

CASTBA: Medidas de tráfico sobre la Red Académica Española de Banda Ancha

M. Alvarez-Campana¹, A. Azcorra³, J. Berrocal¹, J. Domingo²,
D. Larrabeiti³, X. Martínez², J. I. Moreno³, J. R. Pérez¹ y J. Solé Pareta²

¹ Departamento de Ingeniería de Sistemas Telemáticos
Universidad Politécnica de Madrid.
ETSI Telecomunicación, Ciudad Universitaria, 28040 Madrid
www.dit.upm.es

² Centro de Comunicaciones Avanzadas de Banda Ancha
Universitat Politècnica de Catalunya,
Jordi Girona 1-3, Campus Nord, Módulo D6, 08034 Barcelona
www.ccaba.upc.es

³ Area de Ingeniería Telemática
Universidad Carlos III de Madrid,
28911 Leganés (Madrid)
www.it.uc3m.es

Resumen

La variabilidad y la impredecibilidad son probablemente las dos características más comúnmente aceptadas a la hora de describir el comportamiento del tráfico de Internet. Este conocimiento incierto, que dificulta cualquier intento serio de obtener un modelo adecuado a la hora de dimensionar los recursos de la red, lo es aún más cuando nos enfrentamos al empleo de nuevas tecnologías de banda ancha. La conjunción de ambos fenómenos en la Red Académica Española RedIRIS, en donde recientemente se ha adoptado una infraestructura troncal ATM para el soporte de servicios IP, plantea la necesidad de profundizar en el conocimiento sobre las prestaciones de la nueva arquitectura de red y de las características del tráfico que cursa.

El presente artículo recoge algunos de los resultados más significativos de las medidas de tráfico efectuadas en RedIRIS durante 1997 en el marco del proyecto CASTBA del Plan Nacional de I+D. Así mismo, se describen las consideraciones metodológicas utilizadas para la obtención de las medidas y se señalan las principales conclusiones que se derivan de las mismas.

1 Introducción

La popularización de Internet ha traído consigo la alteración intrínseca del tráfico de la red, no sólo en lo que a volumen respecta, si no también en cuanto al tipo de aplicaciones utilizadas y, por consiguiente, a la naturaleza del tráfico que éstas generan. Así, hace tan sólo unos años los principales servicios eran el correo electrónico (e-mail), la transferencia de ficheros (ftp), las noticias (news) y la conexión a ordenadores remotos (telnet). Hoy en día, sin embargo, el tráfico WWW (World Wide Web) es el que predomina sobre el resto, a la vez que se aprecia un sensible crecimiento de otras aplicaciones, como las de audio y vídeo en tiempo real.

La evolución de Internet se ha visto acompañada también de importantes cambios en lo que a infraestructura de red se refiere. Actualmente, uno de los exponentes más significativos de este cambio lo constituye la progresiva incorporación de tecnología ATM (Modo de Transferencia Asíncrona) en los enlaces troncales de la Red. Uno de los ejemplos más cercanos de migración hacia ATM es el que ha tenido lugar recientemente en RedIRIS [1], la rama de Internet que sirve a los centros académicos y de investigación nacionales.

Este contexto de transformación tecnológica y de alteración en los patrones de tráfico en Internet, plantea dificultades extraordinarias a la hora de dimensionar y gestionar los recursos de la red. El hecho de que el tráfico de Internet responda a patrones complejos [2], no bien conocidos, impide disponer de modelos de tráfico fiables. Esta complejidad radica principalmente en la gran diversidad de servicios y perfiles de usuario, que al superponerse en la Red se manifiestan en la agregación de innumerables flujos de tráfico con características muy dispares. Este fenómeno se acentúa en las nuevas redes de transporte de alta velocidad, como las basadas en ATM, en cuyos enlaces troncales de gran capacidad confluyen todos los factores citados.

Hasta la fecha, una gran parte de los estudios acerca del comportamiento del tráfico en Internet en redes de banda ancha, han sido fundamentalmente de carácter teórico. Sin afán de menospreciar este tipo de estudios, si cabe al menos plantearse la validez práctica de los resultados partiendo del conocimiento incierto que se tiene sobre la naturaleza del tráfico IP. En este sentido, la disponibilidad de medidas detalladas sobre escenarios reales, de modo que puedan contrastarse resultados teóricos y prácticos, puede ser el elemento clave que permita profundizar en la caracterización del tráfico de Internet. Este es precisamente uno de los principales objetivos del trabajo que a continuación se expone y de otros similares que recientemente han sido publicados [3].

Dentro del Plan Nacional de I+D, el proyecto CASTBA (CALidad de los Servicios Telemáticos provistos por la red académica de Banda Ancha), desarrollado conjuntamente por el Departamento de Ingeniería de Sistemas Telemáticos de la Universidad Politécnica de Madrid [4] y el Centro de Comunicaciones Avanzadas de Banda Ancha de la Universidad Politécnica de Catalunya [5], ha perseguido como objetivo la obtención de medidas exhaustivas dirigidas a la evaluación de prestaciones y la caracterización del tráfico de RedIRIS. Como ya se ha indicado, a lo largo de 1997 dicha red ha experimentado una migración tecnológica hacia ATM. De la arquitectura IP clásica basada en conmutación de paquetes y enlaces punto a punto dedicados, se ha pasado a una infraestructura de transmisión SDH con nodos de conmutación ATM, sobre la que se ha superpuesto la red IP original.

La adopción de ATM como infraestructura troncal de RedIRIS ha supuesto una importante transición estructural y topológica de la red, ante la cual se plantea la necesidad de conocer con mayor detalle el comportamiento de la nueva arquitectura de red ante el tráfico que generan los usuarios de RedIRIS. Por esta razón, nuestro objetivo ha sido no sólo la obtención de medidas cuantitativas de tráfico y de prestaciones sobre los enlaces troncales de RedIRIS, sino también el análisis cualitativo de dicho tráfico a nivel IP. De esta manera, se ha conseguido disponer de datos concretos para evaluar el comportamiento de la nueva arquitectura ATM de la red. Además, se ha realizado un análisis de las características principales de los servicios IP más utilizados en la Internet, tomando como referencia las Comunidades Autónomas y algunas entidades pertenecientes a las mismas.

El resto del artículo se ha organizado como sigue. El apartado 2 contiene una breve descripción de la arquitectura actual de la red académica española (RedIRIS), así como de la infraestructura y la metodología utilizadas para las medidas. En el apartado 3 se presentan algunos de los resultados más representativos obtenidos a partir del análisis del tráfico cursado por RedIRIS. En el apartado 4 se comentan los aspectos más significativos de las gráficas incluidas en el apartado 3. Finalmente, en el apartado 5 se resumen las conclusiones del trabajo y se exponen las posibles líneas de continuación del mismo.

2 Infraestructura y metodología de medidas para el análisis del tráfico de RedIRIS

La red troncal que soporta los servicios de comunicaciones de RedIRIS tiene una topología en estrella alrededor de un nodo central situado en Madrid (figura 1). A este nodo central se interconectan un total de 17 nodos regionales, uno por cada Comunidad Autónoma, mediante circuitos asimétricos ATM 34/155 Mbps a través de la red pública GigaCom de Telefónica. Los usuarios de la red, centros de investigación y universidades, se conectan a los nodos regionales más próximos mediante líneas dedicadas cuya velocidad varía entre 64 Kbps y 2 Mbps.



Figura 1. Topología de RedIRIS

El tráfico inyectado a la red troncal es 100% IP. Con objeto de transportarlo a través de los enlaces ATM, los paquetes IP se encapsulan según la recomendación RFC1483 [6] utilizando la capa de adaptación AAL5 [7].

Con objeto de poder obtener medidas simultáneas en los 17 enlaces troncales de RedIRIS, la ubicación de los equipos de medida se fijó en el nodo central de la red. Concretamente, se situó un analizador de tráfico entre el conmutador ATM de acceso a GigaCom y el router IP central de RedIRIS. Con objeto de no interferir en el funcionamiento normal de la red, se utilizaron dos divisores pasivos de fibra óptica, uno para cada sentido de transmisión (entrada y salida de los nodos regionales), sobre el par de fibra que agrupa el tráfico del conjunto de nodos. El esquema de conexión de los equipos puede verse en la figura 2.

