

**MECANISMO DE PRECIOS PARA LA PRESTACIÓN DEL SERVICIO DE  
INTERNET BASADO EN LA TEORÍA DE ANCHO DE BANDA EFECTIVO**

**EDISON GIOVANNI MARTÍNEZ CASTRO**

**UNIVERSIDAD DE LOS ANDES  
FACULTAD DE INGENIERÍA  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL  
BOGOTÁ, D.C.  
2003**

MECANISMO DE PRECIOS PARA LA PRESTACIÓN DEL SERVICIO DE  
INTERNET BASADO EN LA TEORÍA DE ANCHO DE BANDA EFECTIVO

EDISON GIOVANNI MARTÍNEZ CASTRO

Trabajo de Tesis para optar al Título de  
Magíster en Ingeniería Industrial

Asesor

FERNANDO BELTRÁN GARCÍA

Ingeniero Eléctrico

UNIVERSIDAD DE LOS ANDES  
FACULTAD DE INGENIERÍA  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL  
BOGOTÁ, D.C.

2003

## TABLA DE CONTENIDO

TABLA DE CONTENIDO .....	2
LISTA DE FIGURAS .....	4
LISTA DE TABLAS .....	5
INTRODUCCIÓN .....	6
1    LOS MODELOS DE FIJACIÓN DE PRECIOS SUGERIDOS EN EL SERVICIO DE INTERNET .....	10
1.1 El esquema de tarifa plana .....	11
1.2 Distribución del tráfico por prioridades .....	13
1.3 Esquema de distribución justa de los recursos .....	14
1.4 Esquema de precios por prioridades basado en los precios Aumann-Shapley y el valor Shapley .....	15
1.5 El esquema de Capacidad Esperada .....	16
1.6 Conclusiones .....	17
2    LA TEORÍA DE ANCHO DE BANDA EFECTIVO .....	18
2.1 Las redes de banda ancha .....	18
2.2 El concepto de ancho de banda efectivo .....	19
2.2.1 Interpretaciones del concepto de ancho de banda efectivo .....	21
2.2.2 La teoría de grandes desviaciones .....	24
2.3 El Ancho de banda efectivo y la Asíntota de muchas fuentes .....	25
2.3.1 Significado de los parámetros $s, t$ .....	29
2.3.2 La asíntota de muchas fuentes .....	32
2.3.3 Utilidad y extensión de la asíntota de muchas fuentes .....	33
2.4 Implicaciones y Conclusiones .....	35
3    APLICACIÓN PRÁCTICA DE UN MODELO DE PRECIOS USANDO EL MÉTODO DE ANCHO DE BANDA EFECTIVO .....	37
3.1 Descripción general de la aplicación práctica .....	38

3.2	Implementación práctica del esquema .....	41
3.2.1	Las trazas de tráfico .....	43
3.2.2	Consideraciones respecto a las trazas de tráfico .....	48
3.2.3	El algoritmo de optimización .....	51
3.3	Resultados y conclusiones .....	52
4	MODELO DE COSTOS Y FIJACIÓN DE PRECIOS .....	55
4.1	El modelo de costos .....	56
4.2	Posibles esquemas de tarifas .....	59
	CONCLUSIONES Y RECOMENDACIONES .....	61
	REFERENCIAS BIBLIOGRÁFICAS .....	64
	LISTA DE ANEXOS .....	67

## LISTA DE FIGURAS

Figura 1. Representación gráfica de la tarifa cobrada al usuario como función de la tasa media del usuario .....	23
Figura 2. Representación de la distribución o repartición del tráfico de todas las fuentes de tipo $j$ a lo largo de la capacidad de la conexión $C^*$ .....	29
Figura 3. Presentación de la página de Diveo Uniandes, con las funciones correspondientes .....	39
Figura 4. Ubicación física del equipo de medición en la RDUA .....	39
Figura 5. Visualización del calendario para seleccionar el día del análisis y la solicitud de la red IP para realizar la consulta .....	44

### Anexo 1:

Figura 1. Representación de un modelo de buffer finito .....	69
--	----

## LISTA DE TABLAS

Tabla 1. Datos de tráfico como se muestran en la aplicación .....	45
Tabla 2. Tráfico de entrada y salida por hora del día, en porcentaje del total...	46
Tabla 3. Tráfico transferido para una hora específica de un día en particular, por aplicación .....	47
Tabla 4. Datos correspondientes al tráfico total de Entrada a la RDUA, para la subred 62 el día Junio 15 de 2003 .....	48

## INTRODUCCIÓN

El concepto de Internet Pricing<sup>1</sup>, que consiste en la fijación de precios en la prestación del servicio de Internet, ha cobrado gran importancia en el área de las comunicaciones. Esto se debe a que en los últimos años la red del Internet se ha vuelto cada vez más grande y compleja y ha permitido que cada vez más servicios de diferentes características sean prestados a través de ésta ([2] Bailey, 1999). Lo anterior ha incentivado a los operadores de la red de Internet <sup>2</sup> a crear un mecanismo a nivel económico que les permita optimizar la utilización de sus recursos de red disponibles, ofreciendo la mejor tarifa, con la mejor calidad del servicio y bajos costos de operación.

Bajo las condiciones actuales del Internet, diversos autores especializados en el tema han propuesto recientemente métodos y planteado herramientas y soluciones en la búsqueda de un mecanismo que determine con exactitud el precio que deben pagar los usuarios por el servicio que reciben, como respuesta a los ineficientes mecanismos convencionales, como la tarifa plana. Muchos de estos mecanismos han centrado su atención en lograr que el usuario tenga la capacidad de manejar la utilización de los recursos de la red acorde con sus necesidades, a través de un acuerdo o contrato de servicio entre el operador de la red y el usuario. De este modo, se busca poder transmitir la mayor cantidad de información en el menor tiempo posible, evitando la pérdida de datos y optimizando así el uso de la red.

---

<sup>1</sup> Término en inglés que podría ser traducido al español como “preciaje del Internet”

<sup>2</sup> También llamados Proveedores de Acceso a Internet (PAI), o Internet Service Providers (ISP) por su significado en inglés.

Un mecanismo ideal de precios debería lograr determinar de forma eficiente y más exacta posible el costo que un usuario de Internet debe asumir por utilizar los recursos de red en una red de banda ancha<sup>3</sup>, permitiendo a su vez que esos recursos se distribuyan de manera justa entre todos los usuarios en beneficio común. En la actualidad este tema aún se encuentra en estudio y no se ha llegado a un sistema determinado de tarificación para redes de banda ancha. Sin embargo, se ha definido una teoría en particular para la medición de los recursos de Internet sobre una red de banda ancha, basada en la distribución estocástica del tráfico proveniente de cada una de muchas fuentes sobre esta red: la teoría de ancho de banda efectivo ([14] Kelly, 1996).

De acuerdo con los estudios previamente hechos sobre el tema y con lo que será expuesto en este documento, esta teoría ha demostrado ser una aproximación no solamente válida sino coherente y precisa – aunque compleja – en la estimación de la porción de red utilizada por cada usuario, en un escenario en el que muchos usuarios originan tráfico de diferentes características que confluyen en una conexión de banda ancha. Sin embargo, no son muchos los análisis que se han hecho en cuanto a la validación experimental del modelo en una red de datos real, que puedan generar resultados y conclusiones importantes que permitan comprobar la efectividad del mecanismo. Sería interesante aplicarlo, por ejemplo, en una red de la cual uno mismo ha sido usuario, con el fin de poder determinar el nivel de calidad en el servicio ofrecido, comprobar de alguna forma si los recursos están siendo utilizados de manera óptima y estimar de este modo el desempeño general - o parcial - de la red.

En el presente trabajo se pretende enunciar un mecanismo de precios usando la teoría del ancho de banda efectivo y validar el modelo creado con datos reales, aplicándolo sobre el canal de tráfico por dependencia de la Facultad de Ingeniería

---

<sup>3</sup> Las redes de banda ancha permiten integrar diversos servicios a través de la misma red, sin importar el tipo de tráfico que cada servicio represente.



de la Universidad de los Andes. Este procedimiento tiene como fin determinar el nivel de utilización de cada una de las terminales de cada dependencia a cierta hora del día (usuarios de la facultad), cuyo tráfico se distribuye estadísticamente sobre el canal de banda ancha disponible entre la universidad y su proveedor de acceso a la red.

El presente documento consta de cuatro capítulos: el primer capítulo expone bs modelos de fijación de precios propuestos por los autores que han trabajado el tema de Internet Pricing, como respuesta a las limitaciones que tiene el modelo simplista de tarifa plana. Cada modelo implica diferentes opciones de servicio en las cuales se pueden llegar a ofrecer tarifas de acuerdo con las necesidades y presupuesto del usuario.

El segundo capítulo explica el concepto de ancho de banda efectivo, comenzando por entender la operación de las redes de banda ancha, sobre las cuales este tema es aplicable. Además, se muestra toda la teoría pertinente, como lo es la teoría de grandes desviaciones y en particular la ecuación de asíntota de muchas fuentes, que es el punto de partida para la simulación práctica de un esquema real, y se muestra conceptualmente la importancia de ciertos parámetros de red.

El tercer capítulo muestra una aplicación práctica utilizando datos de la red de la Universidad de los Andes basado en la teoría expuesta del concepto de ancho de banda efectivo, evaluando el nivel de factibilidad de ésta al momento de usarla en una red de datos real.

El cuarto capítulo presenta un modelo de costos y precios, el cual se tratará de enunciar de modo que sea aplicable en la práctica una vez se ha logrado establecer el parámetro de medición de los recursos de red, como lo fue el Ancho de banda efectivo obtenido en el capítulo 3. Igualmente, se mencionan los lineamientos de un acuerdo de servicio o contrato firmado entre proveedor y usuario y las ventajas de

realizar el proceso de la generación del contrato – una vez se ha prestado el servicio – de manera conjunta entre ambas partes.

Finalmente, se presentan las conclusiones y recomendaciones donde se explica qué resultó de este trabajo, y se muestran las posibilidades de continuar el proyecto con un trabajo futuro.

## **1 LOS MODELOS DE FIJACIÓN DE PRECIOS SUGERIDOS EN EL SERVICIO DE INTERNET**

El principal objetivo de un mecanismo eficiente de fijación de precios en el servicio de Internet es lograr distribuir de manera óptima los recursos de red entre todos los usuarios, satisfaciendo sus necesidades, y a su vez controlando y manejando la congestión. La congestión se entiende como aquel fenómeno en el que el Ancho de banda de la red es usado en su totalidad en un instante dado, pero la demanda por la transmisión (movimiento de información) se mantiene o tiende a aumentar.

Una forma de medir la eficiencia de cierto mecanismo de precios consiste en determinar su capacidad para ofrecer nuevas alternativas de servicio en presencia de congestión, y el principal reto de los ISP es manejar un esquema de precios que resuelva o disminuya el problema de la congestión, el cual produce demoras y degradación en el desempeño general de la red lo que en la práctica se traduce en mayores costos incrementales para el operador y por lo tanto tarifas más altas para el usuario ([10] Delgado, 2002).

En presencia de congestión en la red, los usuarios preferirán escoger la opción de comprar más o menos capacidad de la red de acuerdo con sus necesidades. Sin embargo, históricamente Internet no se ha caracterizado por permitirle al usuario escoger entre uno u otro servicio. Solamente ha implementado una clase de servicio, y ha recurrido a medios técnicos en lugar de económicos para distribuir los recursos de red adecuadamente cuando la red está cargada y congestionada.

Los clientes comenzarán a tener expectativas más sofisticadas, y con el tiempo estarán dispuestos a pagar tasas diferenciadas por la prestación de servicios diferenciados. Hoy en día, muchos usuarios se quejan de la lentitud en la transferencia de datos de gran tamaño, y aunque están dispuestos a pagar por un mejor servicio no pueden hacerlo, porque Internet no posee un mecanismo eficiente que permita mejorar dicho servicio ([6] Clark, 1995).

Un mecanismo de precios eficiente también busca lograr que los usuarios adquieran una cultura y una actitud de utilización de la red que les permita conocer su comportamiento en el uso de los recursos. De esta forma se busca que de alguna forma puedan predecir el efecto del tratamiento que le dan a su tráfico sobre sus respectivos cobros.

A continuación se presentan algunos de los esquemas sugeridos hasta hoy y que buscan con mayor precisión cumplir el objetivo de este trabajo.

## **1.1 EL ESQUEMA DE TARIFA PLANA**

En la actualidad, el escenario de mercado del Internet está dominado por una política simple de fijación de precios en la que la tarifa asignada se basa en una capacidad de red - o de ancho de Banda - que se le provee a un usuario, por la cual se le cobra un tarifa fija o plana. Esta tarifa ha sido previamente acordada entre usuario y proveedor por medio de un contrato o SLA<sup>4</sup> ([2] Bailey, 1999). A nivel técnico, Internet opera como una red de “mejor esfuerzo” en la que los paquetes de datos generados por cada aplicación de cada conexión de la red se transportan a través de ésta dependiendo de su disponibilidad.

---

<sup>4</sup> Service Level Agreement o Acuerdo de Nivel de Servicio por sus siglas en inglés.

El modelo anterior, el cual ha venido aplicándose por muchos años, explica el rápido crecimiento que ha tenido la red, por cuanto promueve la utilización masiva entre los usuarios y permite externalidades de red positivas ([3] Bailey, Nagel, 1999). Además, su aplicación representa ventajas en cuanto a practicidad: el operador se ahorra los costos por *billing* (cobro) y por la eventual implementación y operación de un método para medir los recursos de red.

Sin embargo, este modelo tiene serias limitaciones tanto a nivel técnico como económico. A nivel técnico, el modelo del “mejor esfuerzo” no permite la transmisión de servicios diferenciados porque todo el tráfico se maneja de la misma manera, y por lo tanto todas las aplicaciones tendrán la misma respuesta de parte de la red a pesar de que cada tipo de tráfico tiene un nivel de incidencia diferente sobre ésta ([2] Bailey, 1999). Teniendo en cuenta el desarrollo de nuevas y complejas aplicaciones sobre la red y la necesidad de los usuarios de acceder a más aplicaciones con un mejor servicio por el que están dispuestos a pagar, la aplicación de este modelo origina problemas de congestión en la red que afectan el servicio a los usuarios finales, por el cual el proveedor responde. Actualmente el mecanismo más importante para el control de la congestión a nivel técnico es la técnica TCP<sup>5</sup> Slow Start, pero se considera obsoleta e ineficiente. Además se ha comprobado que usuarios expertos de la red han logrado burlar sus efectos con tal de poder acceder a más servicios ([1] Ahmad, 1999).

Desde el punto de vista económico, el modelo de fijación de precios por capacidad de red fija no permite la utilización eficiente de los recursos de red por cuanto no depende del tráfico enviado. Es decir, dentro del ancho de Banda acordado, no existe control alguno del comportamiento del usuario (nivel de utilización de la red) por parte del proveedor, y el usuario podría usar el ancho de banda disponible al tope de su capacidad o no usarla y aún así tendría que pagar lo mismo al PAI ([2] Bailey, 1999). El modelo TCP no permite explicar el cargo que deben pagar los

---

<sup>5</sup> Transfer Control Protocol, por sus siglas en inglés

usuarios que transmiten mayor cantidad de información por el uso de la red con respecto a los que transmiten menos ([4] Beltrán, García, 2000).

Por las razones expuestas anteriormente, el esquema de tarifa plana ya no es el modelo de precios más conveniente tanto para usuarios como para los proveedores de las redes actuales.

## **1.2 DISTRIBUCIÓN DEL TRÁFICO POR PRIORIDADES**

El esquema propuesto por Gupta ([12], 1998) para la distribución de los recursos de red entre los usuarios consiste en crear clases de servicio con diferentes prioridades para servir a usuarios con diferentes necesidades. Si por ejemplo dos paquetes de prioridad distinta llegan a un nodo al mismo tiempo, el paquete de mayor prioridad siempre se atenderá primero.

Lo anterior puede ser un esquema útil de desarrollo para la discriminación explícita de los servicios de red, excepto que la prioridad no posee medios para satisfacer las demandas para las distintas clases de servicio. Es decir, los paquetes de prioridad alta (los cuales han sido marcados apropiadamente de acuerdo a una correspondiente clase de servicio o aplicación) pueden ser transmitidos más rápidamente en presencia de congestión. Sin embargo, cuando los recursos de red han sido ocupados en el envío de los paquetes de alta prioridad, los paquetes de prioridad baja podrían ser eliminados eventualmente del final de la cola de los paquetes de tráfico que esperan ser enviados. Estos paquetes no llegarían al usuario final y la demanda no sería satisfecha.

Como consecuencia, este mecanismo no le permite al usuario expresar directamente un comportamiento deseado de red, pues éste dependerá de la prioridad del tráfico que los demás usuarios impongan sobre sus paquetes. Por ejemplo, un usuario cuyo comportamiento de tráfico consista en el envío de

paquetes de baja prioridad no siempre podrá ser satisfecho si la red está congestionada.

### **1.3 ESQUEMA DE DISTRIBUCIÓN JUSTA DE LOS RECURSOS**

Otro esquema consiste en la posibilidad de considerar un modelo de distribución justa de los recursos de red, el cual expone Clark ([6], 1995). La pregunta es si la justicia sea en realidad el criterio de servicio deseado por los usuarios, y si existe una distribución adecuada de los recursos basados en la justicia.

Consideremos por ejemplo dos usuarios, en el cual uno envía un archivo 10 veces más grande que el del otro. El hecho de tener un ancho de banda mucho más grande disponible para su uso representa de por sí un beneficio para el usuario con el tráfico más grande, pero de todos modos no afectará significativamente al usuario con el archivo más pequeño. En términos de distribución eficiente de los recursos, el usuario que envía el tráfico más grande está obteniendo más de su “porción justa” de Ancho de banda durante la transferencia. Sin embargo, este esquema le permite a cada usuario utilizar su “porción justa”, no importa cuán grande sea, con tal de completar el envío.

