

## Caracterización y simulación del tráfico de redes LAN mediante el modelo MMPP

### Characterization and simulation of the LAN traffic by MMPP model

*Jaime Alberto Moreno Mogollón\*, Jhon Jairo Padilla Aguilar, Vladimir Escobar Ordóñez, Andrés Fernando Correo Montesino*

Universidad Pontificia Bolivariana seccional Bucaramanga, km 7 Autopista a Piedecuesta, Colombia.

(Recibido el 15 de enero de 2007. Aceptado el 10 de agosto de 2007)

#### Resumen

En este artículo se presenta un método para caracterizar y simular el tráfico de acceso a Internet de cualquier red LAN con comportamiento autosimilar. Para ello se emplea el modelo de cola D-MMPP/D/1 para generar una traza discreta y analizar su comportamiento al pasar por la cola con diferentes capacidades de *buffer*. Este método ha sido probado por otros autores en condiciones controladas de laboratorio y solo para tráfico entrante. Nuestro estudio demuestra que esta metodología puede ser utilizada en redes LAN bajo condiciones reales y actuales tanto para tráfico autosimilar entrante como saliente. Para tal fin se utiliza el tráfico capturado en los enlaces de entrada/salida de Internet de la Red LAN de la Universidad Pontificia Bolivariana, seccional Bucaramanga. Además, esta metodología soluciona el problema de modelar tráfico autosimilar ya que genera una traza simulada aproximada a la traza medida. La exactitud del modelo se ve reflejada en la gráfica de los quantiles (Q-Q plot) entre la traza medida y la traza generada mediante D-MMPP. Para realizar nuestro estudio, el modelo y los algoritmos fueron validados inicialmente utilizando trazas obtenidas en los laboratorios Bellcore. Posteriormente, se realizaron algunos ajustes para tener una traza simulada en términos de bits por segundo y no en términos de paquetes por segundo como se hace en otros trabajos, con el fin de obtener una simulación de la capacidad del canal en bits por segundo.

----- *Palabras clave:* tráfico autosimilar, caracterización, *hurst*, modelamiento, MMPP, paquetes, *quantiles*, red, tráfico, tramas.

---

\* Autor de correspondencia: teléfono: 54+6+679 62 20 Ext. 438, fax: 54+6+679 62 21, correo electrónico: jmorenomo@upbbga.edu.co (J.A. Moreno).

**Abstract**

This paper presents a method to characterize and simulate the traffic of any Local Area Network (LAN) which displays a self-similar behavior. We used the queuing model D-MMPP/D/1 to generate a discrete stream and to analyze its behavior when passing through the queue using different buffer capacities. This method has been tested by other authors under controlled-laboratory conditions for incoming traffic only. We demonstrate that this methodology can be used for LAN networks under real and present conditions for incoming and outgoing traffic. We used the traffic captured on Internet input/output links from Universidad Pontificia Bolivariana Seccional Bucaramanga's LAN. Additionally, this methodology solves the problem of modeling self-similar traffic due to its capacity of generating a simulated stream approximated to the measured stream. The exactitude of the model is shown in the quantile (Q-Q plot) between the measured stream and the one generated by D-MMPP. In our study, the model and algorithms were initially validated using streams provided by Bellcore laboratories. Furthermore, we performed some adjustments in order to obtain a simulated stream in terms of bits per second (bps) as opposed to packets per second in order to perform a link's capacity simulation in bps.

----- *Key words:* Self-similar traffic, Characterization, Hurst, Modeling, MMPP, Packets, Quantiles, Network, Traffic, Frames.

## Introducción

Se ha demostrado recientemente que el tráfico de Internet posee características de autosimilaridad y dependencia a largo plazo. De allí se puede concluir que no se puede modelar tráfico IP mediante Poisson como se realizaba en telefonía, sino que se deben utilizar procedimientos más complejos ya que la característica principal de Poisson es la de ser un proceso sin memoria en donde no depende del pasado. El modelo de tráfico IP utilizado en el presente trabajo es un *Proceso de llegadas de Poisson modulado por cadenas de Markov (MMPP)* [1, 2].