En CASTBA se han llevado a cabo tres tipos de medidas de tráfico sobre la troncal de RedIRIS. Un primer grupo de medidas ha estado dirigido a la obtención de medidas cuantitativas del tráfico y de prestaciones de la red. Con la metodología desarrollada para este tipo de medidas, y con los datos obtenidos en la realización de las mismas, se ha aportado el conocimiento necesario acerca del comportamiento de la red académica en cuanto a sus prestaciones y calidad de servicio (retardos y pérdidas de paquetes IP), así como acerca del adecuado dimensionamiento de los enlaces troncales (utilización media de los enlaces, origen y destino del tráfico, etc.). Este conjunto de medidas es lo que hemos denominado en nuestra metodología como *carga de tráfico y prestaciones*.

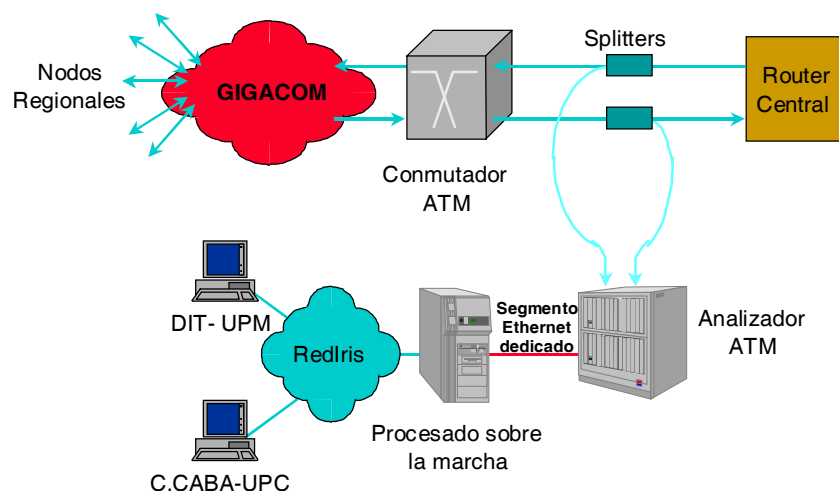


Figura 2. Localización de los puntos de medida

El segundo gran grupo de medidas de tráfico realizadas ha tenido como fin la caracterización del tráfico generado por las aplicaciones más utilizadas en Internet. La metodología desarrollada para este tipo de medidas se ha basado en la *captura y análisis del tráfico* de la red troncal de RedIRIS. El análisis se lleva a cabo mediante la decodificación de las cabeceras IP, TCP, UDP, ICMP, etc., de las tramas AAL5 capturadas. De esta manera, se analizan diferentes características del tráfico IP cursado por los enlaces troncales de RedIRIS: protocolos de nivel de transporte utilizados, porcentajes de utilización de los diferentes servicios o aplicaciones de red (HTTP, FTP, SMTP, TELNET, NEWS, etc.) en base a los puertos TCP/UDP utilizados, longitudes de los datagramas IP fuente, fragmentación de paquetes IP en la red, sobrecarga del tráfico IP sobre ATM, etc.

El tercer grupo de medidas ha tenido como objetivo la determinación de perfiles de utilización de servicios por los distintos grupos de usuarios finales. Ello ha requerido el análisis de tráfico por comunidades autónomas, entidades pertenecientes a las mismas (p.e. universidades), subentidades (p.e. departamentos, centros docentes, etc.) y grupos de usuarios finales (p.e. aulas de estudio y laboratorios de investigación).

3 Resultados obtenidos

En este apartado se incluyen y comentan una selección del conjunto de gráficas que resultaron del análisis del tráfico de RedIRIS. Se pretende que esta selección sea representativa de los distintos tipos de medidas realizadas, de acuerdo con lo expuesto en el apartado precedente.

El apartado 3.1 recoge gráficas referentes a medidas de carga de tráfico y de prestaciones (retardo introducido y la tasa de pérdida de paquetes) de los enlaces troncales de RedIRIS. El apartado 3.2 se centra en medidas relativas a la distribución de la longitud de los paquetes IP, tratando de profundizar en las características intrínsecas del tráfico cursado por la troncal. El análisis de la longitud de paquetes IP por servicios enlaza con el apartado 3.3. En dicho apartado se trata de cuantificar el uso de recursos en la red mediante la diferenciación de servicios y usuarios.

Los datos que a continuación se exponen corresponden a distintos períodos de medidas comprendidos entre marzo y noviembre de 1997, lo que garantiza la validez estadística de las muestras. Cabe señalar también que las medidas basadas en captura de tráfico se han realizado fundamentalmente sobre la troncal de Cataluña. Dicho troncal se escogió por ser uno de los más cargados tal como se demuestra en el siguiente apartado.

3.1 Medidas de carga de tráfico y prestaciones

Las gráficas de la figura 3 muestran el porcentaje medio de ocupación del conjunto de los enlaces troncales de RedIRIS a lo largo de un día laborable (gráfica 3.a) y de una semana (gráfica 3.b). Obsérvese que se han representado por separado los resultados correspondientes a los tráficos de entrada (a las Comunidades Autónomas) y de salida. Como referencia, conviene señalar que el porcentaje de utilización del 100% equivale a 68 Mbit/s y 40 Mbit/s para las curvas de entrada y salida, respectivamente.

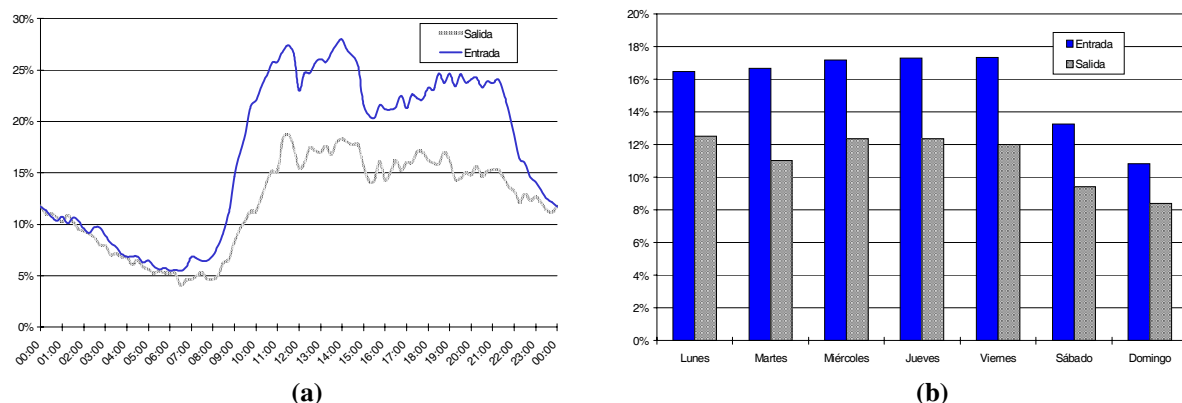


Figura 3. Utilización media de los enlaces troncales
a) evolución diaria b) evolución semanal

En cuanto a la carga de tráfico y porcentajes de utilización de los distintos enlaces regionales, las medidas efectuadas arrojan los resultados de la figuras 4.a y 4.b, respectivamente. Las valores de las gráficas corresponden a promedios obtenidos a partir de las medidas realizadas en días laborables. Cabe indicar que el enlace de la Rioja no se encontraba operativo durante el período de medidas, razón por la que no aparece tráfico en las columnas correspondientes de las gráficas.