Ser justo en cuanto a la prestación del servicio en la red no significa que a cada usuario se le asigne una porción igual de red para la transmisión de sus paquetes, sino que a cada usuario se le permite por igual enviar archivos del tamaño que requieran, percibiendo el mismo nivel de calidad de servicio. Es decir, que todos los usuarios sean tratados de forma proporcionalmente justa. A pesar de que esto puede ser objeto de abuso por parte de algunos usuarios, a la final puede satisfacer los requerimientos de todos ellos.

## 1.4 ESQUEMA DE PRECIOS POR PRIORIDADES BASADO EN LOS PRECIOS AUMANN-SHAPLEY Y EL VALOR SHAPLEY

Este esquema fue sugerido inicialmente por McLean y Sharkey ([17], 1994), y sus conceptos principales fueron extendidos luego por Beltrán y García ([4], 2000). La idea principal consiste en asignar prioridades a los paquetes entrantes en un nodo (*switch*) basados en un sistema de colas con disciplina FIFO (First In First Out).

De acuerdo con lo que propone este esquema, es posible definir un mecanismo de precios que satisface ciertas propiedades de distribución de costos, dentro del cual se encuentra explícito un vector de cantidades  $I$ , compuesto por  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  donde cada  $\lambda_i$  representa la cantidad de información por unidad de tiempo que el tipo de usuario  $i$  pide que sea procesada por un servidor. Como cada  $\lambda_i$  tiene las mismas unidades, se puede representar la cantidad agregada de información por unidad de tiempo como  $\lambda = \sum \lambda_i$ . Los precios Aumann-Shapley corresponden a los *costos marginales* de cada conjunto de requerimientos por unidad de tiempo que tiene una composición similar (o proporcional) a un nivel  $\lambda$  total ([4], Beltrán, García, 2000).

El valor Shapley por su parte corresponde a cierta función  $\psi$  que hace parte de un conjunto de posibles soluciones,  $\psi_i$ , al problema de lograr distribuir eficientemente los costos por concepto de pérdidas por congestión  $F$  entre los  $n$  jugadores o tipos de usuario que participan en un juego cooperativo con costo transferible<sup>6</sup>. La solución  $\psi_i$  será aquella que cumpla con ciertas propiedades y significará la distribución del costo que recibe el usuario de tipo  $i$ .

Los valores anteriores se utilizan por lo tanto para caracterizar un modelo de colas que simule el problema de la congestión. Por eso se calculan los costos que le corresponden a cada usuario  $i$  asumir por utilizar la red en presencia de congestión,

---

<sup>6</sup> Hace referencia a que los costos pueden ser distribuidos entre los  $n$  tipos de usuario de cualquier forma ([4], Beltrán, García, 2000).



lo que sirve para calcular una prioridad para cada usuario. Por medio de lo anterior, los autores pretendían establecer un criterio para determinar precios cuando el orden de prioridades se establece según la valoración de la información a transmitir, y no sobre el "tamaño" promedio de la misma (tiempo de servicio por requerimiento).

La principal ventaja del esquema es poder permitir que un usuario que envíe paquetes de mayor prioridad logre transmitir su información asumiendo un costo correspondiente mayor que el del usuario que envía paquetes de menor prioridad. Este resultado es consistente desde el punto de vista comercial práctico para lo que espera un Proveedor de acceso a Internet. El mecanismo puede asegurar al usuario de mayor prioridad un precio mayor. Sin embargo, esta situación no siempre es deseable. Por otro lado, el mecanismo debe extenderse para un modelo sobre colas finitas que cambiaría los tiempos de espera de los paquetes que llegan, pues el proceso de llegada ya no será el mismo.

## **1.5 EL ESQUEMA DE CAPACIDAD ESPERADA**

Este mecanismo, propuesto por Clark ([6], 1995), consiste en distribuir el servicio entre los usuarios con base en un rango de expectativas que se tienen de la red, y no de garantías. Es decir, sería incoherente garantizar cierto desempeño de la red a cierta hora del día, pues no es posible predecir su comportamiento dada su naturaleza cambiante y dinámica en la transmisión de información a través de ésta.

Por el contrario, es más lógico pensar en el desempeño que un usuario espera de la red (en cuanto a velocidad de transmisión y calidad de servicio), y podría especificarse como la mayor cantidad de información que a un usuario le gustaría enviar bajo cierto nivel de retardo, con una velocidad de transferencia. Esto se puede acordar entre el consumidor y el proveedor, afirmando algo así: "necesito que a cierta hora del día pueda transmitir un objeto de 50 kb con mínimo retardo cada 10 segundos ...".

La principal desventaja de esta propuesta es la inconsistencia que podría existir entre la oferta de la capacidad esperada (o los recursos de red) y la satisfacción del cliente. Es decir, si entre el origen y el destino no se ofrece la capacidad esperada suficiente, el usuario recibirá notificación de congestión en la red a pesar de haber acordado y haber pagado para que no ocurriera. Por otro lado, no se ha definido un modelo de capacidad esperada para la persona que recibe la información, lo que le impide tener pleno conocimiento de las condiciones en que la recibe.

## **1.6 CONCLUSIONES**

Los modelos anteriores representan aportes importantes de varios autores (cada uno para tratar de hallar solución al problema central desde distintos puntos de vista) en la búsqueda de un mecanismo que más se acerque al costo que habrá de asumir un usuario de la red de acuerdo con su comportamiento. Se busca que con este mecanismo se distribuyan los recursos de la red de manera justa, es decir, que cada usuario obtenga la calidad del servicio por la que pagó. En este sentido, de los anteriores probablemente el mecanismo más coherente y exacto sea el de precios por prioridades usando un modelo de colas, pues es capaz de valorar con mayor precisión los requerimientos de un usuario en la utilización de la red.

Sin embargo, una limitante en común que tienen casi todos los modelos expuestos anteriormente – a excepción del modelo de Tarifa Plana que ya se aplica – es que no consideran, incluyen o planean como un trabajo futuro alguna aplicación del modelo en la práctica para de este modo poder comprobar sus resultados en una red de datos real, y así determinar su nivel de validez. Lo que busca este trabajo es identificar un modelo sobre el cual se ha generado mucha teoría (que demuestra que se acerca mucho a lo que es medir los recursos de red utilizados por un usuario) y poderlo aplicar en la práctica, a través de tomas de datos y simulaciones con algoritmos, mostrando sus respectivos resultados y conclusiones.

## 2 LA TEORÍA DE ANCHO DE BANDA EFECTIVO

Los modelos expuestos anteriormente poseen ciertas debilidades a nivel práctico que sugieren la necesidad de proponer un mecanismo de control, aplicable en la realidad, que permita al usuario aprovechar al máximo los recursos de red profundizando en la medición de dichos recursos usados. Este procedimiento aumentará la eficiencia y exactitud en su utilización pero a su vez puede implicar mayor complejidad y posible dificultad en su entendimiento e implementación.

Este nuevo mecanismo es aplicable sobre redes de banda ancha, en las cuales es posible integrar diversos servicios al mismo tiempo sobre el mismo canal de Internet. En este capítulo se muestra en primer lugar que el análisis del tráfico agregado sobre una red de banda ancha es la base para el funcionamiento de este mecanismo en la práctica. En segundo lugar, se explica el concepto de ancho de banda efectivo y la importancia del término multiplexación estadística del tráfico de las fuentes, así como los criterios para determinar la calidad del servicio por medio de este mecanismo. En esta parte, se describe detalladamente el mecanismo y su funcionamiento enfocado hacia un punto de vista propuesto por Courcoubetis ([8], 1999) llamado teoría de grandes desviaciones, cuyo estudio y aplicación son necesarios para validar la aplicación del mecanismo en un escenario real.

### 2.1 LAS REDES DE BANDA ANCHA

A través de las redes de banda ancha puede circular tráfico de diferentes características y requerimientos de calidad en el servicio. Esto ha hecho posible que

el usuario común pueda acceder a más servicios a menores costos, pero a su vez ha implicado un análisis más profundo y complejo en cuanto a la administración de los recursos de red por parte del operador. El diseño y desempeño de estas redes dependerán del comportamiento estadístico del tráfico que origina cada fuente de la red.

El requerimiento de la calidad del servicio para cada uno de los tipos de tráfico presentes en una red de banda ancha difiere de un servicio a otro, no sólo en su valor sino en sus parámetros relevantes. Por ejemplo, algunos servicios requieren que no exista retardo en la transmisión, mientras que otros no son sensibles al retardo. Por otro lado, para otros servicios es indispensable que la pérdida de paquetes durante la transmisión sea mínima ([11] Gaviria, 2000).

Sobre estas redes puede circular tráfico proveniente de aplicaciones de voz, video y datos. Las aplicaciones de vídeo son las que generalmente representan el tráfico más complejo pues tiene una tasa de bits variable, originando retardo en el servicio y posible pérdida de la información por el nivel de congestión que puede producir en el canal de transmisión

Por su parte, las aplicaciones de datos también funcionan con una tasa de bits variable pero los paquetes son de menor tamaño y por lo tanto el nivel de *burstiness* o congestión es más bajo que en las aplicaciones de video.

Las redes de banda ancha son por lo tanto la base para construir la teoría de Ancho de banda efectivo. El objetivo de este trabajo de aquí en adelante será estudiar y simular el comportamiento del tráfico en agregado de una red de banda ancha, usando la teoría del ancho de banda efectivo.

## 2.2 EL CONCEPTO DE ANCHO DE BANDA EFECTIVO

La teoría del ancho de banda efectivo es un concepto desarrollado en los últimos años por diversos autores para la medición de los recursos en Internet, que representa de forma adecuada la interacción entre varias fuentes de información de diferentes clases, teniendo en cuenta las propiedades estadísticas de la red, sus condiciones de operación y los requerimientos de la Calidad del servicio ([13] [14], Kelly, 1995, 1996 ).

Un mecanismo que aplique este concepto tendrá la capacidad de medir de forma eficiente los recursos de red utilizados realmente por cierta fuente de tráfico en un escenario de red dado. Se han obtenido resultados con tráfico real a nivel tanto teórico como experimental que demuestran que el uso efectivo de los recursos depende no solamente de las propiedades estadísticas del tráfico de la fuente, sino de todas las otras fuentes que están siendo distribuidas o *multiplexadas* estadísticamente, así como de la capacidad y el tamaño del buffer del canal de red que contiene el tráfico multiplexado.

El principal concepto detrás de la teoría de ancho de banda efectivo es el de *multiplexación estadística*, que consiste en lograr, a través de un canal de banda ancha, la mezcla o superposición del tráfico de características distintas proveniente de muchas fuentes. En general, el tráfico que proviene de cada fuente no se considera una tasa constante sino que fluctúa alrededor de un valor medio, con un comportamiento estocástico no determinístico. Esto permite que el tráfico que proviene de esas muchas fuentes pueda ser transmitido simultáneamente a través del canal de banda ancha con una baja probabilidad de pérdida de paquetes.

El principal criterio para garantizar la calidad del servicio o cierto desempeño de la red es precisamente la probabilidad de pérdida de paquetes. Este parámetro depende del tamaño del buffer, que es aquella porción de canal que sirve para

almacenar los paquetes que no lograron ser transmitidos porque la tasa total de llegada excedió la capacidad del canal.

### **2.2.1 Interpretaciones del concepto de ancho de banda efectivo**

De acuerdo con Kelly ([13], 1995), este mecanismo permite no sólo distribuir de forma más eficiente la capacidad disponible de la red sino que permite retroalimentación acerca del nivel de congestión que se esté produciendo. Desde su punto de vista, este concepto surge como una elegante y poderosa pero compleja teoría matemática que posee diversas aplicaciones tecnológicas.

El punto de discusión radica en la forma en que debe aplicarse la tarifa que se calcula a partir de la medición de los recursos realmente utilizados por este método. Es decir, cómo y en qué momento se fija el precio a pagar por la prestación del servicio. Según lo afirma el mismo Kelly ([14], 1996), hasta el momento la práctica predominante ha sido asignar la tarifa una sola vez al inicio de la conexión (tarifa plana o Ex Ante ([2] Bailey, 1999)) por parte del proveedor del servicio, pero, como ya se mencionó, ésta no depende del nivel de uso de la red y por lo tanto no representa un método exacto, confiable y eficiente – aunque sencillo - si lo que se busca es optimizar su uso.

Kelly ([13] 1995, [14] 1996) plantea la posibilidad de contabilizar los recursos usados a lo largo del tiempo de conexión, calculando la tarifa en tiempo real y asignándola cada vez que se hace dicha contabilización. Bailey ([2], 1999) denominó este método como fijación de precios Dinámica o aproximación métrica. Propone dividir el tiempo total de conexión en intervalos predefinidos más pequeños, durante los cuales se transfieren paquetes de información entre el origen y el destino y viceversa. En cada intervalo de tiempo los paquetes circularán por la red bajo diferentes parámetros de tráfico, dependiendo de las condiciones de operación, como la congestión (lo que implica mayor uso de ancho de banda). Luego, para

cada intervalo de tiempo se contabiliza el nivel respectivo de los recursos de red realmente utilizados, y de este modo se puede obtener el acumulado de la utilización en todo el tiempo de conexión para finalmente asignar la tarifa correspondiente.

El esquema anterior representa gran complejidad y exige más control por parte de la red durante el proceso, pero permite mayor exactitud y capacidad para evaluar estadísticamente la utilización de la red en cada conexión. Una variante similar de este esquema consiste en imaginar la conexión como un conjunto de conexiones de menor duración, cobrando una cantidad adicional  $c$  por la obtención de cada una de estas conexiones pequeñas que el usuario considere necesarias para completar su transmisión.

Con este esquema es posible lograr que un usuario haga su propia evaluación del beneficio que obtiene al cambiar su régimen tarifario, y evaluar si los beneficios obtenidos de la predicción a corto plazo de la congestión que el usuario provoque en la red justifiquen el costo  $c$ .

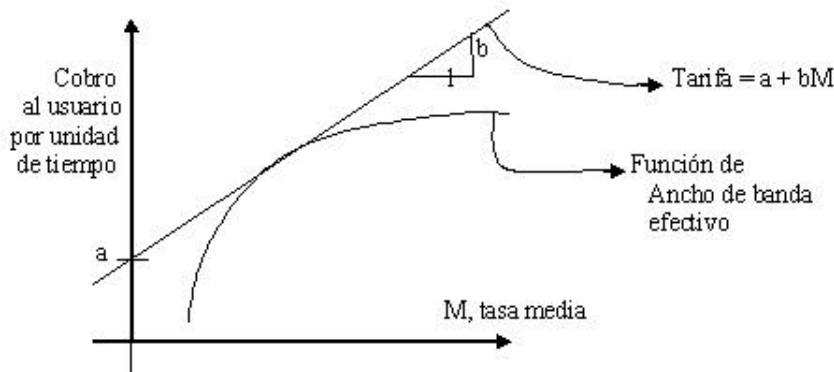
Por otro lado, Nagel ([3], 1999) sugiere la teoría del Ancho de banda efectivo basada en *Ex post pricing*, es decir, se fija el precio por el servicio luego de que éste ha sido consumido en su totalidad durante el período contabilizado. Por lo tanto, para fijar el precio se requiere tener pleno conocimiento del tráfico enviado y recibido por cada usuario.

Siris ([18], 1999) sugiere un esquema en el cual se representa la tarifa que se cobra a cada usuario como una recta en función de la tasa media actual  $M$ , tangente a la función de Ancho de banda efectivo (Figura 1). La ecuación de esta recta se define, por unidad de tiempo:

$$\text{Tarifa por unidad de tiempo} = f(M) = a + b \cdot M \quad (1)$$

donde  $a$  es el cobro por duración (que corresponde al corte de la recta con la ordenada), y  $b$  es el cobro por volumen enviado (que corresponde a la pendiente de la recta). Si multiplicamos ambos lados de la ecuación por el tiempo de conexión  $T$ , nos quedará:

$$\text{Tarifa} = F(T,V) = a \cdot T + b \cdot M \cdot T = a \cdot T + b \cdot V \quad (2)$$



**Figura 1. Representación gráfica de la tarifa cobrada al usuario como función de la tasa media del usuario (tomado de Siris [18], 1999).**

el cual es el cobro o la tarifa cobrada total por realizar una conexión con duración  $T$  y volumen de transferencia  $V$ . El proveedor puede ofrecerle al usuario diversas tarifas, correspondientes a diferentes rectas tangentes a la curva de Ancho de banda efectiva. La elección de tarifa que haga el usuario dependerá de la estimación que él mismo tenga acerca de la tasa media de su conexión,  $M$ .

Un usuario que tenga una tasa media esperada baja debería elegir una tarifa con un cobro por duración,  $a$ , bajo, mientras que un usuario con una tasa media esperada alta debería escoger una tarifa con un cobro por volumen,  $b$ , bajo. Con el fin de minimizar el cobro esperado por la utilización del servicio, el usuario debería escoger una tarifa que mejor se aproxime a su tasa media esperada, lo que permite que sea un cobro aproximadamente proporcional a la curva de Ancho de banda efectivo.



Este esquema de precios provee por lo tanto los incentivos adecuados al usuario para escoger la mejor tarifa, logrando restringir la duración y el volumen de la conexión, satisfaciendo sus necesidades sin afectar la Calidad del Servicio.

## 2.2.2 La teoría de grandes desviaciones

La calidad del servicio en redes de banda ancha permite que la probabilidad de pérdida de paquetes por sobrecarga del buffer sea muy baja (del orden de  $10^{-9}$ ) y por lo tanto este evento se vuelve raro dentro de la operación de la red ([11] Gaviria, 2000). Según Courcoubetis ([8], 1999), el ancho de banda efectivo se obtiene usualmente del análisis asintótico, es decir, tiene que ver con la forma como decrece la probabilidad de sobrecarga del buffer a medida que alguna otra variable del sistema aumenta.