Se han realizado algunos estudios de modelamiento de tráfico y de red mediante este modelo, que indican que el modelo MMPP es capaz de producir tráfico autosimilar [3]. Los algoritmos utilizados en el trabajo son los propuestos por David Lucantoni [1] para generar la traza virtual de tráfico. Antes de caracterizar los datos, es necesario verificar que el tráfico tenga características de autosimilaridad; para esto se estima el valor del parámetro de Hurst ( $H$ ) y se observa si  $H > 0,5$ . Existen diversos métodos para la estimación del parámetro de Hurst [4]; los más usados son varianza-tiempo, R/S plot, estimador Whittle y transformada Wavelet, entre otros. El método utilizado en este proyecto es el método varianza-tiempo en el cual se obtiene un gráfico logarítmico-logarítmico a partir de una serie de datos  $X(t)$  que generan diferentes niveles de agregación  $m$  y luego se calcula la varianza; el resultado es una línea recta con pendiente  $-\beta$ . Valores de pendiente entre  $-1$  y  $0$  sugieren autosimilaridad.

Con el desarrollo del presente trabajo, se verifica que se puede modelar tráfico IP en una red LAN bajo condiciones reales, mediante el modelo MMPP. La herramienta de simulación que se desarrolló contiene una interfaz gráfica que permite que cualquier usuario acceda a ella y pueda generar una traza virtual a partir de cualquier traza real con características de autosimilaridad y además permite realizar pruebas de bondad para la traza virtual obtenida.

Este trabajo está estructurado de la siguiente forma: en la sección 2 se hace una revisión teórica de conceptos básicos del modelado de redes y sistemas de colas enfocado al modelo MMPP y al tráfico autosimilar, en la sección 3 se dan las especificaciones generales de la herramienta implementada para las simulaciones, en la sección 4 se presentan los resultados y se explica cómo se recolectaron y procesaron los datos de tráfico de este estudio, en la sección 5 se presentan las conclusiones obtenidas a partir de los resultados de la caracterización y simulación del tráfico de la red de datos de la UPB, Bucaramanga (RDUPBB) [5].

## Marco teórico

### **Modelamiento de tráfico autosimilar o fractal**

Un modelo de tráfico relativamente actual tiene sus raíces en el modelamiento de fenómenos físicos. En los 70, Benoit Mandelbrot y sus colegas coincidieron en el término autosimilar que describe fenómenos donde una característica particular es similar en diferentes escalas de espacio o tiempo. Un ejemplo de modelos autosimilares en la dimensión espacial, es el de modelos de aviones pequeños para experimentos de túneles de viento. El comportamiento de los modelos más pequeños es similar al del aeroplano real y mucho más barato para fines de experimentación. La autosimilitud ocurre en fenómenos geométricos naturales como se plantea Mandelbrot. Comúnmente, este tipo de propiedad es también llamado *fractal*.

En procesos aleatorios de tráfico de datos, el fenómeno de autosimilitud ocurre en la dimensión del tiempo. Esto tiene consecuencias significantes en la estadística del tráfico. Un parámetro usado para caracterizar el grado de autosimilitud es llamado Hurst (en honor a H. E. Hurst).

A través de una imagen de una traza de tráfico que pueda ampliarse para reducir la granularidad, se podrían visualizar los patrones de tráfico autosimilar como lo han reportado muchos ar-

tículos en la literatura científica [6, 7]. Se debe tener en cuenta que a diferencia de los objetos fractales que tienen una autosimilitud determinística, el tráfico en una red Ethernet tiene una autosimilitud estocástica. Por esta razón cada gráfica no es una reproducción exacta de la gráfica anterior, pero guardan similitud en gran parte. Esto es, si se considera que el tráfico observado es una traza muestral de un proceso estocástico y se restringe la similitud a ciertas estadísticas de las series de tiempo reescalizadas, se descubre fácilmente autosimilitud exacta en los objetos matemático-abstractos y autosimilitud aproximada en las realizaciones específicas [4].

La presentación de este tipo de comparación empezó cuando investigadores analizaron trazas de tráfico de datos de LAN recolectados por Bellcore en los 90 [8].

Estudios realizados recientemente en [9, 10] coinciden en afirmar que el análisis del tráfico de salida a Internet de paquetes en redes LAN presenta estadísticamente un comportamiento autosimilar (fractal). El tráfico autosimilar se caracteriza por la presencia permanente de ráfagas a través de diferentes escalas de tiempo. Las propiedades que caracterizan los procesos autosimilares son dos: la dependencia a corto plazo (SRD: Short Range Dependence) y la dependencia a largo plazo (LRD: Long Range Dependence) [6], las cuales se hacen presentes al aumentar el nivel de agregación.

