP Streams
Tipo de participación: Artículo en Proceedings

Congreso: IEEE International Conference on Computer Communications and Networks (ICCCN'00)
Publicación: Proceedings of ICCCN'00
Lugar celebración: Las Vegas, NV, USA
Fecha: Octubre 2000

Autores: J. Aracil, D. Morató
Título: A-priori flow bandwidth estimates for dynamic bandwidth allocation in ISP access links
Tipo de participación: Artículo en Proceedings
Congreso: International Teletraffic Congress (14th ITC)
Publicación: Proceedings of 14th ITC Specialists Seminar on Access Networks and Systems 2001
Lugar celebración: Girona, España
Fecha: Abril 2001

C.8 Otras Publicaciones

Editores: Pedro J. Lizcano, Javier Aracil Rico, Míkel Izal Azcárate, Eduardo Magaña Lizarrondo, Daniel Morató Osés y Jesús Villadangos Alonso
Título: Proceedings del I Seminario del Programa Nacional de Aplicaciones y Servicios Telemáticos (SPAST-I)
ISBN: 84-95075-28-8

Autores: Javier Aracil, Daniel Morató, Míkel Izal, Juan Ignacio Solana, Teresa Ariste y David Fillmore
Título: IPmiser, sistema de monitorización de enlaces ATM a 155Mbps
Tipo: Poster en el II Congreso Nacional de Ingeniería de Telecomunicación

---

Coautor de más de 9 informes técnicos en proyectos de I+D.

C.9 Experiencia en organización de actividades de I+D. Organización de congresos, seminarios, jornadas, etc., científicos-tecnológicos

Título: I Seminario del Programa Nacional de Aplicaciones y Servicios Telemáticos (SPAST-I)
Tipo de actividad: Congreso
Tipo de participación: Miembro del Comité de Organización

Ámbito: Nacional
Fecha: Diciembre 1999

C.10 Experiencia docente

Asignatura: Arquitectura de Redes, Sistemas de Servicios
Titulación/Curso: Ingeniero de Telecomunicación, 3º
Tipo de docencia: Prácticas
Curso académico: 1997/1998
Facultad: E.T.S. de Ingenieros Industriales y de Telecomunicación
Universidad: Universidad Pública de Navarra

Asignatura: Arquitectura de Computadores
Titulación/Curso: Ingeniero de Telecomunicación, 4º
Tipo de docencia: Prácticas
Curso académico: 1998/1999, 1999/2000 y 2000/2001
Facultad: E.T.S. de Ingenieros Industriales y de Telecomunicación
Universidad: Universidad Pública de Navarra

Asignatura: Redes, Sistemas y Servicios
Titulación/Curso: Ingeniero de Telecomunicación, 5º
Tipo de docencia: Prácticas
Curso académico: 1998/1999 y 1999/2000
Facultad: E.T.S. de Ingenieros Industriales y de Telecomunicación
Universidad: Universidad Pública de Navarra

Asignatura: Redes de Ordenadores
Titulación/Curso: Ingeniero de Telecomunicación, 5º
Tipo de docencia: Prácticas
Curso académico: 2000/2001
Facultad: E.T.S. de Ingenieros Industriales y de Telecomunicación
Universidad: Universidad Pública de Navarra

Asignatura: Laboratorio de Telemática
Titulación/Curso: Ingeniero de Telecomunicación, 4º
Tipo de docencia: Prácticas
Curso académico: 1997/1998
Facultad: E.T.S. de Ingenieros Industriales y de Telecomunicación
Universidad: Universidad Pública de Navarra

Asignatura: Bases de Información: Acceso a través de Sistemas Multimedia
Tipo de docencia: Prácticas
Curso académico: 1998/1999
Universidad: Universidad Pública de Navarra

C.11 Proyectos de fin de carrera dirigidos

Título: Sistema de control de estado de máquinas expendedoras distribuidas geográficamente
Proyectista: Fernando Rio
Titulación: Ingeniero de Telecomunicación

Título: Canal asimétrico TCP
Proyectista: Gorka Sánchez
Titulación: Ingeniero de Telecomunicación

Título: Generación y análisis de tráfico mediante hardware dedicado. Comparativa con hardware de propósito general
Proyectista: Leonardo Larios
Titulación: Ingeniero de Telecomunicación

C.12 Idiomas

Idioma	Habla	Lee	Escribe
Inglés	Bien	Correctamente	Bien
Francés	Regular	Regular	Regular

Ciclo Elemental del Primer Nivel de la Escuela Oficial de Idiomas en Inglés.