Este análisis tiene dos puntos de vista:

### \* La Asíntota de Gran Buffer:

Cuando la cantidad que varía es el tamaño del buffer, tenemos la Asíntota de Gran Buffer. Es decir, la probabilidad de sobrecarga y pérdida de paquetes cae de modo exponencial con el tamaño del búffer,

$$P(\text{sobrecarga}) \approx e^{-\delta B} \quad (3)$$

donde  $\delta$  es una constante que depende de la capacidad del enlace y del comportamiento de las fuentes.

De acuerdo con Courcoubetis ([8], 1999), las definiciones de Ancho de banda efectivo basadas en la Asíntota de Gran Buffer son un tanto optimistas. Esto ocurre

porque este análisis no considera la ventaja que representa multiplexar estadísticamente las muchas fuentes independientes todas juntas a través del ancho de banda de la red. La clave en la definición del Ancho de banda efectivo está en que la cantidad de recursos usados depende no solamente de las propiedades estadísticas y los requerimientos de Calidad del servicio de una sola fuente, sino de las propiedades estadísticas de todas las otras fuentes que están siendo multiplexadas y de los recursos disponibles del canal de red.

\* La Asíntota de Muchas Fuentes:

Este caso se da cuando la probabilidad de sobrecarga del buffer decrece exponencialmente hacia la asíntota  $P(\text{sobrecarga})=0$  cuando aumentan el tamaño del sistema (el canal de transmisión y el parámetro de Calidad de Servicio,  $\gamma$ ) y el número de fuentes multiplexadas hacia un número grande, dejando constantes el tamaño del buffer por fuente y la capacidad por fuente.

La Asíntota de Muchas Fuentes es la base para construir los parámetros de espacio (s) y tiempo (t), cuyo significado se explicará más adelante. Juntos, los parámetros s y t constituyen lo que se llama el Punto de operación de la red, los cuales dependen de los recursos del canal y de las características del tráfico multiplexado. Además, tienen implicaciones importantes para el manejo del tráfico en Internet ([8] Courcoubetis, 1999).

### **2.3 EL ANCHO DE BANDA EFECTIVO Y LA ASÍNTOTA DE MUCHAS FUENTES**

El Ancho de banda efectivo puede definirse como un escalar que se le asigna a cierta fuente que origina tráfico, dependiendo del impacto neto que éste produzca en los recursos de la red. Este número depende en gran medida del volumen de tráfico

enviado, el tiempo de duración de la conexión y la existencia de sobrecarga (*burstiness*), y no simplemente de un conteo de bits o del tiempo total de conexión.

Supongamos que a un canal de red de banda Ancha llegan  $N$  fuentes independientes con  $J$  tipos de tráfico. Sea  $N_j$  el número de fuentes de tipo  $j$ , así:  $N_j = (n_1, n_2, \dots, n_j)$  donde  $n_1$  es el número de fuentes de tipo 1,  $n_2$  es el número de fuentes de tipo 2, etc. También suponemos que después de tomar en cuenta todos los factores económicos (como demanda y competencia) las proporciones del tráfico de cada una de las fuentes permanecen relativamente invariables con respecto a las cantidades expresadas en el vector  $N_j$  ([8] Courcoubetis, 1999). El Ancho de banda efectivo,  $\alpha$ , se define de la siguiente forma ([14] Kelly, 1996):

$$\mathbf{a}_j(s, t) = \frac{1}{s \cdot t} \cdot \log \left[ E \left( e^{s \cdot X_j(0, t)} \right) \right] \quad (4)$$

donde:  $s$  : Parámetro de espacio referente al nivel de multiplexación estadística (refleja que a mayor capacidad, menor probabilidad de pérdida de paquetes),

$t$  : Parámetro de tiempo que significa el tiempo en que el buffer se encuentra en estado de ocupado antes de la ocurrencia de sobrecarga en la red.

$E$  : Operador de Valor esperado

$X_j(0, t)$  : Variable aleatoria que representa la carga de tráfico generada por una fuente de tipo  $j$  en el intervalo  $[0, t)$ .

El Ancho de banda efectivo tiene la propiedad de que varía entre la media y el pico del tráfico originado por la fuente, dado por  $X(t)/t$ , a medida que  $s$  varía entre 0 e  $\infty$ , lo cual la hace una función positiva creciente con  $s$ . Kelly ([14], 1996-págs. 1-3) muestra en su documento una descripción detallada de la teoría detrás de las funciones  $\alpha$  y  $X(0,t)$ . Por otro lado, Gaviria ([11], 2000-págs. 19-25) muestra una

aproximación de un modelo de buffer, por medio del cual utiliza un análisis estocástico para obtener la ecuación (4) de Ancho de banda efectivo. Para esto aplica la función generadora de momentos logarítmica exponencial de la función de carga de tráfico, y saca el límite cuando  $t \rightarrow \infty$  (ver detalles en [7], Courcoubetis, 1996).

El concepto de Ancho de banda efectivo permite que se logren garantizar ciertas especificaciones sobre Calidad del Servicio (Quality of Service, QoS). Para esto, Siris ([18], 1999) muestra que se define una probabilidad  $\gamma$  tal que se limita la posibilidad de pérdida de paquetes de información (o de existencia de sobrecarga), así:

$$P(\text{sobrecarga}) = e^{-\gamma} \quad , \quad \gamma \geq 0 \quad (5)$$

donde  $\gamma$  es una constante predefinida como característica de la red. Dado lo anterior, debe satisfacerse la siguiente restricción para las J fuentes cuyo tráfico está siendo multiplexado en el canal de banda Ancha:

$$\sum_{j=1}^J n_j \mathbf{a}_n(s, t) \leq C + \frac{1}{t} \left( B - \frac{\mathbf{g}}{s} \right) = C^* \quad (6)$$

donde B es el tamaño del Buffer o de almacenamiento temporal de información en un punto de la conexión, C es la capacidad nominal del canal y C\* la capacidad efectiva. Esta ecuación se obtiene a partir de la función Asíntota de Muchas Fuentes que se enunciará más adelante.

Si suponemos que C, B,  $\gamma$  son constantes y que s y t no varían mucho con el tráfico, suponiéndolas igualmente constantes ([14] Kelly, 1996), entonces podemos decir que el valor C\*, que se define como Capacidad efectiva, es también invariable. Por lo tanto, es posible asignar a cada fuente de tráfico que se distribuye en la conexión

una cierta porción de Ancho de banda, de modo que la suma de todas esas porciones en agregado no supere la capacidad  $C^*$ . De igual forma, es posible aplicar linealidad en la distribución del ancho de banda con respecto al número de fuentes. Por ejemplo, si tenemos dos fuentes que poseen el mismo ancho de banda  $\alpha_1$ , podemos reemplazarlas por una sola fuente con un ancho de banda igual a  $2\alpha_1$ .

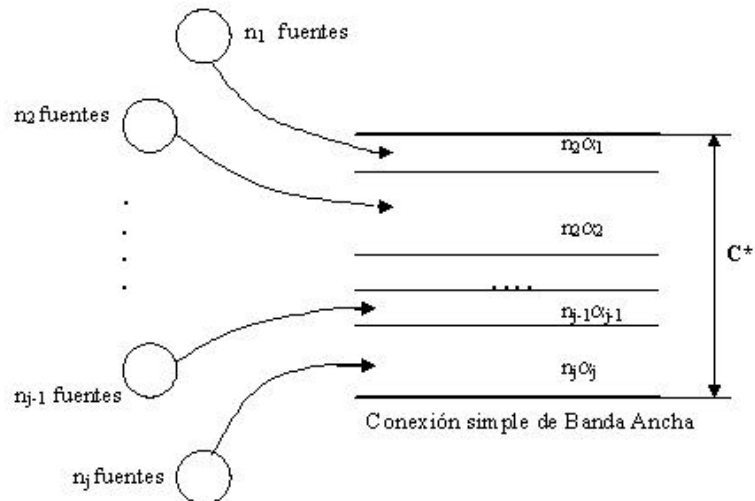
El significado de las ecuaciones (5) y (6) se puede interpretar de la siguiente forma: cuando el parámetro de calidad del servicio,  $\gamma$ , es igual a cero, significa que la probabilidad de pérdida de paquetes o sobrecarga es 1, es decir, el buffer se encuentra totalmente cargado (la sobrecarga es total) y por lo tanto todo el tráfico adicional que se haga circular a través del canal de conexión se perderá porque no hay porción adicional de conexión que guarde los paquetes enviados. En la ecuación (6), esto significa que  $C^*$  será mayor que  $C$  por la cantidad  $B/t$ , y dependerá solamente de  $t$  y no de  $s$ .

Por otro lado, **a medida que  $g$  aumenta** la probabilidad de sobrecarga del buffer decaerá exponencialmente hacia 0 y **la calidad del servicio irá mejorando**. Por lo tanto, como los parámetros  $\gamma$ ,  $s$ ,  $t$  son mayores o iguales a cero, la capacidad efectiva  $C^*$  será mayor que  $C$  por la cantidad  $(B/t) - (\gamma/s*t)$  y  $C^*$  será menor que en el caso anterior.

Según lo anterior,  **$C^*$  va disminuyendo conforme  $g$  va creciendo**, lo que quiere decir que una mejora en la calidad del servicio implica utilizar una porción adicional de conexión y por lo tanto se traduce una reducción proporcional en la capacidad efectiva disponible, lo cual es acorde con la realidad.

Las implicaciones de los parámetros  $s, t$  en la ecuación (6) se explicarán con mayor detalle más adelante.

Volviendo a la definición de Ancho de banda efectivo, la ecuación (4) representa para el operador de la conexión la posibilidad de maximizar su ingreso, puesto que le permite agregar usuarios a la red (fuentes de tráfico) hasta un máximo conocido  $C^*$ , sin afectar el desempeño de ésta (dado por la Calidad de servicio  $\gamma$ ). La figura 2 muestra un bosquejo general del funcionamiento de la distribución (*multiplexing*) aplicando el concepto de Ancho de banda efectivo.



**Figura 2. Representación de la distribución o repartición del tráfico de todas las fuentes de tipo  $j$  a lo largo de la capacidad de la conexión  $C^*$  (multiplexado), asignando a cada uno un Ancho de banda efectivo  $a_n$  ( $n=0,1,\dots,J$ ).**

### 2.3.1 Significado de los parámetros $s,t$

Como se dijo anteriormente, los parámetros  $s,t$  constituyen el punto de operación de la red. Tienen implicaciones importantes para el manejo del tráfico en Internet pues son responsables de caracterizar el contexto de tráfico en el que se localiza el usuario. Hay que recordar que se está considerando un sistema en el que varias fuentes de tráfico de  $J$  tipos están siendo multiplexadas a través de una conexión simple con capacidad  $C$  y tamaño de buffer  $B$  ([3] Bailey, Nagel, 1999).

El parámetro  $s$ , llamado parámetro de espacio, es un indicador del grado de multiplexado y depende de las tasas pico del tráfico de las fuentes multiplexadas. Esto significa que  $s$  es grande ( $s \rightarrow \infty$ ) cuando se multiplexan un número bajo de fuentes con picos de tráfico cercanos a la capacidad de la conexión (tráfico con alta variabilidad). Por otro lado,  $s$  es pequeño ( $s \rightarrow 0$ ) cuando se multiplexan un alto número de fuentes con picos de tráfico mucho menores que la capacidad de la conexión (tráfico con baja variabilidad) ([8] Courcoubetis, 1999). Se mide en unidades de  $\text{bits}^{-1}$ , como por ejemplo  $1/\text{kb}$ .

Como consecuencia, cuando el grado de multiplexado es bajo ( $s \rightarrow \infty$ ), el Ancho de banda efectivo de las fuentes se acerca a su tasa pico. Cuando el grado de multiplexado es alto ( $s \rightarrow 0$ ), el Ancho de banda efectivo de las fuentes se acerca a su tasa media.

Es posible asociar el valor de  $s$  con las variaciones en la probabilidad de pérdida de paquetes ( $\Delta\gamma$ ) y las variaciones en el tamaño del buffer ( $\Delta B$ ), así:

$$s = \frac{dg}{dB} \approx \frac{\Delta g}{\Delta B} = - \frac{\Delta \ln(P(\text{sobrecargar } g))}{\Delta B} \quad (7)$$

La anterior ecuación muestra que el parámetro  $s$  refleja cómo varía la probabilidad de pérdida de paquetes cuando el tamaño del buffer aumenta: de acuerdo con la ecuación (7),  **$s$  es grande ( $s \gg 1$ )** cuando **pequeños cambios en el tamaño del buffer** implican **grandes cambios en la calidad del servicio** (y por lo tanto en la capacidad efectiva del canal). Es decir, el tráfico de las pocas fuentes que están siendo multiplexadas es cercano a la capacidad de la conexión y es de alta variabilidad. Esto significa que el buffer es constantemente usado para el almacenamiento de paquetes propensos a perderse y el más mínimo aumento en el tamaño del buffer será altamente valorado bajo dichas condiciones de sobrecarga y aprovechado al máximo por el operador de la red. Un proveedor de acceso a

Internet podrá prestar una alta calidad en el servicio cuando sea capaz de ofrecer un cierto ancho de banda en condiciones de congestión y sobrecarga de la red.

Por otra parte, según la ecuación (7), **s es pequeño ( $s \ll 0$ )** cuando **grandes cambios en el tamaño del buffer** implican **pequeños cambios en la calidad del servicio**. Esto quiere decir que el tráfico de las muchas fuentes que están siendo multiplexadas es mucho menor a la capacidad de la conexión y es de baja variabilidad. En este caso la red se encuentra suficientemente descongestionada para que el buffer sea muy poco o nada usado para almacenar los paquetes que se puedan perder por sobrecarga. Bajo estas condiciones, un aumento considerable en el tamaño del buffer no incidirá grandemente en el desempeño de la red y por lo tanto en la calidad del servicio.

El parámetro  $t$ , por su parte, corresponde al valor esperado del tiempo en el que el buffer se encuentra en estado de ocupado, antes de la existencia de sobrecarga u *overflow* ([8] Courcoubetis, 1999). Es decir, durante el tiempo total de medición de trazas,  $T$ , es probable que varias veces el buffer se encuentre totalmente ocupado o cargado durante lapsos de tiempo, por lo general pequeños. Este estado implica que a partir de ahí los paquetes adicionales que se hagan circular a través del canal se perderán como consecuencia de la sobrecarga de la red. El parámetro  $t$  corresponde a esos lapsos de tiempo en que el buffer se encuentra ocupado, pero como este estado puede repetirse varias veces durante el tiempo total  $T$ , el parámetro  $t$  se toma como el valor esperado de ocupación del buffer calculando la media estimada (promedio) de todos esos lapsos de tiempo.

Este parámetro corresponde también al intervalo al final del cual se toman las trazas de tráfico y debe ser múltiplo entero del tiempo total  $T$ . Cuando  $t$  no se conoce, se puede inicialmente suponer tan pequeño como se quiera a conveniencia, dependiendo de qué tan exacto se desea el análisis. Mientras  $t$  sea más pequeño y  $T$  más grande, existirá una mayor granularidad en las trazas, se podrán obtener



mayor número de mediciones por unidad de tiempo y por lo tanto tener una mejor idea de la variación del tráfico en el tiempo. La variable  $t$  es una variable discreta.

Los parámetros  $s, t$  tienen implicaciones importantes sobre la ecuación (6) que es la restricción de capacidad para el tráfico de todas las fuentes siendo multiplexadas a través de la conexión. Dicha ecuación se reproduce a continuación:

$$\sum_{n=1}^J N_j \mathbf{a}_j(s, t) \leq C + \frac{1}{t} \left( B - \frac{g}{s} \right) = C^* \quad (6)$$

La capacidad efectiva  $C^*$  del canal es mucho mayor que la capacidad nominal  $C$  cuando  $s$  es grande y  $t$  es pequeño, es decir, se hace más grande cuando existe congestión y el tráfico de las fuentes multiplexadas es de alta variabilidad y se acerca a la capacidad de la conexión. Cuando existe mayor tráfico en la red, la capacidad efectiva tiende a aumentar con el fin de garantizar una mejor calidad en el servicio, y cuando existe menor tráfico ésta tiende a disminuir, lo cual es lógico con la realidad.

### 2.3.2 La asíntota de muchas fuentes

La siguiente ecuación corresponde a la llamada asíntota de muchas fuentes, la cual se obtuvo en el **Anexo 1** y se reproduce aquí:

$$-I = \sup_t \inf_s J(s, t) = \sup_t \inf_s \left[ s \cdot t \cdot \sum_{j=1}^J n_j \mathbf{a}_j(s, t) - s \cdot (B + C \cdot t) \right] \quad (8)$$

donde  $I$  se denomina la función asíntota. Como se dijo antes, esta ecuación es la base para calcular los parámetros  $s, t$  que caracterizan el tráfico de todas las fuentes siendo multiplexadas en el canal de banda ancha, y a partir de ahí poder calcular el Ancho de banda efectivo de todas las fuentes. En el capítulo 3 se intentará aplicar la

ecuación (8) en un escenario real utilizando un algoritmo, para finalmente determinar la factibilidad y utilidad del modelo usando la teoría en cuestión, que es finalmente el objetivo de este trabajo.

La solución de la ecuación (8) involucra dos procesos de optimización ([8] Courcoubetis, 1999), por lo cual contiene el operador sup inf: el primero consiste en minimizar la función  $J(s, t)$  con respecto a  $s$ , dejando fijo  $t$

$$J^*(t) = \min_s J(s, t) \quad (9)$$

y se despeja  $s$ . El segundo proceso consiste en maximizar la ecuación (9) con respecto a  $t$ , fijando el valor de  $s$  que se encontró en el primer proceso

$$-I = \max_t J^*(t) \quad (10)$$

Observando la ecuación (10), para calcular los parámetros  $s, t$  es necesario conocer el Ancho de banda efectivo  $\alpha_j(s, t)$ . Sin embargo, el cálculo se puede simplificar un poco pues el punto de operación varía muy poco en un sistema con muchas fuentes ([11] Gaviria, 2000). Este punto de operación se puede calcular con un estimativo del Ancho de banda de las fuentes, aproximando el operador de Valor esperado por un promedio empírico (estimado). El estimativo puede hacerse a partir de las mediciones de tráfico reales, de la siguiente manera ([8] Courcoubetis, 1999):

$$\mathbf{a}_j(s, t) \approx \frac{1}{s \cdot t} \log \left[ \frac{1}{T/t} \sum_{i=1}^{T/t} e^{s \cdot X_j[(i-1)t, it]} \right] \quad (11)$$

donde  $T$  es el tiempo total de medición. Obsérvese que es necesario que  $t$  sea un múltiplo entero del intervalo total  $T$ .