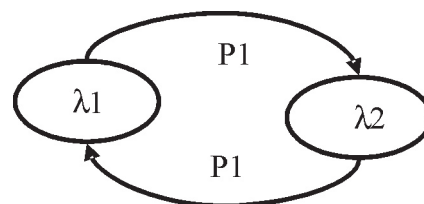
El comportamiento fractal de este tipo de tráfico no coincide con el comportamiento del tráfico modelado tradicionalmente por casi 100 años, el cual corresponde a procesos de Poisson con la característica de la ausencia de ráfagas y la baja variabilidad reflejada en la independencia de las muestras. Estos procesos exhiben una dependencia a corto plazo (SRD).

Estas propiedades se hacen presentes en el tráfico de datos y es posible caracterizarlas mediante estimación del parámetro Hurst ( $0,5 < H < 1$ ), el cual indica el grado de autosimilaridad.

### **Procesos de llegada de Poisson modulado por cadenas de Markov (MMPP)**

Uno de los modelos modulados por Markov es el MMPP [11]. Dicho modelo combina la simplicidad de la conmutación markoviana con la de la generación de Poisson lo que lo hace un atractivo modelo para el tráfico de paquetes en backbones. El mecanismo de modulación estipula que en el estado  $k$  de  $m$  estados la llegada ocurre de acuerdo con un proceso de Poisson con tasa  $\lambda_k$  y a medida que el estado cambia, también lo hace la tasa  $\lambda$  respectiva.

El ejemplo más simple de un modelo MMPP consta de dos estados, el primero es un estado de “encendido” con una tasa de Poisson asociada y el segundo es un estado de “apagado” con una tasa cero asociada. El modelo se muestra en la figura 1. Este modelo se ha usado ampliamente para fuentes de tráfico de voz donde el estado “encendido” corresponde al momento en que la fuente habla y el estado “apagado”, cuando no lo hace. Este modelo básico de MMPP puede ser extendido para representar el tráfico de fuentes independientes multiplexadas debido a la flexibilidad de Poisson para caracterizar el agregado de procesos estadísticamente independientes.



**Figura 1** Modelo sencillo MMPP

Como se observa en la figura 1, el modelo consta de dos estados en los cuales la generación sigue procesos de Poisson de tasas medias  $\lambda_1$  y  $\lambda_2$ . Por su parte, la permanencia en cada estado de la cadena de Markov, se rige por la matriz de transición de probabilidades característica de esta, y

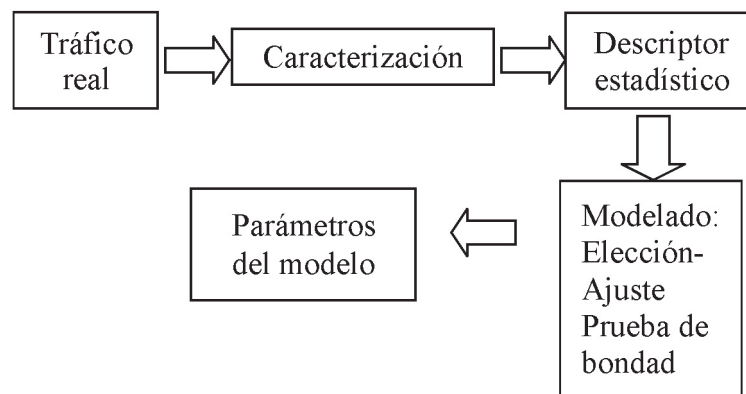
que es definida en función de las probabilidades  $P_1$  y  $P_2$ .

Con el descubrimiento de que el tráfico de Internet exhibe dependencia a largo plazo (LRD), se podría pensar el porqué se propone en este trabajo, así como en Heyman, el usar un MMPP como modelo de tráfico IP, que, por definición, es dependiente a corto plazo (SRD). Aunque el tráfico con características SRD tiene una cola exponencial, la parte exponencial de la cola podría llegar a ser dominante pero más allá de la región de interés para el modelamiento del desempeño. Ejemplos reales se pueden ver en Choudhury, en donde un MMPP presenta una cola pesada hasta más allá del punto donde la probabilidad de bloqueo es menor que  $10^{-40}$ . Mas allá de este punto, la componente exponencial de esta cola es dominante, pero desde que el modelo es aplicado a tráfico de ATM, únicamente se ha estado interesado en regiones donde la probabilidad de bloqueo sea mayor a  $10^{-9}$ . Existen algunos trabajos recientes como los de Andersen y Salvador en donde se modela tráfico IP mediante MMPP tal que el ambiente LRD se ajusta sobre muchos órdenes de magnitud. Además, el estado del arte acerca de medidas de rendimiento com-

putacional con modelos MMPP, va más allá del actual modelamiento de LRD, como se puede ver en Park y Willinger.