C.13 Otros méritos o aclaraciones que se desee hacer constar

- Mención especial en los estudios de Ingeniero de Telecomunicación por la Universidad Pública de Navarra en la convocatoria de premios nacionales de fin de carrera de educación universitaria 1996-97.

- Premio extraordinario de terminación de estudios en la titulación de Ingeniero de Telecomunicación en la Escuela Técnica Superior de Ingenieros Industriales y de Telecomunicación de la Universidad Pública de Navarra en estudios concluidos el año 1997.

- Becario de la Universidad Pública de Navarra en el curso académico 1996-97
- Quinto clasificado en la Olimpiada Química de 1992, Delegación de Aragón y Navarra. Representante y primer puesto por Navarra.
- Premio de enseñanzas medias 1991-92.
- Premio extraordinario de la Fundación de Premios Escolares Don Florencio Ansoleaga a mejor expediente de bachillerato, año 1991.

Bibliografía

[1] John Bellamy. *Digital Telephony*. Wiley Interscience, second edition, 1991.

[2] L. Kleinrock. *Queueing Systems*, volume 1. John Wiley and Sons, 1975.

[3] Mark K. Lottor. RFC 1296: Internet growth (1981-1991), January 1992.

[4] Vern Paxson. Growth trends in wide-area TCP connections. *IEEE Network*, pages 8–17, July/August 1994.

[5] Raj Jain and Shawn A. Routhier. Packet trains-measurements and a new model for computer network traffic. *IEEE Journal on Selected Areas in Communications*, 4(6):986–995, September 1986.

[6] R. Gusella. A measurement study of diskless workstation traffic on an Ethernet. *IEEE Transactions on Communications*, 38:1557–1568, September 1990.

[7] W. Leland and W. Wilson. High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection. In *Proceedings of INFOCOM'91*, 1991.

[8] H. Fowler and W. Leland. Local area network traffic characteristics, with implications for broadband network congestion management. *IEEE Journal on Selected Areas in Communications*, 9:1139–1149, September 1991.

[9] P. Danzig, S. Jamin, R. Cáceres, D. Mitzel, and D. Estrin. An empirical workload model for driving wide-area TCP/IP network simulations. *Internetworking: Res., Experience*, 3(1):1–26, March 1992.

[10] K.Meier-Hellstern, P.Wirth, Y-L. Yan, and D.Hoeflin. Traffic models for ISDN data users: Office automation application. In *Proceedings of the 13th ITC*, 1991.

[11] G. Babic, B. Vandalore, and R. Jain. Analysis and modeling of traffic in modern data communication networks. Technical Report TR OSU-CISRC-1/98-TR02, Ohio State University. Department of Computer and Information Science, February 1998.

[12] S.M. Klivansky, A. Mukherjee, and C. Song. On long-range dependence in NSFNET traffic. Technical Report GIT-CC-94-612, Georgia Institute of Technology, December 1994.

[13] Kevin Thompson, Gregory J. Miller, and Rick Wilder. Wide-Area Internet Traffic Patterns and Characteristics. *IEEE Network*, pages 10–23, November/December 1997.

[14] V. Paxson and S. Floyd. Wide Area Traffic: The Failure of Poisson Modeling. *IEEE/ACM Transactions on Networking*, 3(3):226–244, June 1995.

[15] Vern Paxson. Empirically Derived Analytic Models of Wide-Area TCP Connections. *IEEE/ACM Transactions on Networking*, 2(4):316–336, August 1994.

[16] Mark E. Crovella and Azer Bestavros. Self-Similarity in World Wide Web traffic: evidence and possible causes. *IEEE/ACM Transactions on Networking*, 5(6):835–846, December 1997.

[17] V.S. Frost and B. Melamed. Traffic modeling for telecommunications networks. *IEEE Communications Magazine*, pages 70–81, March 1994.

[18] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the Self-Similar Nature of Ethernet Traffic (Extended Version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, February 1994.

[19] Murad S. Taqqu, Walter Willinger, and Robert Sherman. Proof of a Fundamental Result in Self-Similar Traffic Modeling. *ACM SIGCOMM Compuer Communication Review*, 1997.

[20] J. Beran. *Statistics for Long-Memory Processes*. Number 61 in Monographs on Statistics and Applied Probability. Chapman & Hall, 1994.

[21] B.B. Mandelbrot and J.W. Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM Review*, 10:422–437, 1968.

[22] B. Tsybakov and N. D. Georganas. On self-similar traffic in ATM queues: definitions, overflow probability bound and cell delay distribution. *IEEE/ACM Transactions on Networking*, 5(3):397–409, June 1997.

[23] Ashok Erramilli, Onuttom Narayan, and Walter Willinger. Experimental Queueing Analysis with Long-Range Dependent Packet Traffic. *IEEE/ACM Transactions on Networking*, 4(2):209–223, April 1996.