### 2.3.3 Utilidad y extensión de la asíntota de muchas fuentes

A partir de la definición de la ecuación de Asíntota de muchas fuentes (8), es posible dar una interpretación acerca de su significado, más allá de que su utilización a nivel práctico sea la base para obtener una solución al problema de calcular el Ancho de banda efectivo de cada fuente. Debido a que en la literatura consultada no se le da una interpretación suficientemente clara sobre cómo fue obtenida, sobre lo que realmente quiere decir y sus alcances (para qué se aplican los procedimientos que se hacen sobre ella), a continuación se le trata de dar una desde un punto de vista del autor, con la cual se busca ser lo más objetivo posible.

En primer lugar, la parte entre corchetes cuadrados de la ecuación (8) es una resta entre dos términos: el primer término (el término de Tráfico) representa la suma de los Anchos de banda generados por cada fuente para cada tipo de tráfico, es decir, el nivel de ocupación del canal de transferencia en el que se multiplexan los tráficos de dichas fuentes. El segundo término representa la capacidad de la red para soportar el tráfico que generan las fuentes en cuestión, y es función del tamaño del buffer  $B$  y de la capacidad nominal  $C$ , y por lo tanto se le puede llamar a este el término de Capacidad.

La operación de la resta se puede interpretar como la comparación entre esas dos cantidades, y de este modo observar qué tanta porción de red está siendo usada, o si por el contrario el tráfico de todas las fuentes sobre el canal está superando la capacidad de la red generando congestión (sobrecarga). Obsérvese que ambas cantidades son a su vez función de los parámetros  $s, t$  para que dentro de un proceso de optimización se puedan calcular sus valores.

Los dos procesos de optimización podrían explicarse de la siguiente forma: el primero como la minimización del espacio necesario para transmitir la información deseada (del Ancho de banda resultante de hacer circular el tráfico de todas las fuentes) con respecto al factor espacio, y el segundo buscando que esa situación

(mayor diferencia entre disponibilidad de Ancho de banda y tráfico circulado) se mantenga durante el mayor tiempo posible, con respecto al factor tiempo,  $t$ . Este procedimiento permitirá por lo tanto calcular los parámetros  $s, t$ .

Por otro lado, también es posible obtener algunos resultados importantes a partir de la ecuación de Ancho de banda efectivo que igualmente son usados en el análisis. En particular, la ecuación (6), que corresponde a la restricción de la definición de ancho de banda efectivo, puede obtenerse a partir de la ecuación (8) y se le puede dar una interpretación.

De acuerdo con Kelly ([13], 1996), la restricción (6) es la ecuación del plano tangente a cierta región  $A$  que encierra todas las soluciones a la ecuación de asíntota de muchas fuentes y cuya frontera es diferenciable en cierto punto característico  $n^*$  para el cual la solución será  $(s^*, t^*)$ , es decir, una pareja característica de los parámetros  $s, t$ . Este resultado se demuestra en detalle en el **Anexo 1**.

## 2.4 IMPLICACIONES Y CONCLUSIONES

En este capítulo hemos visto todo el *background* teórico detrás del concepto de Ancho de banda efectivo. Esto incluye comprender y visualizar el funcionamiento de las redes de banda Ancha y entender el análisis del tráfico en agregado que circula por ellas, así como interpretar el modelo teórico que puede ser aplicable en la realidad.

Es muy importante resaltar que antes de plantear una aplicación práctica para la medición de recursos en Internet por medio del modelo de Ancho de banda efectivo, es necesario definir cuáles escenarios a nivel práctico (en una red de tráfico real) pueden ser modelados o simulados a través de la teoría expuesta, y cuáles no. Es decir, la teoría de Ancho de banda efectivo representa un nivel considerable de

complejidad en su implementación y es probable que sea necesario ubicar el escenario real escogido en un contexto especial de tal forma que haya que hacer ciertas suposiciones, algunas de ellas muy fuertes incluso alejando la implementación práctica de lo que realmente sugiere el modelo teórico.

Por ejemplo, en redes de banda Ancha dentro del Internet ocurren muchos efectos que son particularmente difíciles de predecir y modelar a nivel teórico; tal es el caso del reenvío de paquetes o los efectos colaterales de la congestión en la red, entre otros, especialmente cuando entran en juego varios proveedores de acceso a Internet que buscan cada uno por su lado ofrecer una calidad del servicio óptima al menor costo posible. Su manejo le pertenece a una teoría que no le corresponde analizar al trabajo en cuestión.

### **3 APLICACIÓN PRÁCTICA DE UN MODELO DE PRECIOS USANDO EL MÉTODO DE ANCHO DE BANDA EFECTIVO**

Una vez comprendido el marco teórico, el siguiente paso consiste en cómo aplicarlo a un modelo de precios real. Como se explicó anteriormente, un modelo de fijación de precios usando la teoría de Ancho de banda efectivo permite medir los recursos utilizados por varias fuentes a través de conexiones que existen en la realidad, bajo ciertas suposiciones y usando conceptos y resultados básicos que sirven para simplificar su descripción, haciéndolo igualmente válido y perfectamente aplicable. Gracias a esto, el modelo puede usarse para comparar las características del tráfico de otras distintas fuentes en la realidad, como por ejemplo medir los patrones de tráfico de varias conexiones dentro de la red de la Universidad de los Andes, tal como lo hizo Nagel en la Universidad de Maryland ([3], 1999).

Lo anterior implicó buscar y utilizar los medios y herramientas técnicas necesarias para medir las tasas de transferencia existentes y determinar la presencia de congestión u otras características de operación. Luego, modificar o variar los patrones de tráfico para observar cómo cambia el nivel de los recursos usados, es decir, el Ancho de banda efectivo. Aunque no sea posible hablar de cobrar una tarifa por la prestación del servicio de Internet en una universidad, puesto que para los usuarios de instituciones educativas el servicio prestado es gratuito, sí servirá para determinar la porción de red realmente usada y calcular un “precio” como un porcentaje de los costos que representa para el proveedor asignar el Ancho de banda disponible a cada fuente.

### 3.1 DESCRIPCIÓN GENERAL DE LA APLICACIÓN PRÁCTICA

La situación que estamos considerando involucra la multiplexación del tráfico de las fuentes de la red de la Universidad de los Andes sobre el canal de Internet hacia el proveedor de acceso a la red, las cuales corresponden al tráfico proveniente de cada una de las terminales de la universidad conectadas a la red. Este tráfico está compuesto por tráfico de diferentes clases, tales como aplicaciones de Internet (WWW), correo electrónico, FTP (transferencia de archivos) y protocolos de transporte, entre otros.

El objetivo inicial era obtener las herramientas de medición y realizar el procedimiento sobre la red en cuestión. Para esto, se contactó al DTI (Dirección de Tecnologías de la Información) de la universidad, quienes administran el tráfico a través de la red de la universidad y se les pidió autorización para medir el tráfico en la red.

La medición se hizo por medio de la página de Internet de Diveo Uniandes<sup>7</sup>, a través de la cual la universidad tiene un convenio con la compañía Diveo de Colombia para obtener diferentes tipos de análisis del tráfico entrante y saliente de la red de los Andes. Estos análisis se hacen a partir de los datos de tráfico transferido en períodos pasados (días, semanas, meses) que han sido guardados en una base de datos para poder ser procesados. El acceso a esta página está restringido y solamente puede ser usada para fines académicos.

La figura 3 muestra la presentación de la página de Diveo Uniandes, con el menú de funciones para la obtención de datos. La función que nos corresponde utilizar para el proceso en cuestión es **Análisis de Tráfico** (Diagramas de tráfico específicos).

---

<sup>7</sup> La dirección de la página es: <http://noc.diveo.net.co/uniandes> [20].



### Universidad de los Andes

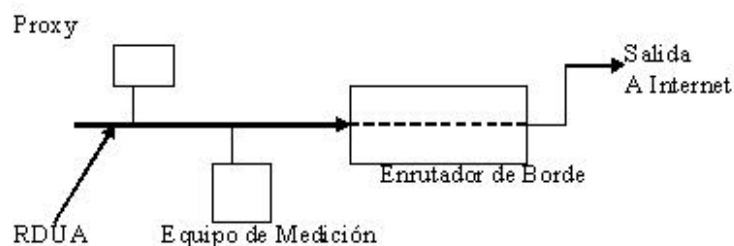
- [Graficas de uso del canal](#)
- [Latencia de última milla](#)
- [Análisis de tráfico Entrante](#)
- [Análisis de tráfico Saliente](#)
- [Análisis de tráfico](#)
- [Reporte de Seguridad](#)
- CPU Cisco 3600

### Diagramas de tráfico

Santafé de Bogotá, Colombia  
Copyright © 2003 DIVEO DE COLOMBIA

**Figura 3. Presentación de la página de Diveo Uniandes, con las funciones correspondientes (tomado de <http://noc.diveo.net.co/uniandes> [20]).**

El procedimiento utilizado por la página para la medición consiste en tomar todo el tráfico que circula a través del canal de salida a Internet y filtrarlo clasificándolo de acuerdo con ciertos criterios de análisis predefinidos. Esto se hace a través de un equipo de medición que se encuentra ligado al enrutador o *gateway* (puerta de entrada) que comunica a la RDU A (Red de la Universidad de los Andes) con el Proveedor de acceso a Internet. La Figura 4 muestra la ubicación física del equipo de medición en la RDU A.



**Figura 4. Ubicación física del equipo de medición en la RDU A (tomado de Betancourt ([5], 2002).**



De acuerdo con los datos del DTI de la Universidad de los Andes, la tasa de tráfico en el canal de Internet entre el proveedor y la red de la Universidad varía entre 6.3 y 7.2 Mbytes/s, y es prácticamente invariable entre las 8 de la mañana y las 5 de la tarde de Lunes a Viernes en el período académico normal (es decir, no vacaciones), período en el cual la demanda de la red es aproximadamente constante. Además, en algunos casos los equipos de algunas dependencias pueden dejarse funcionando a altas horas de la noche y en la madrugada, y domingos y días feriados (cuando las salas se encuentran cerradas al público) para actualizar sus aplicaciones de la red, intercambiar información con ésta o transmitir tráfico de acceso remoto que pasa por la red de los Andes, aprovechando que no están siendo ocupados por usuarios regulares.

A partir de la página de Diveo Uniandes es posible obtener los resultados de varios análisis. En particular, en este trabajo nos concierne un análisis en el cual pueden identificarse varios criterios de medición, clasificados de la siguiente manera:

- Por fecha: en la página se puede identificar todo el tráfico circulado por el canal de tráfico entre el proveedor de acceso a Internet y la red de la Universidad (entrada/salida) en un día específico.
- Por dependencia o zona: la página posee un campo que se llama Filtro IPS, que se usa, en un día en particular, para identificar y separar el tráfico entrante y saliente por dependencia o sala de la Universidad, por medio de su correspondiente dirección IP.
- Por tráfico de Entrada y Salida: para cada dependencia, es posible identificar el número de bytes transferidos entrando y saliendo del canal de tráfico entre el proveedor de acceso a Internet y la red de la Universidad.
- Por tráfico Local, Nacional o Internacional: para el tráfico de Entrada y Salida, se puede identificar que porción pertenece al tráfico que se ha transferido 1)

dentro de la red de la Universidad de los Andes (local), 2) por fuera de la red de los Andes pero dentro del país (tráfico Nacional que pasa a través del NAP Colombia<sup>8</sup>), y 3) a la red de Internet por fuera de Colombia (internacional).

- Por hora del día: es posible determinar el tráfico Local, Nacional e Internacional transferido a cada hora de un día determinado.
- Por tipo de tráfico transferido: para cada hora del día se especifica un porcentaje del tráfico total enviado. A partir de ahí podemos obtener un resumen del tráfico transmitido, donde para cada aplicación se especifican el número de conexiones que maneja y el número de bytes transferidos.

El objetivo global del experimento consiste en tomar las trazas del tráfico que fue transferido durante un período de referencia específico, como por ejemplo, una semana o incluso un mes, si es posible, de tal forma que sea representativo y pueda servir como base para estimar períodos con tráfico de similar comportamiento. Este período debe escogerse de manera que sus días muestren un comportamiento acorde con la realidad. Es decir, si se escoge un período del semestre académico normal de la universidad, se espera que el tráfico de cada uno de los días entre semana tengan un desempeño similar (mayor tráfico), y que lo mismo suceda para los días de fin de semana (menor tráfico), a menos que ocurran eventos inesperados en el funcionamiento de la red que rompan la tendencia.

### **3.2 IMPLEMENTACIÓN PRÁCTICA DEL ESQUEMA**

El análisis propuesto inicialmente consiste en medir los recursos de red utilizados por el tráfico generado por las conexiones dentro de las dependencias que hacen

---

<sup>8</sup> Network Access Point, o Punto de Acceso a Internet, por sus siglas en inglés. Hace referencia al punto de Acceso a la red de Internet en Colombia. Ver [10] Delgado, 2002.

parte del Departamento de Ingeniería de la universidad. La red interna de la universidad (RDUA) se compone de subredes que a su vez comprenden las distintas dependencias. Una sola dirección IP, la cual se usa en el filtro en la página, puede corresponder al tráfico generado por las terminales de varias dependencias, puesto que muchas veces el tráfico que origina cada terminal se redirecciona hacia otras dependencias antes de salir por el canal principal de acceso a la RDUA. Ahora, a cada terminal le corresponde una sola dirección IP. Esta dirección permitirá identificar el tráfico proveniente del grupo de conexiones que hacen parte de una dependencia en particular.

Se debe definir a continuación un programa de trabajo que permita obtener las trazas de tráfico de forma ordenada, basados en los criterios expuestos en el numeral 3.1.

Como se explicó antes, este proceso de medición es bastante complejo, porque implica que el software de medición tome *todo* el tráfico que circuló a través del canal de salida a Internet entre la universidad y el proveedor – cuyos datos se encuentran guardados en la base de datos - y lo filtre clasificándolo de acuerdo con los criterios seleccionados en el numeral 3.1. En la página de Diveo Uniandes, este proceso toma bastante tiempo debido a que el software de la página está dirigido principalmente a obtener análisis más generales y globales, y no tan específicos como el que nos concierne.

Por lo anterior, en la obtención de los datos se requiere de paciencia y probablemente no se logren visualizar todos los datos que se desean. Además, aunque no sea necesariamente lo más adecuado, es preferible hacer las consultas en la página cuando no se encuentre muy congestionada (durante los fines de semana o en horas en la noche).

### 3.2.1 Las trazas de tráfico

Durante el proceso de toma de trazas de tráfico en la red de los Andes por medio de la página de Diveo Uniandes, efectivamente surgieron problemas en la obtención de los datos. La consulta a través de la página para una dirección IP de una dependencia en particular, a una cierta hora de un cierto día, resultó siendo un proceso muy lento y de bajo desempeño. Es decir, no solamente le tomaba bastante tiempo a la aplicación buscar los datos y visualizarlos en pantalla, sino que en el 90% de las consultas realizadas la página mostraba error y no mostraba dato alguno (“communication problem”). Esto ocurría sin importar el tipo de terminal desde la cual se tomaba la medición, o los criterios usados para obtener los datos (fecha, hora, dependencia), si el acceso era de tipo remoto o local, o si la medición se hacía entre semana en horario académico o los fines de semana a altas horas de la noche, cuando más se esperaba que el proceso resultara más sencillo.

Se pudo concluir tras conversaciones sostenidas con el DTI que lo anterior se debió a un problema en la aplicación de Diveo para realizar ese tipo de medición tan compleja – en el sentido de tener que usar tantos criterios filtrando el tráfico global hasta casi su mínima expresión –. Seguramente tuvo que ver también con la estructuración y configuración interna de la red que pudo haber generado restricciones en el manejo de la información, quizá relacionados con acuerdos de confidencialidad. En definitiva, esta aplicación fue diseñada para obtener análisis más globales, como el desempeño de la red medido directamente sobre todo el ancho de banda del canal de tráfico entre la Universidad y el proveedor, y no a granel, como se pretendía con este proceso.

En cierto modo se esperaba desde antes que ocurrieran problemas de este tipo, debido a la dificultad del proceso. Probablemente ésta no es la herramienta más adecuada para la toma de trazas de tráfico desde un punto de vista tan específico, pero quizá también puede ser que este tipo de medición es muy poco común (por la poca difusión que tiene el esquema en estudio) y de ahí que no existan muchos

métodos de hacerlo de forma rápida y confiable, a menos que se tome un caso muy específico y no tan generalizado como el de la red de los Andes.

Sin embargo, fue posible obtener los datos del tráfico de Entrada y Salida generado por las conexiones correspondientes a la dirección IP 157.253.62.0, algunas de las cuales pertenecen a la Sala Tyba de Departamento de Ingeniería (dirección 62), para el día Domingo 15 de Junio de 2003, durante las primeras 11 horas del día, período en el cual se encontraban en funcionamiento dichas terminales.

Los datos se solicitaron en la página de la siguiente manera: en primer lugar se elige el mes en el cual se escoge el día sobre el cual se desea hacer el análisis. Después, para el día correspondiente la aplicación le pide al usuario la dirección IP y se comienza a hacer la consulta. La Figura 5 muestra cómo se visualizan estos pasos.