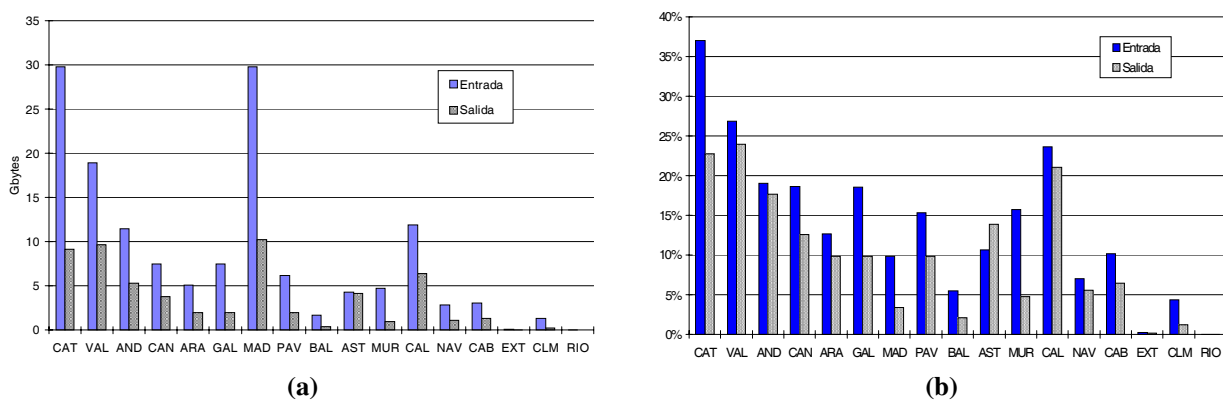
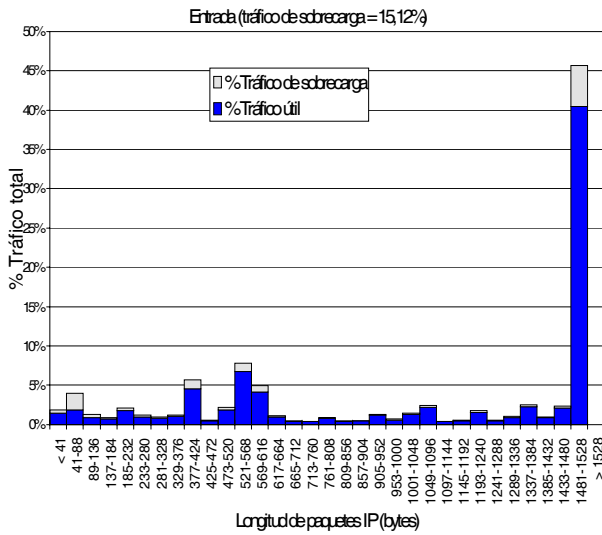
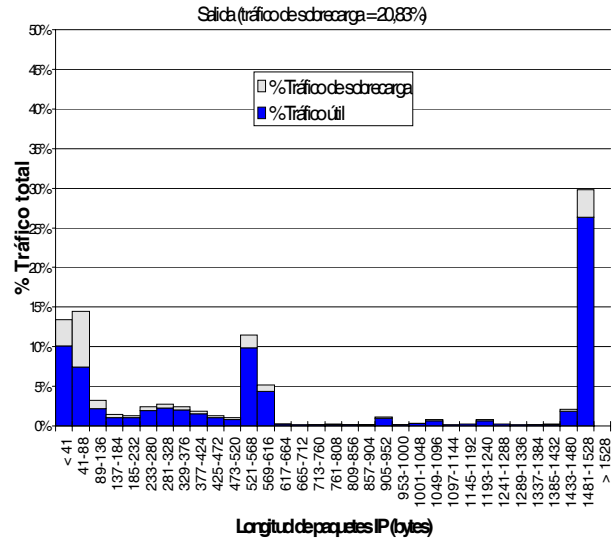


Figura 4. Carga de tráfico por enlace troncal
a) Gbytes diarios transmitidos b) utilización media diaria

En la figura 5 se han representado los porcentajes de tráfico útil (octetos IP) y de sobrecarga (ATM y AAL5) en función de la longitud de los paquetes en bytes. Cada columna del histograma corresponde al conjunto de longitudes de paquetes IP (bytes) que dan lugar a una misma longitud de trama AAL5. La sobrecarga media que resulta para cada sentido de transmisión es del 15,12% (en entrada) y del 20,83% (en salida), siendo del 17,32% en termino medio para el tráfico total (entrada + salida).



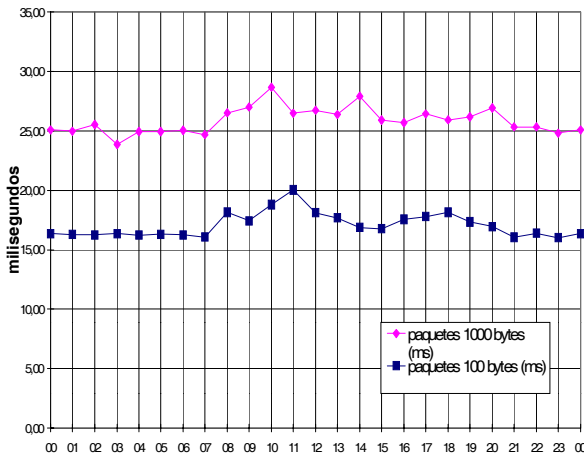
(a)



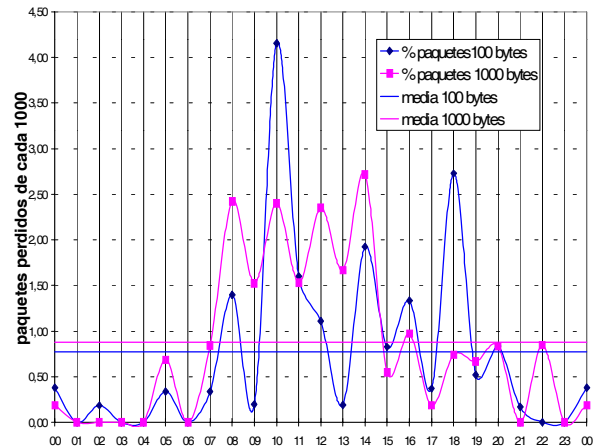
(b)

Figura 5. Sobrecarga media de IP sobre ATM a) entrada b) salida

En la figura 6 se presentan los resultados de las medidas de retardo y pérdida de paquetes a nivel IP. Los valores representados en la gráfica 6.a corresponden al retardo medio de tránsito de un paquete IP (en lazo de ida y vuelta) enviado entre dos máquinas cercanas a los nodos extremos del circuito ATM MAD-CAT. El retardo se expresa en milisegundos en función de la franja horaria. La figura 6.b muestra la tasa media de pérdida de paquetes IP (en lazo de ida y vuelta) del circuito ATM MAD-CAT y se expresa en paquetes perdidos por cada mil en función de la franja horaria.



(a)



(b)

Figura 6. a) Retardo medio de tránsito de paquetes IP
b) Tasa media de pérdidas de paquetes IP

3.2 Medidas de caracterización de tráfico IP

En la figura 7 se muestra la distribución estadística de la longitud de los paquetes IP cursados en el enlace troncal de Cataluña. Cada columna del histograma en la gráfica 7.a corresponde al conjunto de longitudes de paquetes IP que dan lugar a una misma longitud de trama AAL5. La figura 7.b muestra con mayor resolución la distribución de longitudes para paquetes de hasta 100 bytes. En ambas gráficas, los resultados se expresan en tanto por ciento respecto al número total de paquetes capturados. Cabe señalar que los datos presentados no incluyen paquetes IP fragmentados, si bien se ha comprobado que estos paquetes suponen un porcentaje muy bajo (aproximadamente un 0,05%).

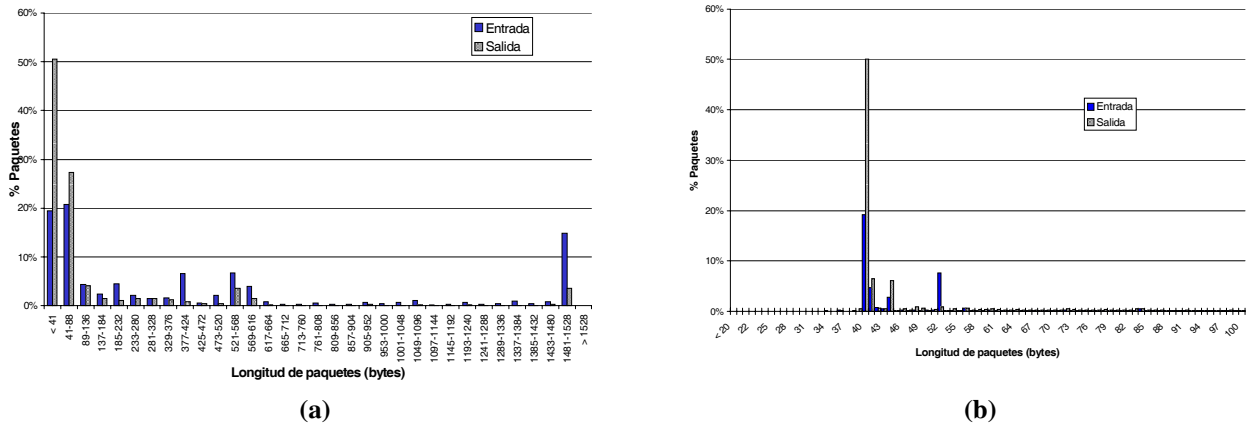


Figura 7. Distribución de longitudes de paquetes IP
a) por tramas AAL5 de igual tamaño b) ampliación para paquetes IP de hasta 100 bytes

Al analizar la longitud de los paquetes IP por servicio se obtienen los resultados que se muestran en la figura 8.a. El gráfico representa los valores medios de la longitud de paquetes IP para cada uno de los servicios estudiados. Como complemento a dicha gráfica, en la figura 8.b se muestra el porcentaje de paquetes capturado para cada servicio.