### Metodología utilizada para el modelamiento

Tal como se ilustra en la figura 2, el diseño de un modelo exige dos etapas: primero, el ajuste de cierto proceso estadístico que suele denominarse caracterización; segundo: seleccionar (en este caso el modelo MMPP) y ajustar el proceso estadístico en cuestión. La caracterización supone analizar la realidad que se va a imitar (el tráfico) con el objeto de describirla mediante un conjunto reducido de variables (descriptor estadístico) que resumen y abstraen sus aspectos más importantes. Tras la definición de este descriptor, que, en ocasiones puede actuar directamente como un modelo en sí, la tarea de modelado se centra en elegir un proceso matemático que aproxime dicho descriptor, ya con independencia de la realidad física detallada que resume. El objetivo de esta fase, tras decidir qué modelo se emplea, es una parametrización adecuada del mismo que asegure un ajuste óptimo del descriptor.



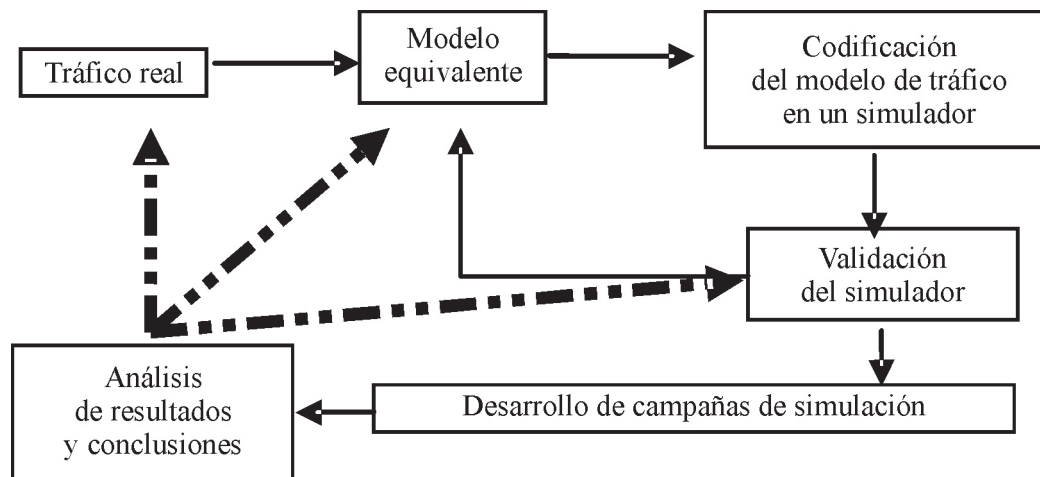
**Figura 2** Proceso de diseño de un modelo

De este modo, la definición del modelo MMPP puede ser múltiple en función de la caracterización seguida.

Después de caracterizar los datos (tráfico), debe realizarse el proceso de simulación mostrado en la figura 3, que consiste en realizar una abstrac-

ción de un modelo real (ambiente real de trazas de tráfico) y luego representarlo por medio de un modelo equivalente, para obtener información

relacionada con el modelo original. Esto con el fin de realizar pruebas que no sería posible realizar directamente sobre el modelo real de tráfico.