**Cliente:** Universidad de los Andes  
**Redes:** 157.253.0.0/16

<< Mayo		Junio 2003					Julio >>	
Domingo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado		
1	2	3	4	5	6	7		
8	9	10	11	12	13	14		
15	16	17	18	19	20	21		
22	23	24	25	26	27	28		
29	30							

[Salir](#)

**Análisis de tráfico total para 2003\_**  
**Cliente** Universidad de los Andes  
**Redes** 157.253.0.0/16  
**Filtro IPS**

**Figura 5. Visualización del calendario para seleccionar el día del análisis y la solicitud de la red IP para realizar la consulta**

A continuación, luego de haberse tomado algunos minutos en la consulta (en los casos en que ésta no mostró error), aparecen los datos mostrados tal como aparecen en la Tabla 1. En esta tabla se muestran, para el tráfico de Entrada y Salida de la red de los Andes, el número de conexiones, de paquetes y bytes

enviados, involucrados en el tráfico que ha sido transferido de forma local, nacional e internacional.

<b>Entrada</b>				<b>Salida</b>			
Origen	Conexiones	Paquetes	Bytes	Origen	Conexiones	Paquetes	Bytes
Local	0	0	0	Local	0	0	0
Nacional	1	3	144	Nacional	0	0	0
Internacional	3,274	3,653	406,712	Internacional	859	990	75,578
<b>Total</b>	<b>3,275</b>	<b>3,656</b>	<b>406,856</b>	<b>Total</b>	<b>859</b>	<b>990</b>	<b>75,578</b>

**Tabla 1. Datos de tráfico como se muestran en la aplicación**

La tabla 1 nos da una idea global del tráfico que se ha manejado por esa dependencia durante ese día. Sin embargo, deseamos poder obtener mediciones de tráfico más específicas, es decir, con trazas de tráfico tomadas al final de intervalos de tiempo más pequeños, como horas (o minutos y segundos, si es posible) en lugar del tráfico por días. Para esto, también se despliega una tabla como la Tabla 2, en la que se muestra el tráfico transmitido a cada hora del día por porcentaje, tanto para la entrada como para la salida.

En la Tabla 2 podemos acceder al tráfico transmitido durante cada hora, haciendo clic en la hora respectiva, ya sea para el tráfico de entrada o salida. De este modo obtenemos algo como lo mostrado por la Tabla 3, en la cual se muestran las distintas aplicaciones manejadas durante la transferencia de los datos en una hora de un día específico (en este caso entre las 0:00 horas y las 0:59 del día, o sea, de la madrugada del 15 de Junio de 2003). Para las demás horas del día se obtienen cuadros similares, que varían obviamente dependiendo del tráfico que ha sido gestionado por las diferentes conexiones.

<b>Entrada Total por hora:</b> <b>406,856 bytes</b>		<b>Salida Total por hora:</b> <b>75,578 bytes</b>	
Hora	Porcentaje	Hora	Porcentaje
00:00	7.44%	00:00	4.97%
01:00	10.84%	01:00	18.59%
02:00	6.98%	02:00	5.20%
03:00	6.48%	03:00	3.55%
04:00	4.67%	04:00	4.97%
05:00	10.23%	05:00	3.55%
06:00	10.46%	06:00	24.65%
07:00	8.63%	07:00	3.82%
08:00	11.58%	08:00	29.41%
09:00	11.14%	09:00	1.28%
10:00	11.55%	10:00	0.00%

**Tabla 2. Tráfico de entrada y salida por hora del día, en porcentaje del total**

Obsérvese que se transmite tráfico de protocolos de distintas clases (que serán explicados más adelante) y que además el tráfico no solamente proviene de la conexión de la subred solicitada (subred 62, en el caso de la Tabla 3) sino de conexiones remotas de las cuales también se especifica su dirección IP.

Por otro lado, el tráfico total (de entrada o salida) para una hora en particular es simplemente la suma de lo transferido por cada dirección IP (tanto para paquetes, conexiones y bytes).

Por lo tanto, para cada hora de cada día en particular, es necesario obtener los datos como los de la tabla 3 y separarlos por aplicación, de acuerdo con los criterios

definidos en el numeral 3.1. De este modo, hemos llegado a la mínima expresión a la que era posible llegar una vez se logró filtrar toda la información por criterios.

00:00		00:59				
Entrada Total: 30,252 bytes						
Local	Remoto	Protocolo	Conex.	Paquetes	Bytes	Porcentaje
157.253.62.33	193.108.154.55	ICMP	10	13	520	1.72%
157.253.62.33	63.211.236.234	ICMP	10	13	520	1.72%
157.253.62.33	207.46.245.10	UDP-DNS	5	6	432	1.43%
157.253.62.10	133.205.179.251	UDP-Other	1	1	404	1.34%
157.253.62.12	81.185.37.25	UDP-Other	1	1	404	1.34%
157.253.62.13	195.43.182.146	UDP-Other	1	1	404	1.34%
157.253.62.17	211.155.234.28	UDP-Other	1	1	404	1.34%
157.253.62.28	12.81.8.142	UDP-Other	1	1	404	1.34%
157.253.62.37	199.0.45.2	UDP-Other	1	1	404	1.34%
157.253.62.40	211.155.234.28	UDP-Other	1	1	404	1.34%
Otros	Otros	Otros	104	116	25,952	85.79%
		SUMA	136	155	30252	100%

**Tabla 3. Tráfico transferido para una hora específica de un día en particular, por aplicación**

En la tabla 4 se especifican las trazas de tráfico que fueron tomadas (solamente para el **tráfico de Entrada**), colocadas en una tabla en la que se pueden observar con mayor facilidad los datos clasificados por criterios, y la cual será usada directamente en el algoritmo usado para calcular los parámetros s,t y el ancho de banda efectivo usado por estas conexiones en el período de tiempo especificado.



### 3.2.2 Consideraciones respecto a las trazas de tráfico

Los datos obtenidos se muestran en la siguiente tabla:

Hora	Conexiones o fuentes						Bytes transmitidos					
	TCP-WWW	ICMP	UDP-DNS	TCP-Otro	UDP-Otro	Otro	TCP-WWW	ICMP	UDP-DNS	TCP-Otro	UDP-Otro	Otro
00:00 - 00:59		20	5		7	104		104	432		2828	25952
01:00 - 01:59	17	12	5		7	298	2304	520	432		2828	38010
02:00 - 02:59					10	118					4040	24374
03:00 - 03:59	31				9	97	3744				3636	18996
04:00 - 04:59		16			8	86		104			3232	14744
05:00 - 05:59	1	9	5	9	6	496	1460	520	432	456	2424	36312
06:00 - 06:59		8	10		7	344		520	864		2828	38360
07:00 - 07:59		9			9	328		520			3636	30968
08:00 - 08:59		20	4		7	411		104	432		2828	42976
09:00 - 09:59			11		4	350			5958		1616	37740
10:00 - 10:59		9			9	359		480			3636	42878
SUMA	49	103	40	9	83	2991	7508	568	8550	456	33532	351310

**Tabla 4. Datos correspondientes al tráfico total de Entrada a la RDU, para la subred 62 el día Junio 15 de 2003**

Por lo explicado anteriormente, solamente fue posible obtener estos datos. Son apenas 11 datos correspondientes a cada hora para cada tipo de tráfico. Los tipos de tráfico especificados en la tabla tienen el siguiente significado (el término común es el protocolo de transporte TCP o Transfer Control Protocol):

- TCP-WWW: Protocolo de transporte para aplicaciones de la Web (Internet utilizando browsers).
- ICMP (Internal Control Messenger Protocol): Es un protocolo usado básicamente para funciones de control.
- UDP-DNS (DNS: Domain Name Server): Protocolo usado para almacenar los códigos asignados a cada sitio en la red, los cuales se guardan como nombres a partir de los cuales se puede identificar las direcciones de dichos sitios.
- TCP-Otro, UPD-Otro, Otro: Comprenden otras aplicaciones de red que son más comunes pero que la página de Diveo Uniandes no las identifica y ni las mide por separado como sucede con las anteriores. Estas aplicaciones corresponden por ejemplo a transferencia de archivos (como FTP), correo electrónico y muchas otras aplicaciones que, como se puede observar, son las que utilizan el mayor número de conexiones y transmiten la mayor cantidad de información, en grandes proporciones con respecto a las otras aplicaciones especificadas.

Al intentar aplicar estos datos al algoritmo pensado inicialmente, se llegó a dos conclusiones: primero, en la tabla 4 las aplicaciones asociadas al mayor número de conexiones son: Otro (de todo tipo) con 2991 conexiones, y en menor proporción ICMP con 103 conexiones y UDP Otro con 83 conexiones. Estas 3 hacen la mayoría sumando con el resto un total de 3275 conexiones, el cual se puede considerar un número grande de conexiones. Este resultado está de acuerdo con el hecho de aplicar la ecuación de asíntota de muchas fuentes para la solución al problema, la cual es válida cuando  $N$  (el número de fuentes) es grande o tiende a infinito.

Por otro lado, son necesarios muchos más datos que los 11 disponibles mostrados en la Tabla 4, los cuales no son suficientes para que el modelo pueda sustentarse sobre una base práctica sólida. Es decir, mientras más datos se tengan, el rango de valores de la función minimizada inf a partir del cual se determinará su valor máximo

con el fin de encontrar el valor de  $t$  se hará más grande, y habrá más posibilidades de encontrar una solución a la ecuación.

Por lo anterior, es necesario hacer una evaluación de los datos de la tabla 4 con el fin de encontrar una estimación a partir de esos datos que genere una mucho mayor cantidad de información. Para esto, se realizó el siguiente análisis:

Se tomaron los datos de Bytes transferidos en todas las aplicaciones para cada hora (6 números), y se tomaron como un patrón de referencia para poder estimar más datos de similares características para esa hora en particular. Es decir, en cada hora en que se obtuvieron los datos lo más probable es que durante intervalos de tiempo más pequeños que 1 hora (como por ejemplo cada minuto o cada segundo) se hayan transmitido cantidades similares de bytes a los que se muestran en la Tabla 4. En este sentido, los datos de la tabla 4 podrían considerarse como un promedio o una media estimada de todos los datos (bytes transferidos) que se generaron durante esa hora en particular.

Según lo anterior, se utilizó una pequeña rutina en Excel que generara datos aleatorios pero cercanos (según estimación del autor) a los de la tabla 4. Es decir, respetando ciertos comportamientos, como que por ejemplo los bytes transferidos para TCP-WWW varían en la Tabla 4 entre 1000 y 4000 aproximadamente, así que no se generarán datos mayores a 5000 bytes, para este caso. O también, se puede observar un patrón en los datos para las aplicaciones ICMP y UDP-DNS, así como que los datos de UDP-Otro son todos múltiplos de 404 y no sobrepasan el valor 4040 (o sea  $404 \cdot 10$ ).

Así que se generaron 100 datos por cada hora especificada en la Tabla 4, lo que se podría interpretar como que cada dato fue tomado cada 36 segundos ( $=3600$

segundos/100) <sup>9</sup>. En total se crearon 1100 datos, pero se dejó en 1000 por cuestiones de practicidad correspondientes a las primeras 10 horas (para tener un mayor número de múltiplos y generar más posibles t). Los espacios en blanco que se pueden notar en la Tabla 4 entre los datos se toman como cero, es decir, que en esa aplicación no se transmitieron bytes a esa hora.

### 3.2.3 El algoritmo de optimización

Una vez generados los datos, se decidió implementar el algoritmo del Ancho de banda efectivo en MATLAB, utilizando la ecuación de Asíntota de Muchas Fuentes para encontrar los valores de los parámetros s,t que caracterizan el tráfico en cuestión. Para esto, se debe primero tener conocimiento de los siguientes parámetros:

- Número de tipos de fuentes que tiene acceso al canal de red.
- Número de fuentes o conexiones desde las cuales se origina tráfico.
- Carga de tráfico que cada fuente impone sobre el canal de red
- Capacidad del canal
- Longitud del buffer
- Calidad del servicio

El valor del Ancho de banda efectivo que cada fuente ocupa sobre el canal sólo puede calcularse si se conocen los parámetros s,t. Para su cálculo, se utiliza el procedimiento descrito en el **Anexo 2** a partir de la ecuación sup inf, que fue obtenida en el **Anexo 1** y se reproduce aquí:

---

<sup>9</sup> Sin embargo, dentro del análisis del algoritmo este valor se considera grande para tomarse como intervalo de medición. Aún así, es posible correr el algoritmo y calcular el ancho de banda efectivo de forma coherente con este dato.

$$\sup_t \inf_s J(s,t) = \sup_t \inf_s \left[ s \cdot t \cdot \sum_{j=1}^J n_j \mathbf{a}_j(s,t) - s \cdot (B + C \cdot t) \right] \quad (8)$$

El primer paso consiste en suponer un valor inicial de t conocido, por ejemplo  $t_0$ , reemplazarlo en la ecuación (8) y luego minimizar la expresión entre corchetes con respecto a s. Esto se hace en MATLAB por medio de la función `fminsearch`, así:

```
s=fminsearch('calculo', Xo)
```

El programa comienza a buscar a partir de un valor inicial para s ( $X_0$ ) y encuentra un mínimo local de la función especificada, usando el método simplex de búsqueda directa de Nelder-Mead.

Una vez se tiene el valor de s, el siguiente paso consiste en suponer una serie de valores de t (que sean múltiplos del tamaño total de la traza, T) y reemplazarlos en la ecuación (8), de tal forma que podamos encontrar el máximo valor de la expresión entre corchetes en dicha ecuación. Con este vector resultante se puede determinar, identificando el valor máximo, el valor de t. A partir de este proceso obtenemos al menos una pareja de valores (s, t) como solución a la ecuación sup inf.

Esta pareja de valores servirá para calcular el Ancho de banda efectivo de una fuente en particular, usando la ecuación (11) de este documento. Una descripción detallada del funcionamiento del algoritmo se encuentra en el **Anexo 1**, e incluye el código de dicho algoritmo en la simulación de los datos obtenidos.

### 3.3 RESULTADOS Y CONCLUSIONES

Para comenzar, se utilizaron los siguientes datos para la ejecución del algoritmo:

- \* Los 1000 datos de tráfico (bytes transferidos) para cada tipo de aplicación.
- \* B=80 Tamaño del buffer, en kilobits
- \* C=40000 Capacidad nominal del canal, en kilobits / s <sup>10</sup>
- \* to=36 Intervalo de medición de trazas, en s
- \* T=36000 Tiempo total de medición, en s (equivalente a 10 horas).
- \* fuentes1=49 Número de fuentes de tipo 1
- \* fuentes2=103 Número de fuentes de tipo 2
- \* fuentes3=40 Número de fuentes de tipo 3
- \* fuentes4=9 Número de fuentes de tipo 4
- \* fuentes5=83 Número de fuentes de tipo 5
- \* fuentes6=2991 Número de fuentes de tipo 6
- \* Xo=0.05 Valor de partida del proceso de minimización

Los resultados del algoritmo descrito en el **Anexo 2** utilizando estos datos, se muestran a continuación luego de su ejecución en MATLAB:

```
> yfinalFIN

    0.0127      Valor de s

1.0e+005 *

Columns 1 through 7  Evaluación de las función entre corchetes en (8)

   -0.0365   -0.3397   -0.1553   -0.1559   -0.3730   -0.8860   -1.0344

Columns 8 through 9

   -2.0712   -2.7695
```

---

<sup>10</sup> Esta cantidad sugerida de la capacidad del canal, 40000 kilobits / s, equivale a 5000 kilobytes / s (dividiendo por 8) que es aproximadamente cercano al tráfico que circula normalmente a través del canal de salida de la red de la Universidad, en horas pico, mencionado en la Sección 3.1 (entre 6300 y 7200 kilobytes / s).

El vector mostrado al final corresponde como se dijo a los valores resultantes de evaluar la función entre corchetes de la ecuación (8) para el  $s$  calculado y para los  $t$  propuestos, que son  $t = [36 \ 72 \ 144 \ 180 \ 360 \ 720 \ 900 \ 1440 \ 1800 \ 3600]$  segundos. El valor de  $t$  para el cual el vector generado por MATLAB tiene su valor máximo (-0.0365) es  $t=36$  segundos.

Lo anterior quiere decir que la pareja de parámetros  $s,t$  es **(0.0127 kbits<sup>-1</sup> , 36 segundos)** es la solución de la ecuación correspondiente de asíntota de muchas fuentes. Con los parámetros  $s,t$  podemos calcular el valor del Ancho de banda efectivo para este caso en particular, usando para tal efecto la ecuación (11) del documento:

$$\mathbf{a}_j(s,t) \approx \frac{1}{s \cdot t} \log \left[ \frac{1}{T/t} \sum_{i=1}^{T/t} e^{s \cdot X_j[(i-1)t, it]} \right] \quad (11)$$

Reemplazando todos los valores finalmente resulta el siguiente valor:

$$\mathbf{a}_j(s,t) \approx 0.182 \frac{\text{kilobits}}{s} = 182 \frac{\text{bits}}{s} \quad (12)$$

Esta será por lo tanto la capacidad de la red utilizada en la transmisión del tráfico proveniente de 3275 conexiones distintas y de 6 diferentes clases, correspondiente a la subred 157.253.0.62 (dentro de la cual se encuentra la sala tyba de Ingeniería) para las primeras 10 horas del domingo 15 de Junio de 2003. Este valor es cerca de 400,000 veces más pequeño que la capacidad nominal de la red, que se definió en 40000 kilobits/s.

Este resultado es lógico y coherente, teniendo en cuenta la dificultad del proceso y las posibles fallas que este pudo haber tenido por las suposiciones y consideraciones especiales que se hicieron. El valor es pequeño y es consistente con el hecho del alto nivel de filtración realizado en los datos para llegar hasta su mínima expresión.