[24] K. Park, G. Kim, and M. Crovella. On the relationship between file sizes, transport protocols, and self-similar network traffic. In *Proceedings of the 1996 Fourth International Conference on Network Protocols (ICNP'96)*, October 1996.

[25] Kihong Park, Gitae Kim, and Mark Crovella. On the effect of traffic self-similarity on network performance. In *Proceedings of the 1997 SPIE International Conference on Performance and Control of Network Systems*, 1997.

[26] P. Karlsson and A. Arvidsson. Modelling of traffic with high variability over long time scales with MMPPs. Technical report, Department of Telecommunications and Mathematics, University of Karlskrona/Wonneby.

[27] Walter Willinger, Murad S. Taqqu, Robert Sherman, and Daniel V. Wilson. Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level. *IEEE/ACM Transactions on Networking*, 5(1), February 1997.

[28] Robert J. Adler, Raisa E. Feldman, and Murad S. Taqqu, editors. *A Practical Guide to Heavy Tails. Statistical Techniques and Applications*. Birkhauser, 1998.

[29] J.M. Peha. Protocols can make traffic appear self-similar. In *Proceedings of the 1997 IEEE/ACM/SCS Communication Networks and Distributed Systems Modeling and Simulation Conference*, 1997.

[30] J.Beran, R.Sherman, M.S.Taqqu, and W.Willinger. Long-range dependence in Variable-Bit-Rate video traffic. *IEEE Transactions on Communications*, 43(2/3/4):1566–1579, February/March/April 1995.

[31] M.W. Garrett and W. Willinger. Analysis, modeling and generation of self-similar VBR video traffic. In *Proceedings of the 1994 ACM SIGCOMM*, September 1994.

[32] I. Norros. On the use of Fractional Brownian Motion in the theory of connectionless networks. *IEEE Journal on Selected Areas in Communications*, 13(6):953–962, August 1995.

[33] C. Huang, M.Devetsikiotis, I.Lambadaris, and A.Roger Kaye. Fast simulation for self-similar traffic in ATM networks. In *Proceedings of the 1995 IEEE International Conference on Communication*, June 1995.

[34] Vern Paxson. Fast, approximate synthesis of fractional gaussian noise for generating self-similar network traffic. *Computer Communication Review*, 27(5):5–18, October 1997.

[35] V. Paxson W. Willinger and M.S. Taqqu. *Self-similarity and heavy tails: structural modeling of network traffic*, chapter 1, pages 3–25. Birkhauser, 1998.

[36] Ashok Erramilli and R.P. Singh. An application of deterministic chaotic maps to model packet traffic. *Queueing Systems*, 20(1-2):171–206, 1995.

[37] A. Veres and M. Boda. The chaotic nature of TCP congestion control. In *Proceedings of the 2000 Infocom*, Abril 2000.

[38] J. Aracil, D. Morató, and M. Izal. IPmiser: Herramienta de medida y análisis de redes de alta velocidad. In *Actas del SPAST-I : Primer seminario del plan nacional de aplicaciones y servicios telemáticos*, pages 190–193, December 1999.

[39] T. Berners-Lee, R. Fielding, and H. Frystyk. RFC 1945: Hypertext transfer protocol – HTTP/1.0, May 1996.

[40] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee. RFC 2616: Hypertext transfer protocol – HTTP/1.1, June 1999.

[41] R. Cáceres, P.B. Danzig, S. Jamin, and D.J. Mitzel. Characteristics of wide-area TCP/IP conversations. In *Proceedings of SIGCOMM'91*, September 1991.

[42] W. R. Stevens. *TCP/IP Ilustrated, Volume 1: The Protocols*, volume 1. Addison-Wesley, 1994.

[43] Edith Cohen and Haim Kaplan. Prefetching the means for document transfer: a new approach for reducing web latency. In *Proceedings of the 2000 Infocom*, Abril 2000.

[44] S. Floyd. Connections with multiple congested gateways in packet switched networks, part 1: One-way traffic. *ACM Computer Communications Review*, 21(5), October 1991.

[45] J. Mahdavdhi, M. Mathis, and J.Semke. The macroscopic behavior of the TCP congestion avoidance algorithm. *Computer Communication Review*, 27(3):67–82, July 1997.

[46] T. V. Lakshman and Upamanyu Madhow. The Performance of TCP/IP for Networks with High Bandwidth-Delay Products and Random Loss. *IEEE/ACM Transactions on Networking*, 5(3):336–350, June 1997.

[47] K. Poduri and K. Nichols. RFC 2415: Simulation studies of increased initial TCP window size, September 1998.