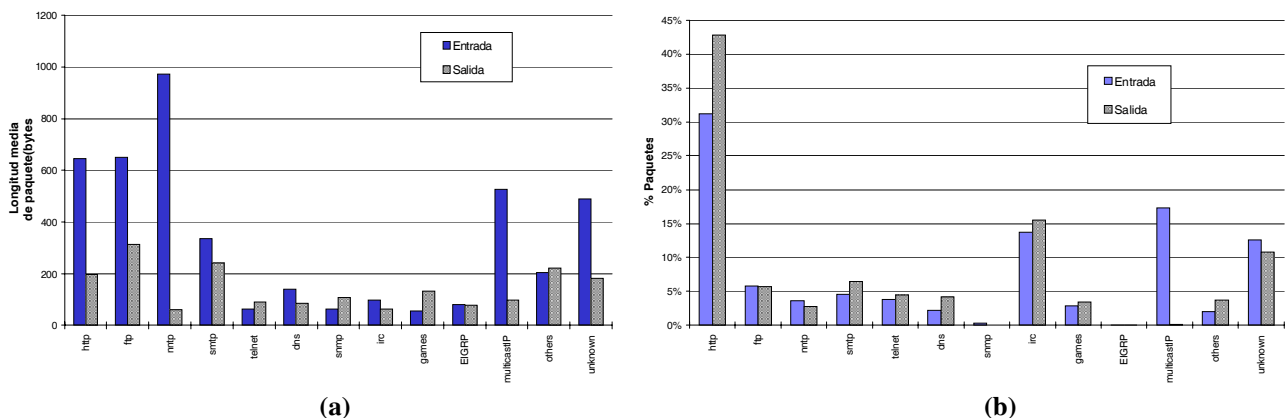
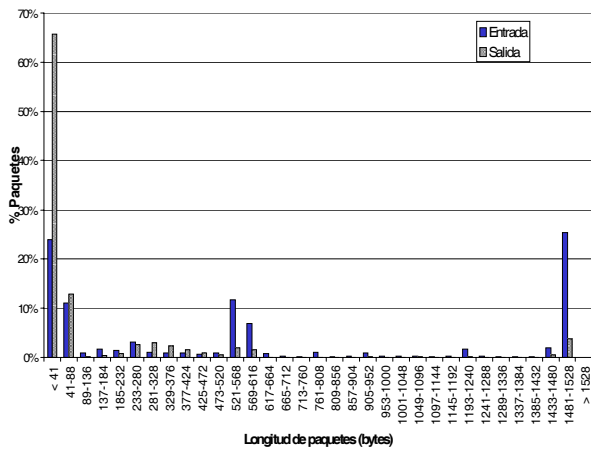
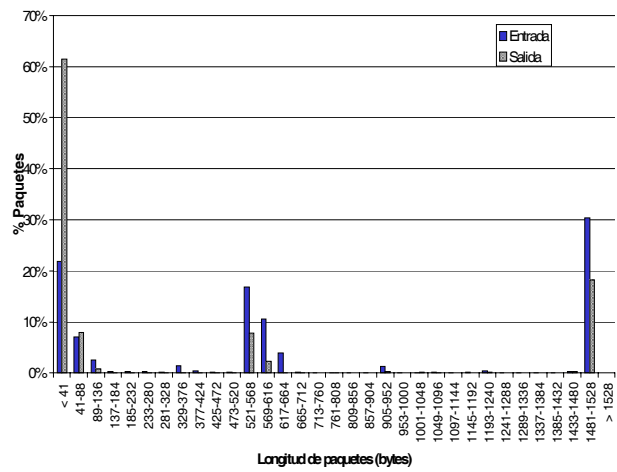


Figura 8. a) Longitud media de paquetes IP por servicio
b) Porcentaje de paquetes IP capturados por servicio

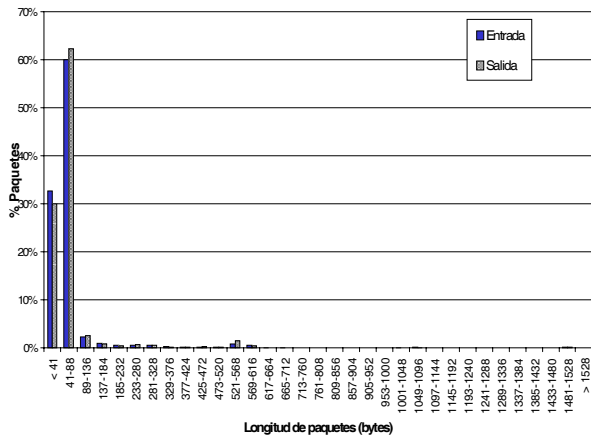
En el apartado siguiente se incidirá con más detalle en medidas de tráfico por servicios. No obstante, a la vista de la figura 8, resulta adecuado analizar la distribución de longitudes de los distintos servicios. Como se observa en la figura 9, las distribuciones varían sustancialmente según el servicio que se considere.



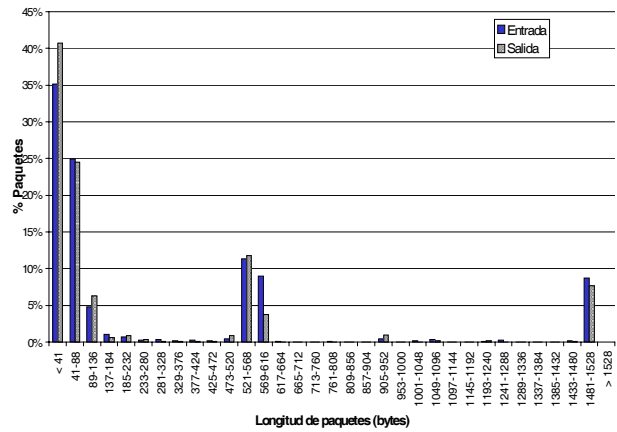
(a)



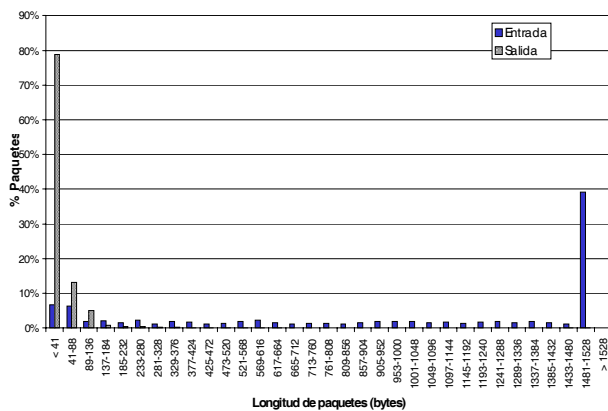
(b)



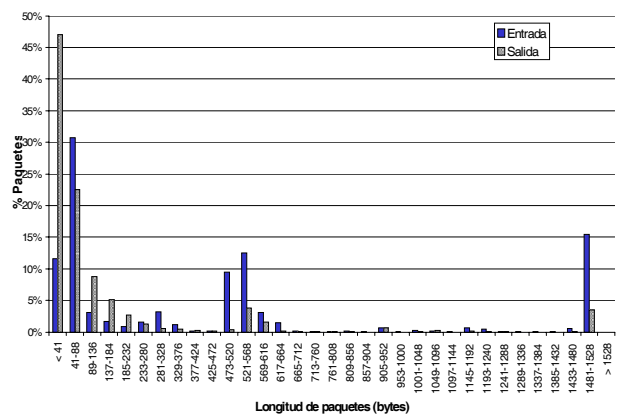
(c)



(d)



(e)



(f)

Figura 9. Distribución de longitudes de paquetes IP por servicio
a) HTTP b) FTP c) TELNET d) SMTP e) NNTP f) Tráfico desconocido

3.3 Medidas de caracterización de tráfico por servicios y perfiles de usuario

En la figura 10 se muestra la evolución diaria (gráficas 10.a y 10.b) y semanal (gráficas 10.c y 10.d) del volumen total de tráfico por servicio medido en el troncal de RedIRIS de Cataluña. Las gráficas de la izquierda corresponden a tráfico de entrada al nodo de Cataluña, mientras que las de la derecha se refieren a tráfico saliente.