## 4 MODELO DE COSTOS Y FIJACIÓN DE PRECIOS

La teoría de Ancho de banda efectivo permite establecer una medida de la utilización de los recursos. Esto permitirá que, por medio de un Acuerdo de Nivel de Servicio (SLA), el administrador de la red decida cuantos contratos ofrecer de manera simultánea para a partir de lo cual crear un mecanismo de tarifas. Por lo tanto, este acuerdo deberá generar en los usuarios incentivos para conocer su comportamiento de red (perfil de tráfico) y así seleccionar los parámetros de contrato que más se ajusten a sus necesidades, haciendo posible un óptimo uso de esta ([9] Courcoubetis, 1999).

La creación de un sistema de tarificación es un aspecto fundamental a considerar tanto para los usuarios como para los proveedores del servicio de acceso a la red, y los involucra a ambos por medio del SLA. Este sistema deberá crearse con base en parámetros conocidos o mesurables, de modo que el proveedor de la red no tenga que crear un sistema adicional para reflejar el uso, lo cual aumentaría los costos.

Para un usuario, un esquema de tarificación adecuado es aquel que refleja realmente la utilización de los recursos de la red, de manera que maximice el beneficio por el servicio al menor costo posible.

Para el proveedor del servicio, el esquema de tarificación debe cumplir dos funciones:

- A nivel económico: la asignación de tarifas busca recuperar la inversión y generar utilidades.



- A nivel técnico: un esquema de tarificación apropiado permite al operador administrar sus recursos de manera óptima, evitando así la congestión en la red y garantizando a sus usuarios una calidad de servicio mínima ([11] Gaviria, 2000).

En este capítulo se presenta un modelo de costos y fijación de precios basado en el concepto de Ancho de banda efectivo. Es posible validar este modelo con los datos reales obtenidos del análisis cuantitativo del capítulo anterior, con el fin de observar el verdadero efecto de la medición de los recursos en el funcionamiento del sistema tanto para los usuarios como para el operador de la red. Esto a su vez servirá para verificar la importancia del concepto de Ancho de banda efectivo como parámetro de cuantificación de los recursos. Este procedimiento podría hacer parte de un trabajo posterior.

#### **4.1 EL MODELO DE COSTOS**

En el modelo de costos consideramos una conexión simple, con una capacidad de canal  $C$  y tamaño de buffer  $B$ , la cual se comparte entre el tráfico de un cierto número de usuarios que han negociado su propio Acuerdo de Nivel de Servicio (SLA) con su respectivo proveedor, es decir, han seleccionado contratos de tráfico.

Para el caso específico de la red de la Universidad, es posible que el administrador de acceso a la red (el proveedor) logre asegurar un nivel predeterminado de Calidad en el servicio, teniendo en cuenta que sus usuarios conforman un grupo de individuos que originan tráfico del mismo tipo, o con parámetros similares (es el caso de estudiantes y profesores). Por lo tanto, se creará un SLA que considere los parámetros específicos de desempeño y de perfil de tráfico del usuario para esta situación.

Para comenzar, suponemos que la conexión soporta solamente una sola clase de servicio con una calidad específica que ha sido acordada en un SLA específico. A nivel económico, el objetivo de los usuarios es maximizar su utilidad, la cual depende de un vector de parámetros de tráfico contratados  $\mathbf{x}_i$  para cada usuario  $i$  (según su comportamiento en cuanto a tasa media, calidad de servicio, capacidad), y de su función de utilidad  $U_i(\mathbf{x}_i)$ . El operador busca distribuir sus recursos a sus usuarios de modo sirva al mayor número de usuarios, maximizando la suma de las utilidades (beneficio social) y garantizando una Calidad en el servicio ([11] Gaviria, 2000).

Esta situación puede expresarse como un problema de optimización. Suponiendo un número  $i$  de usuarios de un solo tipo en la red, se tiene que:

$$\max_{x_i} \sum_i U_i(x_i) \tag{13}$$

Sujeto a:

$$\sum_i \mathbf{a}_i(x_i) \leq C^* \tag{14}$$

donde  $U_i$  es la utilidad del usuario  $i$ ,  $\mathbf{a}_i(x_i)$  es la medida de uso de recursos (ancho de banda efectivo) consumidos por el contrato  $x_i$ ,  $C^*$  es la capacidad efectiva de la red, la cual depende de la capacidad  $C$ , del tamaño del buffer  $B$  y de la probabilidad de pérdida de paquetes ( $\gamma$ ) o calidad en el servicio.

El problema anterior es equivalente a:

$$\max_{x_i} \left[ \sum_i U_i(x_i, C) - p \cdot \left( \sum_i \mathbf{a}_i(x_i) - C^* \right) \right] \tag{15}$$

Las condiciones de primer orden del problema de optimización anterior se pueden expresar así:

$$\frac{\partial U_i(x_i)}{\partial x_i} = p \cdot \frac{\partial a_i(x_i)}{\partial x_i} \quad (16)$$

donde  $p$  es el precio sombra de la restricción del problema inicial. Si la conexión es compartida de manera óptima, este dato puede interpretarse como el precio que los usuarios estarían dispuestos a pagar por la utilización de una unidad adicional de capacidad de enlace, o el precio unitario por utilización efectiva ([8] Courcoubetis, 1999). De este modo, el precio por la utilización por unidad de tiempo para usuario con un contrato  $x_i$  será  $p \cdot a_i(x_i)$ , donde  $p$  comprende los factores económicos, tales como demanda, competencia, etc, mientras que  $a_i(x_i)$  comprende los factores operativos, es decir, recursos de red, calidad del servicio, etc.

Para el caso en que la conexión ofrece varios tipos (2) de fuentes, el problema se extiende de esta forma:

$$\max_{\{x_i, y_i\}} \sum_i U_i(x_i, y_i) \quad (17)$$

Sujeto a:

$$\sum_i a_1(x_i) \leq C_1^* \quad (18)$$

$$\sum_i a_2(x_i) \leq C_2^* \quad (19)$$

A nivel económico en la práctica, la ventaja del método de Ancho de banda efectivo es que ofrece la posibilidad de discriminar los precios para cada clase de fuente, asignándolos de forma acorde con la utilización real de los recursos. Si se tarifara

de acuerdo con una tasa media constante a todas las fuentes, la red se subutilizaría y congestionaría más rápidamente.

## 4.2 POSIBLES ESQUEMAS DE TARIFAS

Estos esquemas permiten limitar el tráfico de la fuente (tasa pico, capacidad) por medio de un contrato ISP-usuario, según el comportamiento de éste último:

- El primero permite asignar tarifas según el tipo de contrato: según el Ancho de banda efectivo de cada fuente de una clase, cobra proporcional a ese valor.
- El segundo consiste en cobrar solamente según las mediciones de tráfico realizadas, lo que aumenta su complejidad.
- El tercero consiste en comprometer en un solo mecanismo los parámetros estáticos (estipulados en el contrato) y los dinámicos (producto de las mediciones de tráfico y características de la red), permitiendo una interacción mutua entre proveedor y usuario.

En el tercer esquema, el operador calcula el Ancho de banda efectivo para cada tipo de fuente como función de los parámetros estáticos (fuentes de igual clase tiene el mismo contrato); el usuario le reporta el valor esperado de los parámetros (indica el punto de operación), con lo cual el operador calcula la función de cobro trazando la recta tangente a la función  $a(s,t)$ . Debido a la concavidad de esta función, el cobro a los usuarios se minimizará cuando su reporte coincide con su punto de operación, logrando un esquema adecuado y eficiente.

En primer lugar, es un esquema justo pues el cobro asignado a los usuarios es proporcional al uso que hacen de la red; y en segundo lugar, es un esquema que

incentiva a los usuarios a conocer su comportamiento de forma precisa, minimizando los costos y proveyendo al operador de información real, administrando mejor sus recursos.

## CONCLUSIONES Y RECOMENDACIONES

La limitante en común que tienen casi todos los modelos de fijación de precios – a excepción del modelo de Tarifa Plana que ya se aplica – es que no consideran, incluyen o planean como un trabajo futuro alguna aplicación del modelo en la práctica para de este modo poder comprobar sus resultados en una red de datos real, y así determinar su nivel de validez. Lo que busca este trabajo es identificar un modelo sobre el cual se ha generado mucha teoría y poderlo aplicar en la práctica, a través de tomas de datos y simulaciones con algoritmos, mostrando sus respectivos resultados y conclusiones.

Es muy importante resaltar que antes de plantear una aplicación práctica para la medición de recursos en Internet por medio del modelo de Ancho de banda efectivo, es necesario definir cuáles escenarios a nivel práctico (en una red de tráfico real) pueden ser modelados o simulados a través de la teoría expuesta, y cuáles no. Es decir, la teoría de Ancho de banda efectivo representa un nivel considerable de complejidad en su implementación y es probable que sea necesario ubicar el escenario real escogido en un contexto especial de tal forma que haya que hacer ciertas suposiciones, algunas de ellas muy fuertes incluso alejando la implementación práctica de lo que realmente sugiere el modelo teórico.

Las fallas en la obtención de muchos datos se debió a un problema en la aplicación de Diveo para realizar ese tipo de medición tan compleja. Tal vez tuvo que ver también con la estructuración y configuración interna de la red que pudo haber generado restricciones en el manejo de la información, quizá relacionados con

acuerdos de confidencialidad. En definitiva, esta aplicación fue diseñada para obtener análisis más globales, como el desempeño de la red medido directamente sobre todo el ancho de banda del canal de tráfico entre la Universidad y el proveedor, y no a granel, como se pretendía con este proceso.

En cierto modo se esperaba desde antes que ocurrieran problemas de este tipo, debido a la dificultad del proceso. Probablemente ésta no es la herramienta más adecuada para la toma de trazas de tráfico desde un punto de vista tan específico, pero quizá también puede ser que este tipo de medición es muy poco común (por la poca difusión que tiene el esquema en estudio) y de ahí que no existan muchos métodos de hacerlo de forma rápida y confiable, a menos que se tome un caso muy específico y no tan generalizado como el de la red de los Andes.

Se puede concluir que la metodología de usar la página de Internet de Diveo Uniandes es muy ineficiente si estos problemas siguen persistiendo de forma generalizada como ocurrió en el transcurso de este proyecto, como la visualización de los datos como se requieren. En ese sentido, aunque probablemente no sea una tarea fácil, sería mejor buscar una forma alterna para la medición de la información a granel, y en eso deberían concentrarse esfuerzos.

En este documento se ha logrado crear un modelo que permita medir la porción de red que ha sido usada realmente por un usuario bajo una aplicación en particular (y bajo ciertas condiciones también). De este modo, si se aplica en la realidad de forma correcta, le servirá al proveedor y operador de la red como la herramienta para crear un mecanismo que sirva para manejar sus recursos de manera óptima. A pesar de su alto nivel de complejidad, los miembros de la comunidad del Internet debería hacer un esfuerzo adicional en la implementación y ejecución de un mecanismo como éstos, que se verá reflejado, junto con la creación de un acuerdo de servicio de manera conjunta con el usuario, en la posibilidad de ofrecer la mejor tarifa, con la mejor calidad del servicio y bajos costos de operación, y de este modo maximizar sus utilidades.

Queda como la posibilidad de realizar en un trabajo futuro una medición más completa, en la que se pueda tener acceso a más datos reales sobre una red en particular y de este modo sacar conclusiones importantes sobre el desempeño del mecanismo y de la teoría en especial, y también en la que se logre estudiar el desempeño de redes cada vez más grandes y complejas. Una vez se obtengan estos datos se calculan los parámetros relevantes, como el Ancho de banda efectivo y así se puede usar en el modelo de precios del capítulo 4. De este modo se logrará cuantificar un contrato de servicio entre el proveedor y el usuario.

Por otro lado, sería muy interesante aplicar este mecanismo para redes en las que, a diferencia de las instituciones educativas y sin ánimo de lucro, sí se requiere fijar una tarifa en pesos para cobrar por el servicio. Esto obviamente aplica para los proveedores del acceso a la red que existen en el mercado, de manera que logren obtener una tarifa por el servicio, calculada por este método, y compararla con la que venían usando antes (como la tarifa plana) y observar los beneficios económicos a corto, mediano y largo plazo.



## REFERENCIAS BIBLIOGRÁFICAS

1. AHMAD, Waqas ; SANGMANEE, Sangboon. *Ex Post Internet Pricing: Contractual Pricing and Hurst Parameter Pricing*. Documento de Investigación. Telecommunications Project Class. M.S. in Telecommunications program. University of Maryland. September, 1999.
2. BAILEY, Joseph. ; NAGEL, José. ; RAGHAVAN, S. *Ex Post Internet Pricing*. Documento de investigación. University of Maryland. 1999.
3. BAILEY, Joseph. ; NAGEL, José. ; RAGHAVAN, S. *An Internet Ex Post Pricing Scheme based on the Effective bandwidths theory*. Documento de investigación. Telecommunications Project Class. M.S. in Telecommunications program. University of Maryland. August, 1999.
4. BELTRÁN, Fernando; GARCÍA, César. *Precios por congestión en redes de datos con prioridades basados en los precios Aumann-Shapley y el valor Shapley*. Documento del Centro de Estudios para la Gestión de Servicios en Redes, CGSR Departamento de Ingeniería Industrial. Universidad de los Andes. Bogotá, 2000.
5. BETANCOURT, Diego. *Dimensionamiento de Capacidad de canal de un enlace de Internet en redes congestionadas*. Tesis de Grado (Ingeniería Eléctrica). Universidad de los Andes, Bogotá, Mayo de 2002.

6. CLARK, David D. *A Model for Cost Allocation and Pricing in the Internet*. Documento de Investigación presentado en el MIT Workshop on Internet Economics. Marzo, 1995.
7. COURCOUBETIS, Costas; WEBER, Richard. *Buffer overflow asymptotics for a Buffer handling many traffic sources*. ICS-FORTH. 1996.
8. COURCOUBETIS, Costas; SIRIS, Vasilios; STAMOULIS, George. *Application of the many sources asymptotic and effective bandwidths to traffic engineering*. ICS-FORTH. Enero, 1999.
9. COURCOUBETIS, Costas; SIRIS, Vasilios. *Managing and Pricing Service Level Agreements for Differentiated Services*. ICS-FORTH. Mayo, 1999.
10. DELGADO, Juan Pablo. *¿Cómo establecer mejores esquemas de costos y precios del producto de acceso a Internet dedicado de AT&T Colombia, para utilizarlos en la definición de estrategias de negocio de la compañía?* Trabajo de Investigación. Especialización en Gerencia de Empresas de Telecomunicaciones. Universidad de los Andes, 2003.
11. GAVIRIA, Natalia. *Esquemas de tarificación en redes de banda Ancha*. Tesis de Maestría de Ingeniería Eléctrica. Departamento de Eléctrica. Universidad de los Andes, 2000.
12. GUPTA, A., STAHL, D., WHINSTON, A. *Priority Pricing of Integrated Services Networks*. En *Internet Economics*, Ed. por L. McKnight y J. Bailey. The MIT Press. 1998.
13. KELLY, Frank. *Charging and Accounting for Bursty Connections*. Documento de Investigación presentado en el MIT Workshop on Internet Economics. Marzo, 1995.

14. KELLY, Frank. *Notes on Effective bandwidths*. Documento de investigación. University of Cambridge, 1996.
15. MACKIE-MASON, J., VARIAN, H. *Pricing the Internet*. University of Michigan, 1993.
16. MACKIE-MASON, J., VARIAN, H. *Pricing Congestible Network Resources*. University of Michigan, 1994.
17. McLEAN, R.; SHARKEY, W. *An Approach to the Pricing of Broadband Telecommunications Services*. Telecommunications Systems, 1994.
18. SIRIS, Vasilios; SONGHURST, David; STAMOULIS, George. *Usage-based charging using effective bandwidth: studies and reality*. ICS-FORTH. Enero, 1999.
19. Página de Internet del Instituto ICS-FORTH (Departamento de Ciencias Computacionales de la Universidad de Creta. Heraklion, Grecia): <http://www.ics.forth.gr/netgroup/>.
20. Página de Internet de Diveo Uniandes (Aplicación para el análisis del tráfico del canal de red de los Andes): <http://noc.diveo.net.co/uniandes>

## LISTA DE ANEXOS

ANEXO 1. PROCEDENCIA Y SIGNIFICADO DE LA ECUACIÓN ASÍNTOTA DE MUCHAS FUENTES .....	68
ANEXO 2. TRAZAS DE TRÁFICO Y ALGORITMO PARA LA APLICACIÓN DE LA TEORÍA DE ANCHO DE BANDA EFECTIVO.....	72

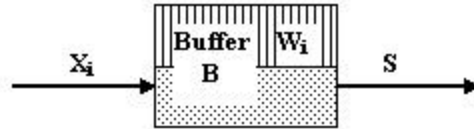
## **Anexo 1. Procedencia y significado de la ecuación asíntota de muchas fuentes**

En este Anexo se explica cómo se obtiene la ecuación de Asíntota de muchas fuentes, punto de partida para aplicar en la práctica un modelo de precios midiendo los recursos con precisión por medio del esquema de ancho de banda efectivo. Además, se le da un significado adicional a dicho resultado.