**Figura 3** Proceso de simulación

La metodología usada en este trabajo se puede resumir como sigue: 1) Análisis del comportamiento semanal del tráfico en la red que se va a analizar con el fin de establecer el tiempo de mayor tráfico. 2) Captura de trazas de tráfico en el canal de salida de Internet que se desea analizar en el tiempo determinado en el paso 1. 3) Procesamiento de los datos de tal forma que se obtengan las variables tiempo y tamaño de cada trama capturada. 4) Determinación de la autosimilitud de la traza de tráfico dentro de las restricciones dadas por el parámetro Hurst. 5) Caracterización del tráfico para estimar los parámetros propios del modelo MMPP como matriz de probabilidades del proceso de Markov, vector estacionario y vector de lambda (tasas de generación de paquetes de los estados del proceso de Markov). 6) Generación de la traza virtual. 7) Evaluación de la exactitud de la traza virtual mediante la prueba de bondad de Q-Q. 8) Simulación del comportamiento del canal de Internet haciendo pasar la fuente de tráfico MMPP por una cola MMPP/D/1 para evaluar el comportamiento del canal de Internet

frente a cambios en la capacidad dada por el proveedor de servicios. Esto permite analizar la probabilidad de pérdida y tiempo de espera para *buffer* infinito y finito cuando se hace variar la utilización del canal. Además, se muestra la aplicación de este método sobre una red real y congestionada como lo es la RDUPBB (Red de datos de la Universidad Pontificia Bolivariana, seccional Bucaramanga).

A continuación se describen brevemente los conceptos generales más relevantes utilizados en el modelamiento del tráfico de datos mediante MMPP/D/1.

### **Gráfica quantil-quantil (Q-Q plot)**

A menudo se desea asignar algunos números a una distribución de frecuencia para tener mayor facilidad al realizar una comparación con otras distribuciones. El método más directo es usar los quantiles. Los quantiles son valores que dividen la distribución, tal que halla una proporción dada



de observaciones debajo del quantil. Por ejemplo, la mediana es un quantil. La mediana es el valor central de la distribución, a partir de esa mitad los puntos son menos o se igualan a ella y la mitad es mayor que o igual a ella. Se puede estimar cualquier quantil fácilmente de la distribución de frecuencia acumulativa.

En general, Q-Q plot es lo que su nombre sugiere: una gráfica de los quantiles de los datos contra los quantiles teóricos correspondientes de la distribución especificada por el modelo estadístico. Esta idea es muy utilizada para el caso especial de verificar el ajuste de datos a un modelo.

### ***Estimación del parámetro Hurst***

Se han desarrollado numerosos métodos para determinar si una serie dada es autosimilar y para estimar el parámetro H de autosimilaridad. Aquí se resumen algunos de los métodos más comúnmente utilizados:

Gráfica varianza-tiempo. La varianza de los agregados describe la propiedad de decaimiento lento de la varianza de los agregados del tráfico autosimilar [6, 10].

Estimaciones R/S. Describe la propiedad de dependencia a largo plazo, mediante un estudio de autocorrelaciones del tráfico agregado [12].

Índice de dispersión de conteos. Su uso más común es el análisis de la variabilidad de las muestras en tiempo [9].

Estimador Wavelet. Brinda la estimación del parámetro Hurst realizando un análisis en frecuencia y tiempo simultáneamente, siendo óptimo en comparación con las herramientas de Laplace y Fourier, los cuales solo realizan análisis en frecuencia.

Por último, para caracterizar el tráfico en redes de datos, es necesario conocer la variación de la demanda durante largos intervalos de tiempo (perfil de tráfico), por ejemplo un día. Esto permite determinar el lapso de tiempo en el cual el tráfico presente baja variabilidad pero sea máxima la de-

manda para que la estimación sea de una alta confiabilidad y lograr el correcto dimensionamiento o determinación de rendimiento de los componentes de la red que se va a estudiar [9, 13].

### ***Modelo de colas para los componentes de red***

La teoría de colas es la principal herramienta para el análisis y predicción de una red de telecomunicaciones. Consiste en una serie de usuarios (en este caso paquetes) llegando en tiempos aleatorios a un lugar donde reciben algún tipo de servicio y luego parten. Los usuarios representan transacciones (flujo de información que va de un lado a otro).

El análisis del comportamiento de los componentes de la red para la capa física es independiente del tipo de funciones que cumpla cada uno de ellos en las demás capas. Por esta razón, con un solo modelo se logra representar matemáticamente el comportamiento de todos los elementos de la red, considerando cada uno de ellos como un nodo [14].

En el presente trabajo se utiliza la cola MMPP/D/1 para modelar tanto el canal de entrada como de salida de Internet. Para el canal de salida de Internet se incluye al router de borde dentro del modelo, el cual tiene una entrada y una salida, es decir el tiempo del proceso de enrutamiento es casi despreciable. También se utiliza la cola MMPP/D/1 para representar matemáticamente el canal de entrada de Internet que incluye el comportamiento de retardo o latencia dentro de la Red. Lo anterior se debe a que los datos fueron tomados entre la red y el router de borde, quedando claro que los datos son de LAN (red de área local) y no de WAN (red de área amplia).