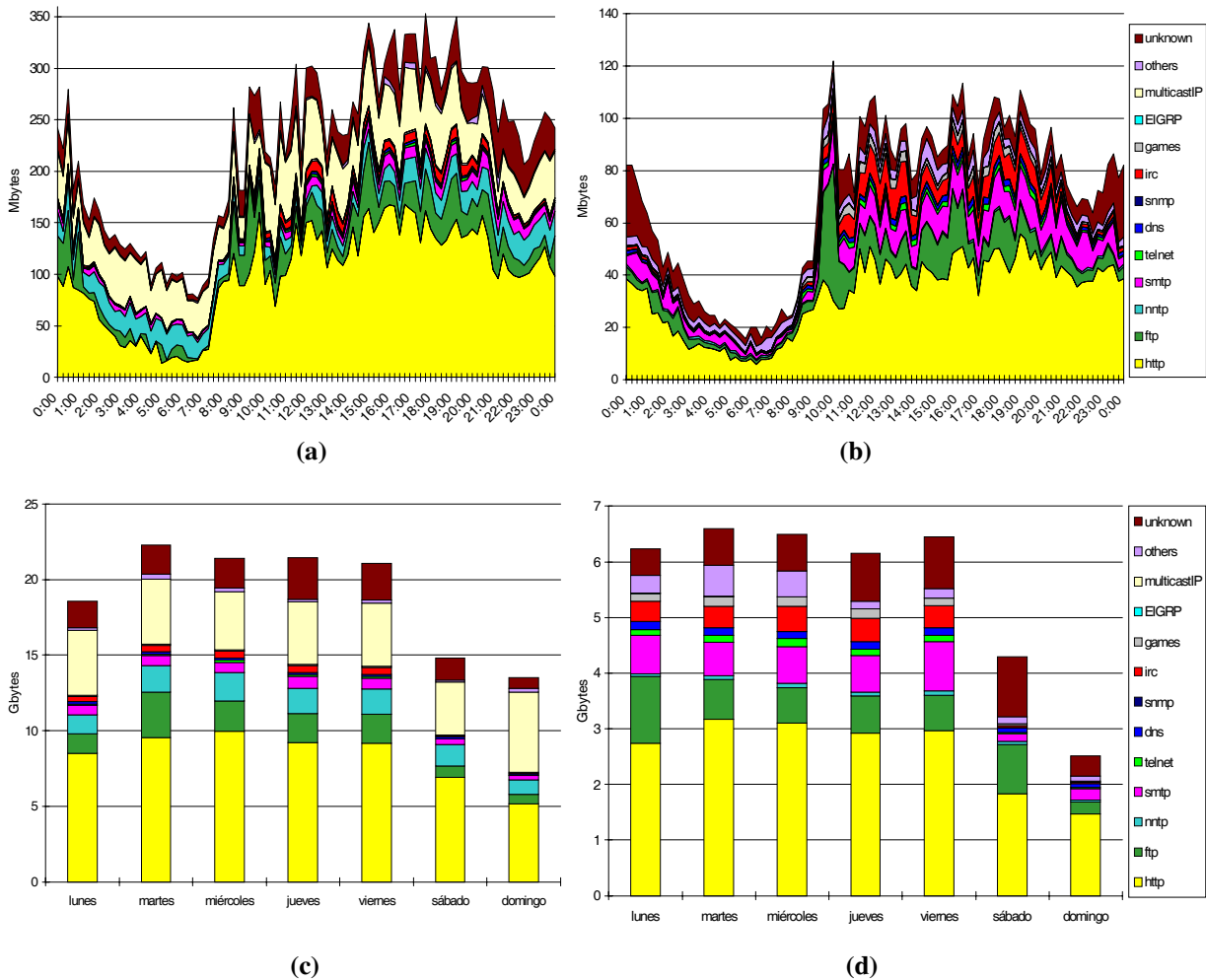


Figura 10. Distribución de tráfico por servicios

- a) evolución horaria del tráfico de entrada b) evolución horaria del tráfico de salida
 c) evolución semanal del tráfico de entrada d) evolución semanal del tráfico de salida

Como complemento de las gráficas anteriores, en la figura 11 se muestra la distribución del volumen tráfico por entidades conectadas al troncal de Cataluña. Por motivos de claridad, sólo se han diferenciado algunas de las entidades más significativas, habiéndose representado el volumen de tráfico restante de manera agregada, como si de una única entidad se tratase. La figura 11.a corresponde a la distribución del tráfico de entrada a las instituciones seleccionadas, mientras que la figura 11.b corresponde a la distribución del tráfico de salida. Las figuras 11.c y 11.d contiene la misma información que las figuras 11.a y 11.b pero referidas a la franja diaria.

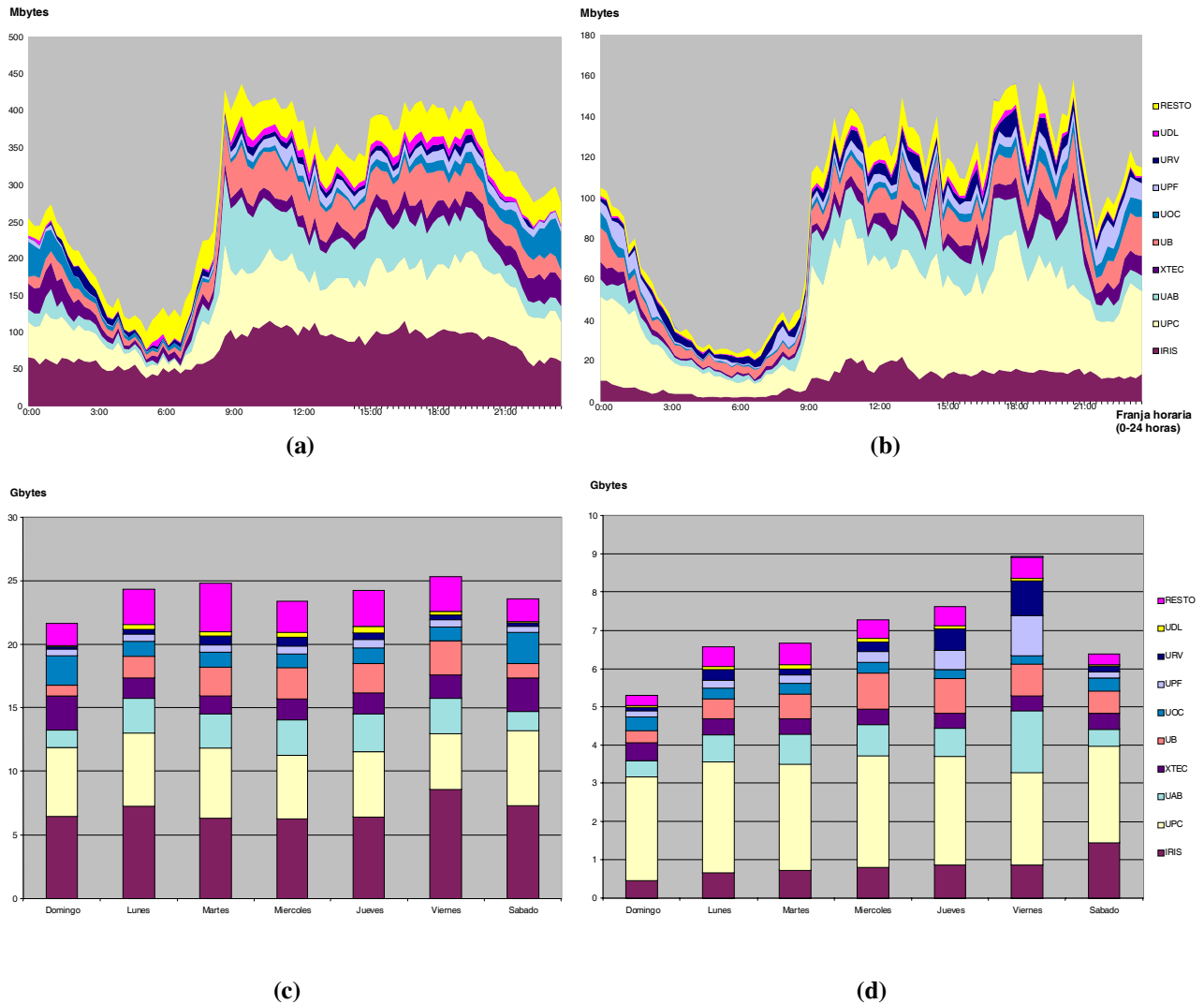


Figura 11. Distribución de tráfico por entidades
a) evolución horaria del tráfico de entrada b) evolución horaria del tráfico de salida
c) evolución semanal del tráfico de entrada d) evolución semanal del tráfico de salida

Como muestra de distribución de la utilización de un determinado servicio por las distintas entidades del troncal en estudio se ha escogido el servicio WWW que, como se puede ver observando la distribución del tráfico por servicios que presentan las gráficas de la figura 10, es el que presenta un mayor volumen de tráfico. Así, la figura 12 muestra el volumen de tráfico http (servicio WWW) de distintas entidades (las mismas que en el caso anterior, figura 11) del troncal de Cataluña. La figura 12.a corresponde a la distribución obtenida en el tráfico de entrada a las distintas entidades, mientras que la figura 12.b corresponde a la distribución en el tráfico de salida.