De acuerdo con Courcoubetis ([7], 1996), el análisis para comprender el significado de la Asíntota de muchas Fuentes parte de un modelo de buffer finito dentro de un *switch* ATM (conmutador). En este buffer se forma una cola con paquetes de tráfico de prioridad igual, a donde llegan por cada intervalo de tiempo  $i$  ( $i=1,2,3,\dots$ ) con una tasa de llegada  $X_i$ ; la tasa de servicio o de salida es  $S$ , y la carga al inicio del intervalo  $i$  es  $W_i$ . La carga de tráfico en el conmutador al inicio del intervalo  $i+1$  será:

$$W_{i+1} = \max\{0, \min[(W_i + X_i - S), B]\} \quad (1)$$

Es decir, las celdas se pierden cuando el buffer se llena y la carga al inicio del intervalo  $i+1$  será el tamaño del buffer,  $B$ . Por otro lado, en la ecuación (1) se evita por medio del proceso de maximización que  $W_{i+1}$  represente una longitud de cola negativa, lo cual sería imposible. La ecuación (1) puede representarse por medio de la Figura 1:



**Figura 1. Representación de un modelo de buffer finito. La parte sombreada con rayas corresponde a la carga dentro del buffer al inicio del intervalo i. Toda la parte sombreada (rayada y punteada) corresponde al buffer B.**

Sean  $b$  y  $c$  las cantidades de tamaño de buffer y capacidad utilizada por fuente, y  $N$  el número de fuentes. Por  $b$  tanto, el buffer disponible para todas las fuentes será  $B=Nb$ , y la capacidad utilizada globalmente es  $C=Nc$ . A partir de la representación anterior, la tasa de pérdida de paquetes puede expresarse como:

$$L(c,b,N) = E( \max [ W_i + X_i - S - B, 0 ] ) \quad (2)$$

Este valor está asociado directamente con una función  $Q$  que representa la probabilidad de que en un buffer la longitud de la cola de paquetes iguale el límite  $B=Nb$  y llene el buffer. Esta probabilidad equivaldrá así mismo a la probabilidad de sobrecarga de éste:

$$Q(Nc,Nb,N) = P(W_i=B) = P(\text{sobrecarga}) \quad (3)$$

Es importante recordar que el principio de la teoría de Grandes Desviaciones aplicable a redes de Banda Ancha es el comportamiento exponencial de la probabilidad de sobrecarga, que decrece asintóticamente hacia la recta  $P(\text{sobrecarga})=0$  a medida que otra variable del sistema aumenta. A partir de esto podemos sugerir el caso para cuando el número de fuentes  $N$  aumenta asintóticamente., así:

$$Q(Nc, Nb, N) = P(\text{sobrecarga}) \approx e^{[-N \cdot I(c,b) + g_1(c,b,N)]} \quad (4)$$

donde  $I(c,b)$  es alguna función, llamada asintótica, que por depender del búffer  $b$  y de la capacidad  $c$  será una aproximación clave para caracterizar el tráfico a través de la red por medio de los parámetros  $s,t$ . La función  $g_1$  es un término que completa el formato lineal del exponente de  $e$ . Sacando logaritmo a ambos lados de la igualdad (4) se obtiene:

$$\ln Q(Nc, Nb, N) = -N \cdot I(c, b) + g_1(c, b, N) \quad (5)$$

Dividiendo todo por  $N$ :

$$\frac{1}{N} \ln Q(Nc, Nb, N) = -I(c, b) + \frac{g_1(c, b, N)}{N} \quad (6)$$

Como el número de fuentes  $N$  aumenta de manera asintótica, se saca el límite a ambos lados de la ecuación cuando  $N \rightarrow \infty$ , con lo cual el segundo término a la derecha de la ecuación (6) se vuelve cero:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln Q(Nc, Nb, N) = -I(c, b) \quad (7)$$

De nuevo,  $I$  es la función asintótica, la cual está en función del buffer  $b$  y la capacidad  $c$  pero igualmente de los parámetros  $s,t$  que son el punto de operación de la red. Esta función se define de la siguiente forma:

$$-I = \sup_t \inf_s J(s, t) = \sup_t \inf_s \left[ s \cdot t \cdot \sum_{j=1}^J n_j \mathbf{a}_j(s, t) - s \cdot (B + C \cdot t) \right] \quad (8)$$

Courcoubetis ([7], 1996, págs. 5-6) plantea las ecuaciones (7) y (8) como parte de un teorema. La demostración de este teorema así como la obtención de la ecuación (8) se encuentran en detalle en dicho documento y su análisis no hace parte de este trabajo.

Ahora, por otro lado, sea  $n=(n_1, n_2, \dots, n_J)$  el conjunto de los números de fuentes multiplexadas, desde 1 hasta J tipos de fuentes. Si definimos a la función  $A(\gamma, C, b)$  como un subconjunto de los  $R_+^J$  que representa todas las soluciones a la ecuación de asíntota de muchas fuentes - donde  $\gamma$  es el parámetro de calidad en el servicio, C es la capacidad del canal y b el tamaño del buffer -, tal que  $n \in A(\gamma, C, b)$ , entonces implica el siguiente resultado, similar al de la ecuación (4) de este Anexo:

$$\log Q(Nc, Nb, N) = \log P(\text{sobrecarga}) \leq -\mathbf{g} \quad (9)$$

El límite en la ecuación (7) es estrictamente creciente en cada componente del vector n, así que se cumple:

$$\lim_{N \rightarrow \infty} \frac{A(NC, Nb, N\mathbf{g})}{N} = A_t \quad (10)$$

donde

$$A = \bigcap_{0 < t < \infty} A_t \quad (11)$$

y donde a su vez,

$$A_t = \left\{ n : \inf_s \left[ s \cdot t \cdot \sum_{j=1}^J n_j \mathbf{a}_j(s, t) - s \cdot (B + C \cdot t) \right] \leq -\mathbf{g} \right\} \quad (12)$$

Entonces, si la frontera de la region A es diferenciable en algún punto característico  $n^*$ , la ecuación del plano tangente a esa frontera en ese punto es:

$$\sum_{n=1}^J n_j \mathbf{a}_j(s^*, t^*) = C + \frac{B}{t^*} - \frac{\mathbf{g}}{s^* \cdot t^*} \quad (13)$$

donde la pareja  $(s^*, t^*)$  es la solución de (8) donde n ha sido reemplazado por  $n^*$ . La ecuación (13) corresponde por lo tanto a la restricción de la ecuación de ancho de banda efectivo (6) pero surgida a partir de un límite asintótico (tomado de [14] Kelly, 1996).



## Anexo 2. Trazas de tráfico y algoritmo para la aplicación de la teoría de ancho de banda efectivo

En este anexo se describe el procedimiento para la obtención del algoritmo que permite calcular los parámetros  $s, t$  y posteriormente el Ancho de banda efectivo  $\alpha$ , utilizando la ecuación sup inf especificada en la expresión (8) del **Anexo 1**. Además, se muestra una simulación de este algoritmo, comprobando su desempeño y verificando sus resultados.

El primer paso consiste en dejar la ecuación (8) que contiene el operador sup inf en términos de los parámetros  $s, t$ , y luego reemplazar la ecuación del estimativo del Ancho de banda efectivo (11) del documento. De esta forma, se puede aplicar el algoritmo y encontrar las parejas de valores  $(s, t)$  que dan solución a dicha ecuación, así:

$$\sup_t \inf_s J(s, t) = \sup_t \inf_s \left[ s \cdot t \cdot \sum_{j=1}^J n_j a_j(s, t) - s \cdot (B + C \cdot t) \right] = -I \quad (8)$$

Reemplazando la ecuación (11) del documento en la anterior, se tiene:

$$\sup_t \inf_s \left[ s \cdot t \cdot \sum_{j=1}^J n_j \left\{ \frac{1}{s \cdot t} \log \left( \frac{1}{T/t} \sum_{i=1}^{T/t} e^{s X_j[(i-1)t, it]} \right) \right\} - s \cdot (B + C \cdot t) \right] = -I \quad (14)$$

Cancelando  $s \cdot t$ ,

$$\sup_t \inf_s \left[ \sum_{j=1}^J \left\{ n_j \cdot \log \left( \frac{1}{T/t} \sum_{i=1}^{T/t} e^{s \cdot X_j[(i-1)t, it]} \right) \right\} - s \cdot (B + C \cdot t) \right] = -I \quad (15)$$

El procedimiento de maximización para la búsqueda de los valores s,t se hará por lo tanto sobre la ecuación (15). El algoritmo comprende los datos tomados de la red de la Universidad de los Andes y analizados para encontrar la utilización real de los recursos.

El algoritmo se divide en dos partes, y funciona de la siguiente manera:

\* En la primera parte se calculará el término entre corchetes cuadrados de la ecuación (15). Obsérvese que este término es calculable si se conocen todos los datos (número de fuentes por tipo de fuente (ver Tabla 4 del documento)), tiempo total de la traza (10 horas), intervalo de medición t (36 segundos), tamaño de buffer B y capacidad nominal, que se nombrarán más adelante). El único valor no conocido es el parámetro s y es el que se va a calcular con el segundo programa (algoritmo) en MATLAB, usando la función de búsqueda fminsearch. Por otro lado, la parte de este término entre corchetes que requiere de atención es:  $X_j[(i-1)t, it]$  o sea la carga de tráfico en el intervalo  $[(i-1)t, it]$  para una fuente j, lo que significa la creación de una pequeña rutina.

Primero se toman todos los datos dentro de una matriz de tamaño 1000 filas x 6 columnas (número de tipos de tráfico). Luego se divide en vectores asignando a cada columna un vector (correspondiente a tráfico de cada aplicación). Este valor  $X_j$  se obtiene precisamente de esos vectores: se necesita que el intervalo de medición de trazas (t) sea divisor entero del intervalo total de medición, T, que en este caso es 10 horas o 36000 segundos, para que este valor  $X_j$  se calcule sobre intervalos de igual tamaño abarcando todos los datos.

Es decir,  $T$  será igual a 36000, pero  $t$  deberá tomar valores de divisores enteros, como 36 (el mismo intervalo inicial de medición de trazas) en cuyo caso los  $X_j$  serán los mismos elementos del vector inicial (1000 elementos). Para  $t=36*2$  (el siguiente divisor entero), cada  $X_j$  será la suma por parejas, es decir, igual a  $(2^{i-1})+(2^i)$ , es decir, el primer elemento más el segundo, luego el tercero más el cuarto y así (este vector de los  $X_j$  tendrá entonces 500 elementos). Para  $t=36*4$ , cada  $X_j$  será la suma de cada 4 valores del vector grande y por tanto este vector  $X_j$  tendrá  $1000/4=250$  elementos. Para  $t=36*10$ , cada  $X_j$  será la suma de cada 10 valores del vector grande y por tanto este vector  $X_j$  tendrá  $1000/10=100$  elementos. Y así sucesivamente.

Lo anterior debe ser así también por el hecho de que en la ecuación (11) del documento la sumatoria va de  $i=1$  a  $T/t$ , y como se sabe los subíndices de una sumatoria deben ser enteros y por tanto  $T/t$  debe ser entero (positivo).

Así que a  $t$  se le dieron los siguientes valores:  $t=36, 72, 144, 180, 360, 720, 900, 1440, 1800, 3600$ . El algoritmo para calcular el término entre corchetes cuadrados de la ecuación (15) solamente se usa para hacer esos cálculos numéricos, en particular el de  $X_j$  que toma mayor código, pero no usa ninguna operación compleja en especial.

\* En la segunda parte del algoritmo se realizan los dos procesos de optimización (maximización y minimización). La primera parte corresponde a una función (llamada cálculo) a la cual se le entra el valor del parámetro  $s$  y calcula únicamente el término entre corchetes cuadrados de la ecuación (15). En la segunda parte esa función se llama dentro de otra función interna de MATLAB, llamada `fminsearch` la cual calcula el valor de  $s$  en el cual la función tiene un mínimo. Se usa de la siguiente forma

```
s=fminsearch('calculo',Xo)
```

La función comienza a buscar a partir de un valor inicial para  $s$  ( $X_0$ ) el cual es supuesto por el usuario y que para este caso es 0.05, y encuentra un mínimo local de la función especificada, usando el método simplex de búsqueda directa de Nelder-Mead.

Una vez se calcula el parámetro  $s$ , se procede a calcular  $t$  de la siguiente forma: simplemente lo que se hace es reemplazar este valor de  $s$  y los valores posibles de  $t$  (36, 72, etc) en la función (el término entre corchetes cuadrados), obteniendo tantos valores de dicha función como valores de  $t$  hemos generado inicialmente. Luego, observando estos valores se determina cuál de ellos es el máximo y así mismo se determina el  $t$  correspondiente, para el cual la función resultó ser máxima.

De este modo se obtiene una pareja de valores ( $s,t$ ) como posible solución a la ecuación sup inf. Se dice que es una posible solución porque no es la única; hay que recordar que la función de minimización encuentra un mínimo local y así como pudo haber uno alrededor de  $X_0=0.05$  también puede haber otro. Además, los valores de  $t$  fueron supuestos y quizá si se prueban con otros encontremos otros valores.

A continuación se muestran los códigos de los algoritmos en MATLAB que fueron usados. Los resultados se muestran en el numeral 3.3. del documento (Resultados y conclusiones).

Primer algoritmo:

```
function y=calculoFIN(s)

X=[0 1040 432 0 2828 25952 % Datos generados
0 1040 1069 0 3636 48478
....
0 0 7222 0 2828 52261
];
```

```
for i=1:1000,
    tipo1_t0(i)=8*X(i,1);
    tipo2_t0(i)=8*X(i,2);
    tipo3_t0(i)=8*X(i,3);
    tipo4_t0(i)=8*X(i,4);
    tipo5_t0(i)=8*X(i,5);
    tipo6_t0(i)=8*X(i,6);
end

B=50;    % Tamano del buffer, en kilobits
C=1000;  % Capacidad nominal del canal, en kilobits / s
t0=36;  % Intervalo de medición de trazas, en s
T=36000; % Tiempo total de medición, en s
fuentes1=49; % Numero de fuentes de tipo 1
fuentes2=103; % Numero de fuentes de tipo 2
fuentes3=40; % Numero de fuentes de tipo 3
fuentes4=9; % Numero de fuentes de tipo 4
fuentes5=83; % Numero de fuentes de tipo 5
fuentes6=2991; % Numero de fuentes de tipo 6

% s en 1/kilobits
```

```

t1=72;
for i=1:1000/2,
    tipo1_t1(i)=X(i*2-
1,1)+X(i*2,1);
    tipo2_t1(i)=X(i*2-
1,2)+X(i*2,2);
    tipo3_t1(i)=X(i*2-
1,3)+X(i*2,3);
    tipo4_t1(i)=X(i*2-
1,4)+X(i*2,4);
    tipo5_t1(i)=X(i*2-
1,5)+X(i*2,5);
    tipo6_t1(i)=X(i*2-
1,6)+X(i*2,6);
end

t2=144;
tipo1_t2=[];
tipo2_t2=[];
tipo3_t2=[];
tipo4_t2=[];
tipo5_t2=[];
tipo6_t2=[];
for i=1:1000/4,
    valor=[0 0 0 0 0 0];
    for j=1:4,

valor(1)=valor(1)+tipo1_t0(i*4-
(j-1));

valor(2)=valor(2)+tipo2_t0(i*4-
(j-1));

valor(3)=valor(3)+tipo3_t0(i*4-
(j-1));

valor(4)=valor(4)+tipo4_t0(i*4-
(j-1));

valor(5)=valor(5)+tipo5_t0(i*4-
(j-1));

valor(6)=valor(6)+tipo6_t0(i*4-
(j-1));
end
    tipo1_t2(i)=valor(1);
    tipo2_t2(i)=valor(2);
    tipo3_t2(i)=valor(3);
    tipo4_t2(i)=valor(4);
    tipo5_t2(i)=valor(5);
    tipo6_t2(i)=valor(6);
end

t3=180;
tipo1_t3=[];
tipo2_t3=[];
tipo3_t3=[];

```

```

tipo4_t3=[];
tipo5_t3=[];
tipo6_t3=[];
for i=1:1000/5,
    valor=[0 0 0 0 0 0];
    for j=1:5,

valor(1)=valor(1)+tipo1_t0(i*5-
(j-1));

valor(2)=valor(2)+tipo2_t0(i*5-
(j-1));

valor(3)=valor(3)+tipo3_t0(i*5-
(j-1));

valor(4)=valor(4)+tipo4_t0(i*5-
(j-1));

valor(5)=valor(5)+tipo5_t0(i*5-
(j-1));

valor(6)=valor(6)+tipo6_t0(i*5-
(j-1));
end
    tipo1_t3(i)=valor(1);
    tipo2_t3(i)=valor(2);
    tipo3_t3(i)=valor(3);
    tipo4_t3(i)=valor(4);
    tipo5_t3(i)=valor(5);
    tipo6_t3(i)=valor(6);
end

t4=360;
tipo1_t4=[];
tipo2_t4=[];
tipo3_t4=[];
tipo4_t4=[];
tipo5_t4=[];
tipo6_t4=[];
for i=1:1000/10,
    valor=[0 0 0 0 0 0];
    for j=1:10,

valor(1)=valor(1)+tipo1_t0(i*10-
(j-1));

valor(2)=valor(2)+tipo2_t0(i*10-
(j-1));

valor(3)=valor(3)+tipo3_t0(i*10-
(j-1));

valor(4)=valor(4)+tipo4_t0(i*10-
(j-1));

```

```

valor(5)=valor(5)+tipo5_t0(i*10-
(j-1));

valor(6)=valor(6)+tipo6_t0(i*10-
(j-1));
end
tipo1_t4(i)=valor(1);
tipo2_t4(i)=valor(2);
tipo3_t4(i)=valor(3);
tipo4_t4(i)=valor(4);
tipo5_t4(i)=valor(5);
tipo6_t4(i)=valor(6);
end

t5=720;
tipo1_t5=[];
tipo2_t5=[];
tipo3_t5=[];
tipo4_t5=[];
tipo5_t5=[];
tipo6_t5=[];
for i=1:1000/20,
    valor=[0 0 0 0 0 0];
    for j=1:20,

valor(1)=valor(1)+tipo1_t0(i*20-
(j-1));

valor(2)=valor(2)+tipo2_t0(i*20-
(j-1));

valor(3)=valor(3)+tipo3_t0(i*20-
(j-1));

valor(4)=valor(4)+tipo4_t0(i*20-
(j-1));

valor(5)=valor(5)+tipo5_t0(i*20-
(j-1));

valor(6)=valor(6)+tipo6_t0(i*20-
(j-1));
end
tipo1_t5(i)=valor(1);
tipo2_t5(i)=valor(2);
tipo3_t5(i)=valor(3);
tipo4_t5(i)=valor(4);
tipo5_t5(i)=valor(5);
tipo6_t5(i)=valor(6);
end

t6=900;
tipo1_t6=[];
tipo2_t6=[];
tipo3_t6=[];
tipo4_t6=[];