De varios tipos de modelos matemáticos estudiados para modelar tráfico se ha escogido MMPP/D/1. Este modelo es un tipo de cola que tiene una fuente de tipo MMPP, un solo servidor y el tiempo de atención o de servicio es determinístico. El tipo de cola MMPP/D/1 es utilizada por Herman y Lucantoni, en donde se ha demostrado su eficacia a la hora de modelar el tráfico en el canal

de entrada de Internet de un red LAN; claro está que se ha realizado en un ambiente de laboratorio. Ahora bien, en este proyecto se utiliza el modelo pero se prueba en una red real con todas sus perturbaciones y problemas. La cola MMPP/D/1 es un tipo especial de la cola BMAP/D/1 [15]; como se aprecia en su nomenclatura la diferencia está en la fuente de tráfico. La fuente de tráfico MMPP (proceso de Poisson modulado por Markov) es un proceso de no renovación a diferencia de los procesos Erlang e hiperexponenciales que son de renovación. BMAP (*batch markovian arrival process*) incluye los dos anteriores. Ahora, el modelo de la cola BMAP/G/1, matemáticamente es una generalización de matriz simple de las colas M/G/1, M/D/1, BMAP/D/1 y MMPP/D/1. Se decide trabajar con MMPP/D/1 debido a que se asume que todos los paquetes o tramas tienen un solo tamaño igual al promedio de los paquetes o tramas capturadas. Por esta razón el tiempo de servicio será una constante de tipo determinístico. La cola MMPP/G/1 es más general e incluye la anterior, pero requeriría modelar los tamaños de los paquetes presentes en cada traza de tráfico que se analice, para determinar la distribución de probabilidad de los tiempos de servicio. Si bien la cola MMPP/D/1 cae dentro de la cola general MMPP/G/1, esta estructura general requerirá el cálculo y el almacenamiento de una gran cantidad de matrices las cuales estarán contenidas en la matriz de transición de probabilidades de la cadena de Markov embebida en transiciones.

Es bien sabido que los modelos básicos de colas M/G/1 y GI/M/1 pueden ser analizados por medio de cadenas de Markov embebidas. Estas son llamadas cadenas de Markov tipo M/G/1 y tipo GI/M/1. En este proyecto se trabajan las cadenas de Markov para obtener la fuente MMPP. Específicamente se tiene una cadena de Markov que contiene varios procesos de Poisson distribuidos en sus estados. Es bien sabido en el ámbito del modelamiento de tráfico de redes de datos que existen formas de generar tráfico autosimilar como movimiento browniano y otras. Sin embargo, en este trabajo se genera MMPP por que se ajusta a la necesidad del proyecto (medir el tráfico a la salida de Internet).

En Kleinrock se muestran los parámetros que caracterizan el comportamiento de las colas:

- La tasa de arribo  $\lambda$ , número de usuarios activos, número de servidores, tamaño de la cola (en el caso que la cola sea finita), tasa de servicio de la cola  $\mu$ .
- Intensidad del tráfico. Esta variable relaciona la tasa de llegadas con la tasa de servicio. También es llamado el “factor de utilización” de la cola.
- *Throughput*. Tasa promedio de atención de usuarios en el sistema. Esta variable mide la proporción con que son atendidos los usuarios en el sistema.
- Longitud del sistema. Es la longitud promedio de una cola e incluye todos los usuarios que están en el sistema, tanto los que están en la cola como los que están recibiendo el servicio.
- Tiempo de retardo. Es el tiempo total que gasta un usuario en el sistema. Este corresponde al tiempo que el usuario está en la cola más el tiempo de servicio.

### **Especificaciones generales del software desarrollado**

Los algoritmos desarrollados para el análisis del tráfico mediante el modelo MMPP/D/1, fueron implementados bajo la plataforma Matlab versión 7.1. El software consiste en una GUI (interfaz de usuario gráfica) con su respectivo manual de usuario la cual permite caracterizar y simular trazas de tráfico de cualquier canal de Internet de cualquier red LAN. Esta herramienta permite analizar el comportamiento de los paquetes (probabilidad de pérdida y tiempo de espera promedio de los paquetes) frente al cambio en la capacidad del canal, el cual generalmente contiene un elemento activo de red como un enrutador de borde. Se debe apreciar que el enrutador de borde hace parte del canal en este modelo. El manual contiene la función de cada uno de los componentes de la interfaz del *software*. En la figura 4 se muestra la interfaz del usuario.