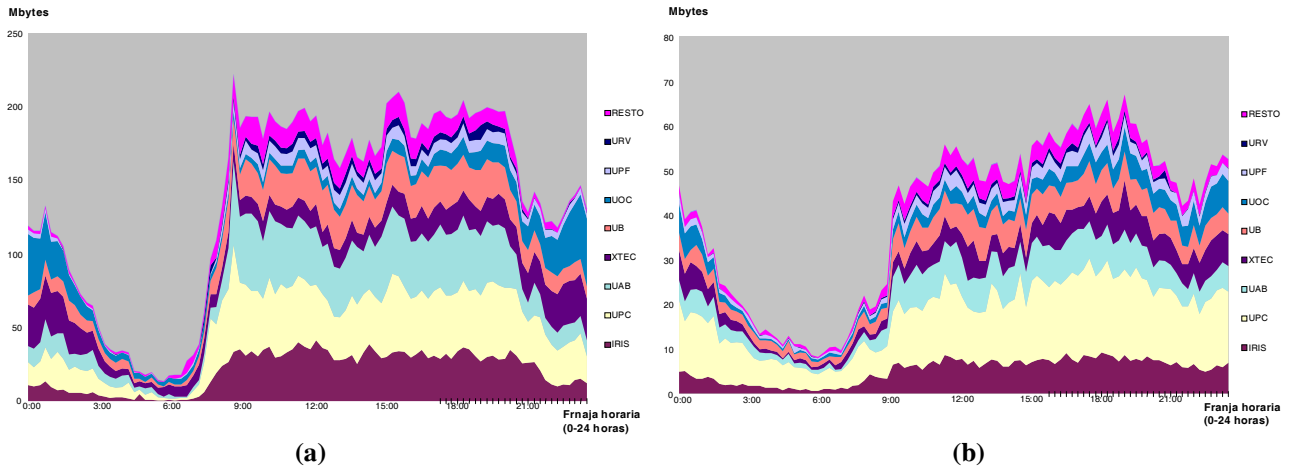


Figura 12. Distribución de tráfico HTTP por entidades
a) tráfico de entrada b) tráfico de salida

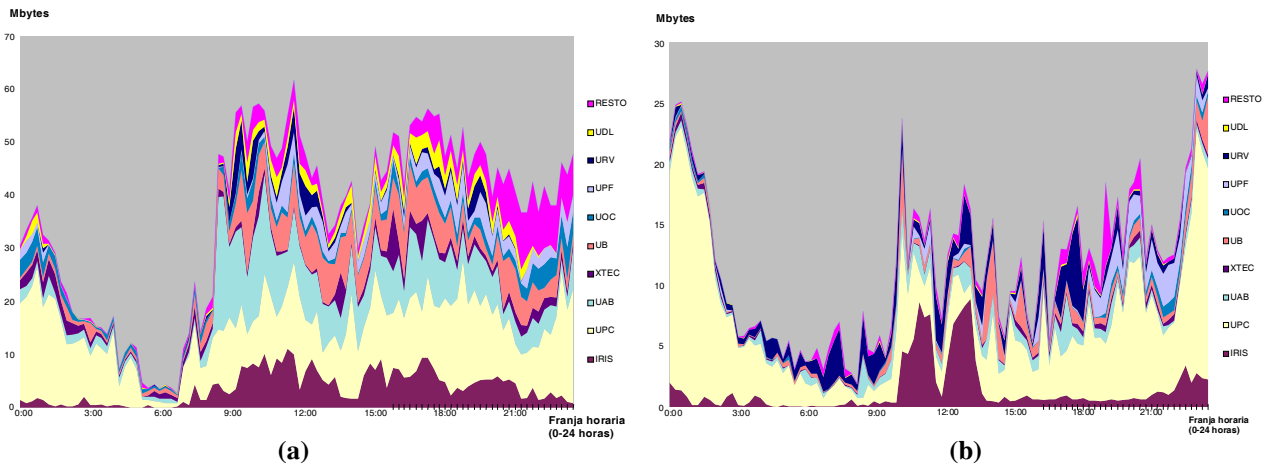


Figura 13. Distribución de tráfico desconocido por entidades
a) tráfico de entrada b) tráfico de salida

También es interesante hacer mención la distribución del volumen de tráfico correspondiente a servicios que hemos llamado desconocido y en realidad corresponde a servicios no catalogados, es decir tráfico en el que se detectan puertos TCP no estándares. En este sentido, la figura 13 muestra el volumen de tráfico desconocido de las distintas entidades del troncal de Cataluña seleccionadas en las gráficas anteriores (figuras 11 y 12). La figura 13.a corresponde a la distribución obtenida en el tráfico de entrada, mientras que la figura 13.b corresponde a la distribución en el tráfico de salida.

Por otro lado, como muestra de distribución de tráfico por servicios las distintas entidades del troncal de Catalunya se han escogido la UPC (Universidad Politècnica de Catalunya) y la UOC (Universitat Oberta de Catalunya), una (la UPC) por ser de las que, como se puede ver observando la distribución del tráfico por entidades que presentan las gráficas de la figura 11, genera un mayor volumen de tráfico, y la otra (la UOC) por presentar una distribución según franja horaria singular debido a las características propias de la entidad. Así, la figura 14 muestra la distribución del volumen de tráfico por servicios de la UPC y la figura 15 la distribución del volumen de tráfico por servicios de la UOC. Como siempre, las figuras 14.a y 15.a corresponden a la distribución obtenida en el tráfico de entrada de las dos entidades, mientras que las figuras 14.b y 15.b corresponden a la distribución en el tráfico de salida.

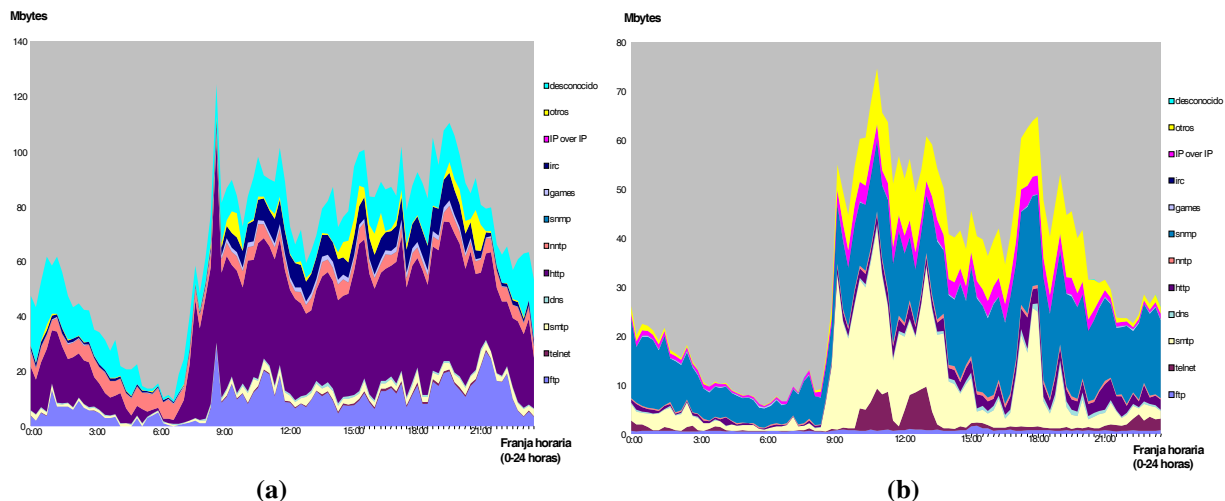


Figura 14. Distribución del tráfico de la UPC por servicios
 a) tráfico de entrada b) tráfico de salida