```

```

tipo5_t6=[];
tipo6_t6=[];
for i=1:1000/25,
    valor=[0 0 0 0 0 0];
    for j=1:25,

valor(1)=valor(1)+tipo1_t0(i*25-
(j-1));

valor(2)=valor(2)+tipo2_t0(i*25-
(j-1));

valor(3)=valor(3)+tipo3_t0(i*25-
(j-1));

valor(4)=valor(4)+tipo4_t0(i*25-
(j-1));

valor(5)=valor(5)+tipo5_t0(i*25-
(j-1));

valor(6)=valor(6)+tipo6_t0(i*25-
(j-1));
end
tipo1_t6(i)=valor(1);
tipo2_t6(i)=valor(2);
tipo3_t6(i)=valor(3);
tipo4_t6(i)=valor(4);
tipo5_t6(i)=valor(5);
tipo6_t6(i)=valor(6);
end

t7=1440;
tipo1_t7=[];
tipo2_t7=[];
tipo3_t7=[];
tipo4_t7=[];
tipo5_t7=[];
tipo6_t7=[];
for i=1:1000/40,
    valor=[0 0 0 0 0 0];
    for j=1:40,

valor(1)=valor(1)+tipo1_t0(i*40-
(j-1));

valor(2)=valor(2)+tipo2_t0(i*40-
(j-1));

valor(3)=valor(3)+tipo3_t0(i*40-
(j-1));

valor(4)=valor(4)+tipo4_t0(i*40-
(j-1));

valor(5)=valor(5)+tipo5_t0(i*40-
(j-1));

```

```

valor(6)=valor(6)+tipo6_t0(i*40-
(j-1));
end
tipo1_t7(i)=valor(1);
tipo2_t7(i)=valor(2);
tipo3_t7(i)=valor(3);
tipo4_t7(i)=valor(4);
tipo5_t7(i)=valor(5);
tipo6_t7(i)=valor(6);
end

t8=1800;
tipo1_t8=[];
tipo2_t8=[];
tipo3_t8=[];
tipo4_t8=[];
tipo5_t8=[];
tipo6_t8=[];
for i=1:1000/50,
    valor=[0 0 0 0 0 0];
    for j=1:50,

valor(1)=valor(1)+tipo1_t0(i*50-
(j-1));

valor(2)=valor(2)+tipo2_t0(i*50-
(j-1));

valor(3)=valor(3)+tipo3_t0(i*50-
(j-1));

valor(4)=valor(4)+tipo4_t0(i*50-
(j-1));

valor(5)=valor(5)+tipo5_t0(i*50-
(j-1));

valor(6)=valor(6)+tipo6_t0(i*50-
(j-1));
end
tipo1_t8(i)=valor(1);
tipo2_t8(i)=valor(2);

```

```

tipo3_t8(i)=valor(3);
tipo4_t8(i)=valor(4);
tipo5_t8(i)=valor(5);
tipo6_t8(i)=valor(6);
end

t9=3600;
tipo1_t9=[];
tipo2_t9=[];
tipo3_t9=[];
tipo4_t9=[];
tipo5_t9=[];
tipo6_t9=[];
for i=1:1000/100,
    valor=[0 0 0 0 0 0];
    for j=1:100,

valor(1)=valor(1)+tipo1_t0(i*100-
(j-1));

valor(2)=valor(2)+tipo2_t0(i*100-
(j-1));

valor(3)=valor(3)+tipo3_t0(i*100-
(j-1));

valor(4)=valor(4)+tipo4_t0(i*100-
(j-1));

valor(5)=valor(5)+tipo5_t0(i*100-
(j-1));

valor(6)=valor(6)+tipo6_t0(i*100-
(j-1));
end
tipo1_t9(i)=valor(1);
tipo2_t9(i)=valor(2);
tipo3_t9(i)=valor(3);
tipo4_t9(i)=valor(4);
tipo5_t9(i)=valor(5);
tipo6_t9(i)=valor(6);
end

```

```

if (t9/T)*sum(exp((s/1000)*tipo1_t9)) &
(t9/T)*sum(exp((s/1000)*tipo2_t9)) & (t9/T)*sum(exp((s/1000)*tipo3_t9))
~= 0,
%
y=fuentes1*log((t0/T)*sum(exp((s/1000)*tipo1_t0)))+fuentes2*log((t0/T)*su
m(exp((s/1000)*tipo2_t0)))+fuentes3*log((t0/T)*sum(exp((s/1000)*tipo3_t0)
))+fuentes4*log((t0/T)*sum(exp((s/1000)*tipo4_t0)))+fuentes5*log((t0/T)*s
um(exp((s/1000)*tipo5_t0)))+fuentes6*log((t0/T)*sum(exp((s/1000)*tipo6_t0
)))-s*(C*t0+B);
%
y=fuentes1*log((t1/T)*sum(exp((s/1000)*tipo1_t1)))+fuentes2*log((t1/T)*su
m(exp((s/1000)*tipo2_t1)))+fuentes3*log((t1/T)*sum(exp((s/1000)*tipo3_t1)

```



```

))+fuentes4*log((t1/T)*sum(exp((s/1000)*tipo4_t1)))+fuentes5*log((t1/T)*sum(exp((s/1000)*tipo5_t1)))+fuentes6*log((t1/T)*sum(exp((s/1000)*tipo6_t1)))-s*(C*t1+B);
%
y=fuentes1*log((t2/T)*sum(exp((s/1000)*tipo1_t2)))+fuentes2*log((t2/T)*sum(exp((s/1000)*tipo2_t2)))+fuentes3*log((t2/T)*sum(exp((s/1000)*tipo3_t2)))+fuentes4*log((t2/T)*sum(exp((s/1000)*tipo4_t2)))+fuentes5*log((t2/T)*sum(exp((s/1000)*tipo5_t2)))+fuentes6*log((t2/T)*sum(exp((s/1000)*tipo6_t2)))-s*(C*t2+B);
%
y=fuentes1*log((t3/T)*sum(exp((s/1000)*tipo1_t3)))+fuentes2*log((t3/T)*sum(exp((s/1000)*tipo2_t3)))+fuentes3*log((t3/T)*sum(exp((s/1000)*tipo3_t3)))+fuentes4*log((t3/T)*sum(exp((s/1000)*tipo4_t3)))+fuentes5*log((t3/T)*sum(exp((s/1000)*tipo5_t3)))+fuentes6*log((t3/T)*sum(exp((s/1000)*tipo6_t3)))-s*(C*t3+B);
%
y=fuentes1*log((t4/T)*sum(exp((s/1000)*tipo1_t4)))+fuentes2*log((t4/T)*sum(exp((s/1000)*tipo2_t4)))+fuentes3*log((t4/T)*sum(exp((s/1000)*tipo3_t4)))+fuentes4*log((t4/T)*sum(exp((s/1000)*tipo4_t4)))+fuentes5*log((t4/T)*sum(exp((s/1000)*tipo5_t4)))+fuentes6*log((t4/T)*sum(exp((s/1000)*tipo6_t4)))-s*(C*t4+B);
%
y=fuentes1*log((t5/T)*sum(exp((s/1000)*tipo1_t5)))+fuentes2*log((t5/T)*sum(exp((s/1000)*tipo2_t5)))+fuentes3*log((t5/T)*sum(exp((s/1000)*tipo3_t5)))+fuentes4*log((t5/T)*sum(exp((s/1000)*tipo4_t5)))+fuentes5*log((t5/T)*sum(exp((s/1000)*tipo5_t5)))+fuentes6*log((t5/T)*sum(exp((s/1000)*tipo6_t5)))-s*(C*t5+B);
%
y=fuentes1*log((t6/T)*sum(exp((s/1000)*tipo1_t6)))+fuentes2*log((t6/T)*sum(exp((s/1000)*tipo2_t6)))+fuentes3*log((t6/T)*sum(exp((s/1000)*tipo3_t6)))+fuentes4*log((t6/T)*sum(exp((s/1000)*tipo4_t6)))+fuentes5*log((t6/T)*sum(exp((s/1000)*tipo5_t6)))+fuentes6*log((t6/T)*sum(exp((s/1000)*tipo6_t6)))-s*(C*t6+B);
%
y=fuentes1*log((t7/T)*sum(exp((s/1000)*tipo1_t7)))+fuentes2*log((t7/T)*sum(exp((s/1000)*tipo2_t7)))+fuentes3*log((t7/T)*sum(exp((s/1000)*tipo3_t7)))+fuentes4*log((t7/T)*sum(exp((s/1000)*tipo4_t7)))+fuentes5*log((t7/T)*sum(exp((s/1000)*tipo5_t7)))+fuentes6*log((t7/T)*sum(exp((s/1000)*tipo6_t7)))-s*(C*t7+B);
%
y=fuentes1*log((t8/T)*sum(exp((s/1000)*tipo1_t8)))+fuentes2*log((t8/T)*sum(exp((s/1000)*tipo2_t8)))+fuentes3*log((t8/T)*sum(exp((s/1000)*tipo3_t8)))+fuentes4*log((t8/T)*sum(exp((s/1000)*tipo4_t8)))+fuentes5*log((t8/T)*sum(exp((s/1000)*tipo5_t8)))+fuentes6*log((t8/T)*sum(exp((s/1000)*tipo6_t8)))-s*(C*t8+B);
y=fuentes1*log((t9/T)*sum(exp((s/1000)*tipo1_t9)))+fuentes2*log((t9/T)*sum(exp((s/1000)*tipo2_t9)))+fuentes3*log((t9/T)*sum(exp((s/1000)*tipo3_t9)))+fuentes4*log((t9/T)*sum(exp((s/1000)*tipo4_t9)))+fuentes5*log((t9/T)*sum(exp((s/1000)*tipo5_t9)))+fuentes6*log((t9/T)*sum(exp((s/1000)*tipo6_t9)))-s*(C*t9+B);
%
y=fuentes1*log((t10/T)*sum(exp((s/1000)*tipo1_t10)))+fuentes2*log((t10/T)*sum(exp((s/1000)*tipo2_t10)))+fuentes3*log((t10/T)*sum(exp((s/1000)*tipo3_t10)))+fuentes4*log((t10/T)*sum(exp((s/1000)*tipo4_t10)))+fuentes5*log((t10/T)*sum(exp((s/1000)*tipo5_t10)))+fuentes6*log((t10/T)*sum(exp((s/1000)*tipo6_t10)))-s*(C*t10+B);
end

```

## Segundo algoritmo:

La primera parte hasta el calculo de los vectores  $X_j$  (hasta  $t_9=3600$ ) es exactamente igual en ambos algoritmos, así que a continuación no se repite.

```
% yfinalFIN

X=[0 1040 432 0 2828 25952 % Datos generados
0 1040 1069 0 3636 48478
....
0 0 7222 0 2828 52261
];

tipo1_t0=[];
tipo2_t0=[];
tipo3_t0=[];
tipo4_t0=[];
tipo5_t0=[];
tipo6_t0=[];

for i=1:1000,
    tipo1_t0(i)=8*X(i,1);
    tipo2_t0(i)=8*X(i,2);
    tipo3_t0(i)=8*X(i,3);
    tipo4_t0(i)=8*X(i,4);
    tipo5_t0(i)=8*X(i,5);
    tipo6_t0(i)=8*X(i,6);
end

B=50; % Tamano del buffer, en kilobits
C=1000; % Capacidad nominal del canal, en kilobits / s
t0=36; % Intervalo de medición de trazas, en s
T=36000; % Tiempo total de medición, en s
fuentes1=49; % Numero de fuentes de tipo 1
fuentes2=103; % Numero de fuentes de tipo 2
fuentes3=40; % Numero de fuentes de tipo 3
fuentes4=9; % Numero de fuentes de tipo 4
fuentes5=83; % Numero de fuentes de tipo 5
fuentes6=2991; % Numero de fuentes de tipo 6

Xo=0.05;

s=fminsearch('calculoFIN',Xo);
disp(s);
```

(IGUAL QUE EN EL ALGORITMO ANTERIOR, DESDE  $t_1$  HASTA  $t_9$ )

```
y0=fuentes1*log((t0/T)*sum(exp((s/1000)*tipo1_t0)))+fuentes2*log((t0/T)*sum(exp((s/1000)*tipo2_t0)))+fuentes3*log((t0/T)*sum(exp((s/1000)*tipo3_t0)))+fuentes4*log((t0/T)*sum(exp((s/1000)*tipo4_t0)))+fuentes5*log((t0/T)*sum(exp((s/1000)*tipo5_t0)))+fuentes6*log((t0/T)*sum(exp((s/1000)*tipo6_t0)))-s*(C*t0+B);
```

```
y1=fuentes1*log((t1/T)*sum(exp((s/1000)*tipo1_t1)))+fuentes2*log((t1/T)*sum(exp((s/1000)*tipo2_t1)))+fuentes3*log((t1/T)*sum(exp((s/1000)*tipo3_t1)))+fuentes4*log((t1/T)*sum(exp((s/1000)*tipo4_t1)))+fuentes5*log((t1/T)*sum(exp((s/1000)*tipo5_t1)))+fuentes6*log((t1/T)*sum(exp((s/1000)*tipo6_t1)))-s*(C*t1+B);
```

```
y2=fuentes1*log((t2/T)*sum(exp((s/1000)*tipo1_t2)))+fuentes2*log((t2/T)*sum(exp((s/1000)*tipo2_t2)))+fuentes3*log((t2/T)*sum(exp((s/1000)*tipo3_t2)))+fuentes4*log((t2/T)*sum(exp((s/1000)*tipo4_t2)))+fuentes5*log((t2/T)*sum(exp((s/1000)*tipo5_t2)))+fuentes6*log((t2/T)*sum(exp((s/1000)*tipo6_t2)))-s*(C*t2+B);
```

```
y3=fuentes1*log((t3/T)*sum(exp((s/1000)*tipo1_t3)))+fuentes2*log((t3/T)*sum(exp((s/1000)*tipo2_t3)))+fuentes3*log((t3/T)*sum(exp((s/1000)*tipo3_t3)))+fuentes4*log((t3/T)*sum(exp((s/1000)*tipo4_t3)))+fuentes5*log((t3/T)*sum(exp((s/1000)*tipo5_t3)))+fuentes6*log((t3/T)*sum(exp((s/1000)*tipo6_t3)))-s*(C*t3+B);
```

```
y4=fuentes1*log((t4/T)*sum(exp((s/1000)*tipo1_t4)))+fuentes2*log((t4/T)*sum(exp((s/1000)*tipo2_t4)))+fuentes3*log((t4/T)*sum(exp((s/1000)*tipo3_t4)))+fuentes4*log((t4/T)*sum(exp((s/1000)*tipo4_t4)))+fuentes5*log((t4/T)*sum(exp((s/1000)*tipo5_t4)))+fuentes6*log((t4/T)*sum(exp((s/1000)*tipo6_t4)))-s*(C*t4+B);
```

```
y5=fuentes1*log((t5/T)*sum(exp((s/1000)*tipo1_t5)))+fuentes2*log((t5/T)*sum(exp((s/1000)*tipo2_t5)))+fuentes3*log((t5/T)*sum(exp((s/1000)*tipo3_t5)))+fuentes4*log((t5/T)*sum(exp((s/1000)*tipo4_t5)))+fuentes5*log((t5/T)*sum(exp((s/1000)*tipo5_t5)))+fuentes6*log((t5/T)*sum(exp((s/1000)*tipo6_t5)))-s*(C*t5+B);
```

```
y6=fuentes1*log((t6/T)*sum(exp((s/1000)*tipo1_t6)))+fuentes2*log((t6/T)*sum(exp((s/1000)*tipo2_t6)))+fuentes3*log((t6/T)*sum(exp((s/1000)*tipo3_t6)))+fuentes4*log((t6/T)*sum(exp((s/1000)*tipo4_t6)))+fuentes5*log((t6/T)*sum(exp((s/1000)*tipo5_t6)))+fuentes6*log((t6/T)*sum(exp((s/1000)*tipo6_t6)))-s*(C*t6+B);
```

```
y7=fuentes1*log((t7/T)*sum(exp((s/1000)*tipo1_t7)))+fuentes2*log((t7/T)*sum(exp((s/1000)*tipo2_t7)))+fuentes3*log((t7/T)*sum(exp((s/1000)*tipo3_t7)))+fuentes4*log((t7/T)*sum(exp((s/1000)*tipo4_t7)))+fuentes5*log((t7/T)*sum(exp((s/1000)*tipo5_t7)))+fuentes6*log((t7/T)*sum(exp((s/1000)*tipo6_t7)))-s*(C*t7+B);
```

```
y8=fuentes1*log((t8/T)*sum(exp((s/1000)*tipo1_t8)))+fuentes2*log((t8/T)*sum(exp((s/1000)*tipo2_t8)))+fuentes3*log((t8/T)*sum(exp((s/1000)*tipo3_t8)))+fuentes4*log((t8/T)*sum(exp((s/1000)*tipo4_t8)))+fuentes5*log((t8/T)*sum(exp((s/1000)*tipo5_t8)))+fuentes6*log((t8/T)*sum(exp((s/1000)*tipo6_t8)))-s*(C*t8+B);
```

```
y9=fuentes1*log((t9/T)*sum(exp((s/1000)*tipo1_t9)))+fuentes2*log((t9/T)*sum(exp((s/1000)*tipo2_t9)))+fuentes3*log((t9/T)*sum(exp((s/1000)*tipo3_t9)))+fuentes4*log((t9/T)*sum(exp((s/1000)*tipo4_t9)))+fuentes5*log((t9/T)*sum(exp((s/1000)*tipo5_t9)))+fuentes6*log((t9/T)*sum(exp((s/1000)*tipo6_t9)))-s*(C*t9+B);
```

```
yfinal=[y0 y1 y2 y3 y4 y5 y6 y7 y8 y9];
```

```
disp(yfinal);
```