**SOFTWARE PARA ANALISIS DE TRAFICO DE UNA TRAZA CUALQUIERA**

DEBES DAR CLICK EN 'HELP' PARA OBSERVAR EL PROCEDIMIENTO PARA USAR EL SOFTWARE

DEBES DAR CLICK EN 'SELECCIONE EL ARCHIVO' PARA CARGAR EL ARCHIVO QUE CONTIENE EL TRAFICO

**PARAMETROS DE ENTRADA**

INGRESE EL INTERVALO DE TIEMPO EN (SEG) CON EL CUAL DESEA DISCRETIZAR LA TRAZA

INGRESE EL TAMAÑO DEL BUFER

INGRESE EL AUMENTO DEL TRAFICO

INGRESE LA CAPACIDAD EN (bps) A LA QUE DESEA LIMITAR LA TRAZA

**HELP**

SELECCIONE EL ARCHIVO

**-RESULTADOS-**

TASA DE TRANSFERENCIA REAL (bps)

TASA DE TRANSFERENCIA VIRTUAL (bps)

Promedio paquetes de llegada por unidad de tiempo (REAL)

Promedio paquetes de llegada por unidad de tiempo (Virtual)

PARAMETRO HURST

TASA DE TRANSFERENCIA PARA EL TRAFICO CON CAPACIDAD LIMITADA DEL CANAL

**-BYTES PROMEDIO DE LA TRAZA-**

Este boton permite observar el promedio de los bytes de llegada de la traza seleccionada

**CALCULAR BYTES**

Bytes promedio de la traza

Total paquetes de llegada de la traza seleccionada

**Análisis con trafico sin limite en la Capacidad**

CALCULAR TRAZA REAL EN (bps)

CALCULAR TRAZA VIRTUAL EN (bps)

CALCULAR TRAZA REAL Y VIRTUAL EN (paquetes / s)

CALCULAR EL PARAMETRO HURST

CALCULAR TIEMPO DE ESPERA PROMEDIO Y PROBABILIDAD DE BLOQUEO

QUANTILES TRAFICO EN (bps)

**Análisis de trafico con limite en la capacidad y variacion en el trafico**

CALCULAR TRAZA VIRTUAL CON LIMITE EN LA CAPACIDAD

CALCULAR TIEMPO DE ESPERA PROMEDIO Y PROBABILIDAD DE BLOQUEO

**INICIALIZAR**

**SALIR**

**Figura 4** Interfaz del usuario

## Resultados obtenidos

Con el fin de validar los algoritmos propuestos por Lucantoni en [1], se simula el tráfico para la traza BC\_pAug89 de Bellcore Labs tomada de [2] y los resultados obtenidos se comparan con los resultados obtenidos por Lucantoni en [1] y con algunos trabajos de análisis de tráfico realizados con esta traza [16, 17] y otros. Las simulaciones de tráfico que se han desarrollado utilizando los algoritmos propuestos por Lucantoni [1], apuntan y coinciden en tener en cuenta la cantidad de paquetes que transitan por la red por unidad de tiempo ( $\Delta t$ ). Esta unidad de tiempo se debe escoger teniendo en cuenta la aproximación que se produzca en el modelo. Para el caso de este trabajo, el delta de tiempo que ajusta mejor la traza es de 500 ms. Para esto se realizaron pruebas con varios valores de intervalos de análisis.

### *Traza BC\_pAug89 de Bellcore Labs*

#### *Características de la Red Ethernet Bellcore [2]*

Cada una de las trazas disponibles en la web para análisis, contiene un millón de paquetes de llegada

tomados en la red Ethernet de “Bellcore Morristown Research and Engineering facility”. Dos de estas trazas son LAN (con una pequeña porción de tráfico WAN) y las otras dos son tráfico WAN.