[48] R. Cohen y S. Ramanathan. Tuning TCP for high-performance in hybrid fiber coaxial broadband access networks. *IEEE/ACM Transactions on Networking*, 6(1), February 1998.

[49] John Heidemann, Katia Obraczka, and Joe Touch. Modeling the Performance of HTTP Over Several Transport Protocolos. *IEEE/ACM Transactions on Networking*, (5):616–630, October 1997.

[50] A. Romanow and S. Floyd. Dynamics of TCP traffic over ATM networks. *IEEE Journal on Selected Areas in Communications*, 13(4):633–641, May 1995.

[51] Kevin Fall and Sally Floyd. Simulation-based Comparisons of Tahoe, Reno and SACK TCP. *Computer Communication Review*, 26(3):5–21, July 1996.

[52] Anurag Kumar. Comparative Performance Analysis of Versions of TCP in a Local Network with a Lossy Link. *IEEE/ACM Transactions on Networking*, 6(4), August 1998.

[53] N. Cardwell, S. Savage, and T. Anderson. Modeling the performance of short TCP connections. Available in http://www.cs.washington.edu/homes/cardwell/quals/quals-paper.ps, October 1998.

[54] J.C. Mogul. Observing TCP dynamics in real networks. Technical Report WRL 92/2, Abril 1992.

[55] S. Glassman. A caching relay for the world wide web. In *Proceedings of the First International Conference on the WWW*, 1993.

[56] M. Abrams, C.R. Standridge, G. Abdulla, S. Williams, and E.A. Fox. Cahing proxies: Limitations and potentials. In *Proceedings of the Fourth International Conference on the WWW*, 1996.

[57] E. Cruz. A calculus of network delay, part I: Network elements in isolation. *IEEE Transactions on Information Theory*, 37, 1991.

[58] G. Kesidis. *ATM Network Performance*. Kluwer Academic Publishers, MA, 1996.

[59] J. Reynolds and J. Postel. RFC 1700: Assigned numbers, October 1994.

[60] K. Nagami, H. Esaki, Y. Katsube, and O. Nakamura. Flow aggregated, traffic driven label mapping in label-switching networks. *IEEE Journal on Selected Areas in Communications*, 17(6):1170–1177, June 1999.

[61] M. Ilvesmaki, M. Luoma, and R. Kantola. Flow classification schemes in traffic-based multilayer IP switching - comparison between conventional and neural approach. *Computer Communications*, (21):1184–1194, May 1998.

[62] K. C. Claffy, H. Braun, and G. C. Polyzos. On parametrizable methodology for internet traffic flow profiling. *IEEE Journal on Selected Areas in Communications*, 13(8):1481–1494, October 1995.

[63] M. Ilvesmaki, K. Kilkki, and M. Luoma. Packets or ports - the decisions of IP switching. In *Proceedings of SPIE Broadband Networking Technologies*, November 1997.

[64] Mark E. Crovella and Azer Bestavros. Explaining World Wide Web Traffic Self-Similarity. Technical Report TR-95-015, Computer Science Department, Boston University, October 1995.

[65] J. Aracil, D. Morató, and M. Izal. Analysis of Internet services in IP over ATM networks. *IEEE Communications Magazine*, 37(12):92–97, December 1999.

[66] M. Nabe, M. Murata, and H. Miyahara. Analysis and modeling of World Wide Web traffic for capacity dimensioning for Internet access lines. In *SPIE Video, Voice and Data Communications Conference*, Dallas, TX, November 1997.

[67] C. Qiao and M. Yoo. Optical burst switching (OBS) - A new paradigm for an optical Internet. *Journal of High-Speed Networks*, 8(1), 1999.

[68] V. Paxson. *Measurement and enalysis of end-to-end Internet dynamics*. PhD thesis, University of California, Berkeley, 1997.

[69] IEEE Communications Magazine, special issue on MPLS, December 1999.

[70] Y. Rekhter, B. Davie, D. Katz, E. Rosen, and G. Swallow. "Tag switching architecture overview". RFC 2105, February 1998.

[71] John G. Proakis. *Digital Communications.* McGraw-Hill, 1995.

[72] D.Morató, J.Aracil, E.Maga na M.Izal, and L.A.Díez-Marca. Explaining the impact of optical burst switching on self-similarity. *En proceso de revisión para el IEEE Communications Letters*, 2001.

[73] Vern Paxson. Automated packet trace analysis of TCP implementations. In *Proceedings of SIGCOMM '97*, June 1997.

[74] J. Aracil, D. Morató, and M. Izal. Analysis of internet services in IP over ATM networks. In *Actas del ICATM'99*, pages 258–266, June 1999.