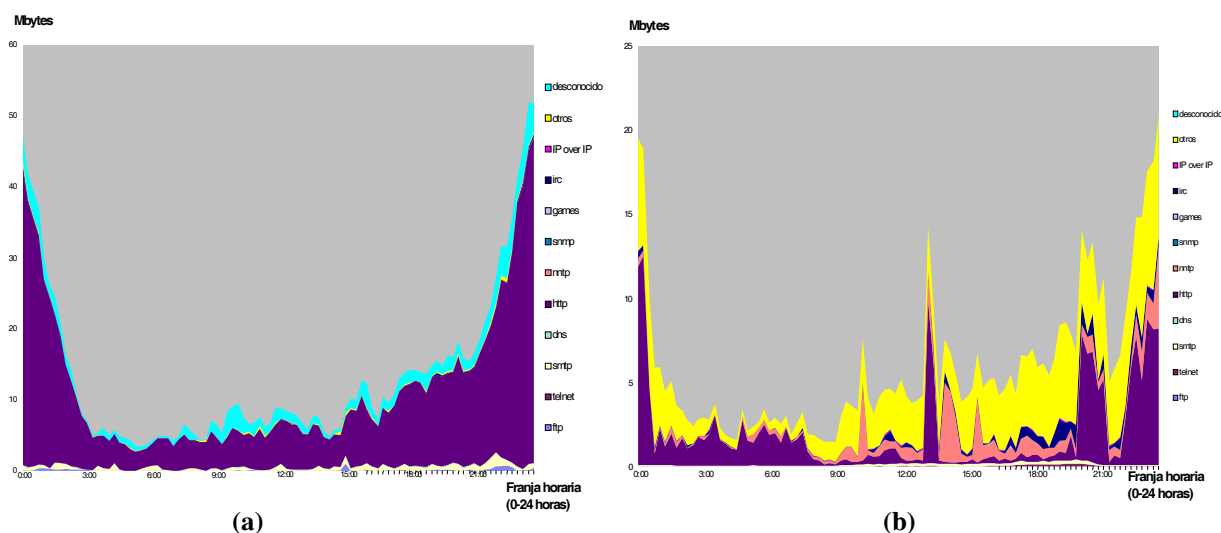


Figura 15. Distribución del tráfico de la UOC por servicios

Finalmente, el análisis del tráfico en la red se ha centrado en colectivos de usuarios más concretos. Como ya se ha visto, la UPC presenta un volumen de tráfico considerable y por ello se eligió para analizar grupos de usuarios. En esta selección de gráficas se ha incluido la correspondiente a la distribución de los volúmenes de tráfico generados por distintas unidades de la UPC. Para este estudio se han seleccionado grupos de usuarios (departamentos, centros docentes, laboratorios de investigación y laboratorios docentes) que se creía que generaban un mayor volumen de tráfico. Adviértase, sin embargo, que no ha resultado ser así, observando las gráficas de la figura 16 se puede ver como el volumen de tráfico correspondiente a los grupos seleccionados en su conjunto no llega a la mitad del volumen total de tráfico medido.

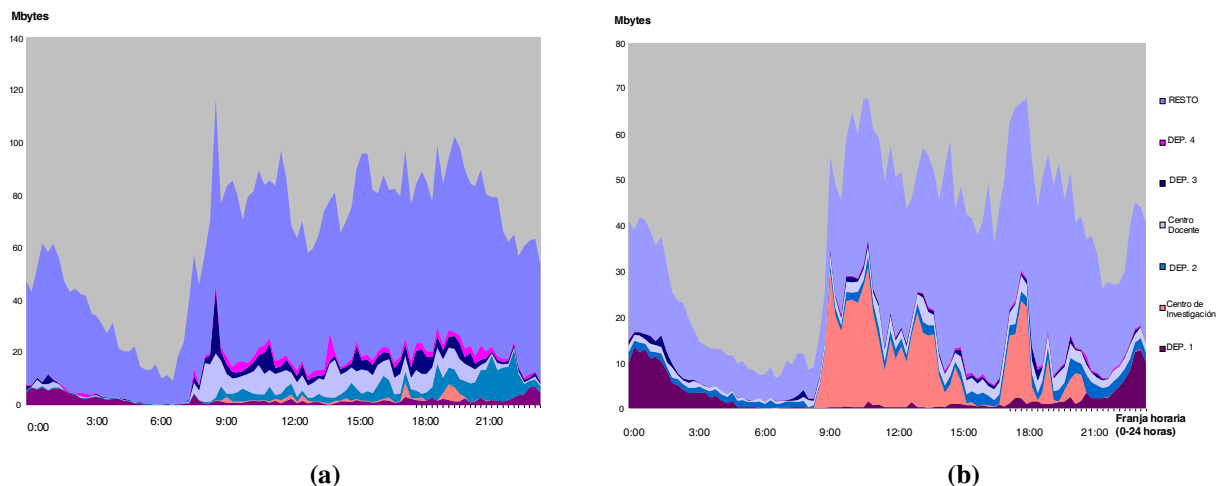


Figura 16. Distribución de tráfico por grupos de usuarios de la UPC
a) tráfico de entrada b) tráfico de salida

4 Discusión de los resultados

Como conclusiones que directamente se derivan de las medidas realizadas en el proyecto y, en particular, de las presentadas en este artículo destacaremos las siguientes:

Carga de tráfico

- Las medidas de carga horaria y diaria de tráfico de la red troncal de RedIris (figura 3) revelan que la ocupación máxima del conjunto de enlaces ATM es aproximadamente del 25% en entrada y del 17% en salida a las horas de mayor tráfico, es decir, entre las 9 h y las 21 h. Cabe señalar además que no se han observado picos de tráfico de corta duración que sobrepasen ampliamente esta ocupación en ninguno de los enlaces.
- El tráfico agregado transferido a diario sobre los enlaces troncales de RedIris equivale a unas tasas medias de 14,5 Mbit/s en el sentido de entrada hacia los nodos regionales y de 5,8 Mbit/s en el sentido de salida. Dichas tasas medias equivalen a la transferencia diaria de 146 Gbytes y 59 Gbytes, respectivamente.
- En la gráfica 4 se aprecia que las comunidades autónomas con mayor volumen de tráfico son Madrid, Cataluña, Valencia, Castilla-León y Andalucía, por este orden. También se observa que, en general, el tráfico de entrada a los nodos regionales es aproximadamente el doble que el de salida. Ello sugiere que los nodos regionales son sumideros de información proveniente del exterior (USA, Europa, Ibernet).

Sobrecarga de IP sobre ATM

- En cuanto a las medidas de sobrecarga, cabe destacar la ineficiencia introducida por las capas ATM y AAL5. Aproximadamente el 17% del tráfico cursado es de sobrecarga. Esta sobrecarga se debe a las cabeceras de célula y a los campos de trailer y relleno añadidos por AAL5. En la gráfica 5 se observa que el desperdicio de ancho de banda ocasionado por el total de paquetes cortos es comparable al ocasionado por el total de paquetes largos. Esto se explica porque los paquetes cortos son los más numerosos (~60%) y los que soportan mayor sobrecarga (~50%), mientras que los paquetes largos, aunque en menor proporción (~10%) y con menor sobrecarga (~10%), suponen aproximadamente el 40% del tráfico en octetos.
- La diferencia entre la sobrecarga de entrada (15,12%) y de salida (20,83%) es debida a la asimetría de ambos tráficos en sus distribuciones de longitudes de paquetes. El de entrada tiene una mayor proporción de paquetes IP grandes y por tanto, una eficiencia mayor. Por el contrario, en el de

salida abundan paquetes pequeños (típicamente de control TCP y peticiones de transferencia de ficheros http), dando lugar a una eficiencia mucho menor.

Retardo y pérdidas

- Las medidas de retardo de paquetes IP (figura 6.a) muestran valores muy estables y bajos (17 ms para paquetes de 100 bytes y 26ms para los de 1Kbyte). Las pérdidas de paquetes (figura 7.b) también resultan muy bajas, con una media menor al 1 por mil y, en todo caso, siempre inferiores al 4 por mil.

Longitud de paquetes IP

- A la vista de la gráfica de distribución de longitudes de paquetes IP (figura 7.a), se observan tres grupos de tamaños predominantes : paquetes cortos (1 a 3 células), paquetes largos (32 células) y de tamaño medio (9, 12 y 13 células).
- La proporción de paquetes cortos es muy alta. Más del 75% de los paquetes de salida y más del 40% de los de entrada tienen una longitud menor de 100 octetos. Al analizar el tamaño de estos paquetes en octetos (figura 7.b), se aprecia una cierta dispersión debido a la variedad de servicios y de sus PDUs de control. Cabe mencionar que la longitud característica de los mensajes de control de TCP es 40 octetos, que el servicio telnet es el principal causante del valor 41, el http del 44 y que el tamaño 52 se debe a tráfico desconocido (unknown).
- Los paquetes de tamaño medio se deben a los mecanismos habitualmente utilizados por las implementaciones de TCP con objeto de minimizar el grado de segmentación. Así, muchas implementaciones de BSD requieren que el MSS (max tcp segment size) sea múltiplo de 512. Precisamente este tamaño, sumado a los 20 octetos de cabecera IP y a los 20 octetos de cabecera TCP, da lugar a 552 octetos de datagrama IP, correspondiente a la franja 521-568. Por otro lado, si en el establecimiento de la conexión la entidad TCP no especifica MSS se utiliza el valor por defecto de 536 octetos, es decir 576 octetos IP, que se corresponde con el máximo relativo en el intervalo 569-616 y con la MTU (maximum transmission unit) mínima recomendada para IP. Este comportamiento de TCP justifica el bajo grado de fragmentación encontrado en la muestra. Por último, señalar la existencia de una fracción significativa de paquetes de entrada en la franja 373-424, debida a tráfico multicast.
- Los paquetes largos (intervalo 1481-1528) se corresponden con los 1500 octetos de la MTU de un interfaz ethernet que es el tipo de red local más extendido. El MSS correspondiente es 1460. Algunos sistemas muy extendidos como SunOS, Solaris, AIX, anuncian por defecto una MSS de 1460 para direcciones IP en el dominio considerado local, en caso contrario utilizan 536. Otro aspecto a destacar es el hecho de que no aparecen prácticamente paquetes de longitud superior a 1500 octetos, como las correspondientes a una red ATM o a una red local FDDI.
- Las distribuciones de entrada y de salida son diferentes. El tráfico de entrada tiene una mayor proporción de paquetes grandes debido a que las transferencias masivas de datos se concentran en el sentido de entrada, utilizándose con frecuencia el máximo tamaño de paquete disponible (1500 octetos, como se ha explicado anteriormente). En otras palabras, los nodos regionales de RedIris son netamente sumideros de información y la mayor parte de tráfico de salida que generan son paquetes cortos de peticiones a servidores externos de información.

Distribución de tráfico por servicios

- En cuanto a la distribución del tráfico por servicios (en octetos) para el enlace de Cataluña (figura 10), se observa que el tráfico predominante es de HTTP (web) con un 45% del tráfico tanto en entrada como en salida. A continuación se encuentra el tráfico multicast con un 18% de tráfico de entrada. Siguiendo en importancia son los tráficos de FTP y UNKNOWN, ambos en torno al 10% del tráfico de entrada y de salida. Otros servicios a destacar, son los de SMTP (e-mail) con casi el 12% del tráfico de salida, el de NEWS con el 7% en entrada y el de IRC (chats) con el 6% del tráfico de salida del nodo de Cataluña.

- Cabe aclarar que, a pesar del elevado porcentaje de tráfico HTTP en el enlace de Cataluña, no se aprecia un volumen destacable de tráfico de DNS, como podría esperarse. Ello se debe a que la resolución de URLs se lleva a cabo en servidores de DNS locales.
- Al analizar la longitud media de los paquetes IP en función del servicio (figura 8.a) destacan las de los servicios de transferencia masiva de datos: nntp, ftp y http, por este orden. Nótese que, aunque estos servicios poseen características comunes (implican volúmenes de tráfico masivo y emplean el mismo protocolo de transporte), los tamaños medios de paquete que utilizan difieren apreciablemente. El motivo ha de buscarse en las configuraciones particulares utilizadas por los servidores.

Perfiles de tráfico de usuarios

- Las gráficas presentadas también ponen de manifiesto la consabida asimetría entre el tráfico de cliente a servidor y el de servidor a cliente, ello se puede observar contrastando los volúmenes de tráfico de entrada y salida de una misma entidad y/o unidad estructural, una vez sabido que es eminentemente consumidor o productor de información. Véase por ejemplo el caso de un centro proveedor de servicios como el volumen de tráfico que genera (figura 16.b, centro de investigación) es mucho más elevado del que recibe (figura 16.a).
- En la mayoría de las gráficas presentadas en este artículo pueden verse plasmada la gran dispersión estadística típica de Internet (más cuanto más nos acercamos al usuario final) en lo que se refiere a patrones horarios, uso de aplicaciones y longitud de los paquetes IP. Este aspecto es de capital importancia para no perder de vista que el conjunto de medidas realizadas no se pueda considerar concluyentes sino tan sólo una fotografía del período de tiempo en el que se realizaron las capturas de tráfico sobre las que se han sacado las estadísticas.

5 Conclusiones y trabajos futuros

En el marco del proyecto CASTBA, los grupos de trabajo implicados han obtenido un conjunto significativo de medidas de tráfico de la red académica española RedIRIS. Con ello se ha logrado profundizar en el conocimiento de las prestaciones de la nueva arquitectura de red IP/ATM y de las características del tráfico que actualmente cursa. En cuanto al primero de los aspectos, los resultados de las medidas ponen de manifiesto las excelentes prestaciones de la red, que se reflejan en valores de retardo y pérdida de paquetes muy bajos.

Otros resultados de interés son los referentes a la adecuación de la arquitectura de protocolos utilizada. En las medidas realizadas ha podido constatarse que la sobrecarga introducida por ATM y el protocolo AAL5 no son despreciables, representando aproximadamente un 17% del tráfico cursado en la troncal de RedIRIS. Este hecho se debe, en gran medida, a la elevada proporción de paquetes cortos, generalmente paquetes de control, que son los que mayor sobrecarga soportan. Si tenemos en cuenta que la aplicación más utilizada, el WWW, requiere el establecimiento de múltiples conexiones TCP por página, fácilmente se deduce que el diseño del protocolo HTTP no es el más indicado sobre una red IP/ATM.

La conclusión general más importante que se desprende a la vista de los resultados obtenidos durante los distintos períodos de medida del proyecto, es la variabilidad que exhibe el tráfico de RedIRIS. Esta consideración cobra mayor importancia, si cabe, en un entorno como el de RedIRIS, en el que se supone un colectivo de usuarios homogéneo. Probablemente, la principal fuente de inestabilidad del tráfico radique en que no existe una política de control del uso de recursos en la red. De este modo, cualquier usuario (o máquina) puede, en un momento dado, originar flujos de tráfico impredecibles que alteren de manera apreciable el comportamiento global de la red.

Desde el punto de vista práctico, esperamos que los resultados del proyecto CASTBA sean de utilidad a la hora de plantearse un posible redimensionamiento de la red ante las nuevas demandas que puedan surgir en el futuro. En el campo teórico, creemos que el análisis de las distribuciones de longitudes de paquete por servicio puede servir de punto de partida para la especificación de modelos de tráfico y la propuesta de mecanismos de control de tráfico en los puntos de acceso a la troncal.

Para finalizar, señalar que las labores de CASTBA tienen su continuación en un nuevo proyecto cuyo principal objetivo es el desarrollo de una herramienta de monitorización y análisis de tráfico de bajo coste. Se espera que la aplicación del sistema a RedIRIS permita en todo momento conocer aspectos cuantitativos y cualitativos del tráfico cursado en la red y, de esta manera, contribuir a mantener los niveles de calidad de servicio que actualmente tiene nuestra red.

Agradecimientos

El proyecto CASTBA ha sido subvencionado por la Comisión Interministerial de Ciencia y Tecnología (CICYT) a través de las Acciones Especiales TEL96-2508-E y TEL96-2509-E.

Además los autores agradecen la inestimable colaboración de RedIRIS, Telefónica I+D, C⁴ (Centre de Computació i Comunicacions de Catalunya) y UOC (Universitat Oberta de Catalunya) que han participado como EPOs (Entes Promotores y Observadores) de este proyecto

Referencias

- [1] RedIRIS, “Infraestructura de la Red Sideral (Servicio de Interconexión de Redes de Area Local)”, <http://www.rediris.es>.
- [2] W. Leland, M. Taqqu, W. Willinger, D. Wilson, “On the self-similar nature of Ethernet traffic”, IEEE/ACM Trans. on Networking, Febrero 1994.
- [3] K. Thompson, G. Miller, R Wilder, “Wide-Area Internet Traffic Patterns and Characteristics”, IEEE Network Magazine, Noviembre 1997.
- [4] DIT-UPM, “CASTBA-M. Supervisión de Calidad de Servicios Telemáticos sobre la Red Académica de Banda Ancha”, Programa Nacional de I+D, Informe Final, Noviembre 1997.
- [5] CCABA-UPC, “CASTBA-C. Supervisión de Calidad de Servicios Telemáticos sobre la Red Académica de Banda Ancha”, Programa Nacional de I+D, Informe Final, Noviembre 1997.
- [6] J. Heinanen, “Multiprotocol Encapsulation over ATM Adaptation Layer 5”, RFC 1483, Julio 1993.
- [7] ITU-T, “Recommendation I.363.5 - B-ISDN ATM Adaptation Layer specification: Type 5 AAL”, Agosto 1996.