Las trazas se encuentran en dos columnas, en formato ASCII, veinte *bytes* por línea. La primera línea muestra el tiempo en segundos desde el comienzo de la traza. La segunda columna muestra el tamaño del paquete en *bytes*, incluyendo parte del encabezado y el código de redundancia cíclica (CRC) como se observa en la figura 5, además se debe notar que con el protocolo Ethernet los paquetes poseen un tamaño mínimo de 64 *bytes* y un máximo de 1.518 *bytes*.

La traza Bc-pAug89 comienza a las 11:25 el 29 de agosto de 1989, capturó datos alrededor de 3.142,82 s (aproximadamente 1.000.000 paquetes fueron capturados). El *hardware* de captura no dejó perder ningún paquete, pero los paquetes parásitos (por ejemplo, colisiones de Ethernet) no fueron incluidos en la muestra.

El 99,5% de los paquetes capturados eran IP (protocolo de Internet). El trazado se hizo en la

Preámbulo	Delimitador de inicio de trama	Dirección MAC origen	Dirección MAC destino	Tipo	Datos	CRC
7 bytes	1 byte	6 bytes	6 bytes	2 bytes	46 @ 1.500 bytes	4 bytes

64 @ 1.518 bytes

**Figura 5** Características de las tramas capturadas en laboratorios Bellcore

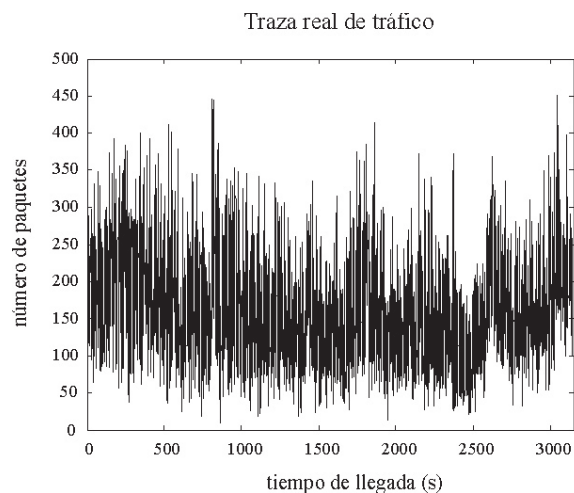
Bellcore Morristown Research and Engineering Facility. Mientras los tiempos capturados son reportados a 6 cifras decimales, estos tiempos tienen 4  $\mu$ s de precisión, y el análisis extenso indica que la exactitud real es aproximadamente 10  $\mu$ s (principalmente debido al tráfico presente en la red).

**Generación y análisis de la traza de tráfico virtual mediante MMPP (paquetes por unidad de tiempo)**

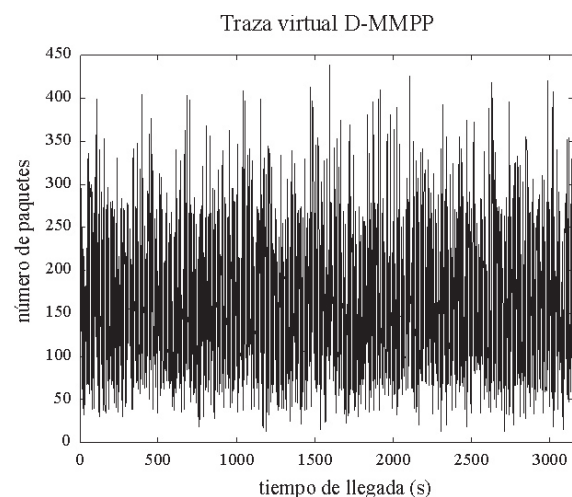
Esta traza es popular en el ámbito del modelamiento de tráfico en redes de datos. Aquí se grafica la traza real con un  $\Delta t$  de 500 ms, que permite contar el número de paquetes que se capturaron cada 50 ms. También se grafica la traza generada a través del modelo MMPP presentando resultados semejantes a los antecedentes de esta traza, lo cual valida el correcto funcionamiento de la herramienta de simulación. En las figuras 6 y 7 se muestra el tráfico original y el tráfico generado mediante el modelo MMPP respectivamente.

La aproximación del modelo MMPP para representar las características de la traza real, se mide por medio de la gráfica de quantiles-quantiles entre la traza original y la traza generada mediante el modelo MMPP.

En la figura 8 se muestra la aproximación del modelo para representar las características del tráfico real capturado en los laboratorios de Bellcore, por medio de la gráfica de Q-Q entre la traza original y la traza generada mediante MMPP. El criterio

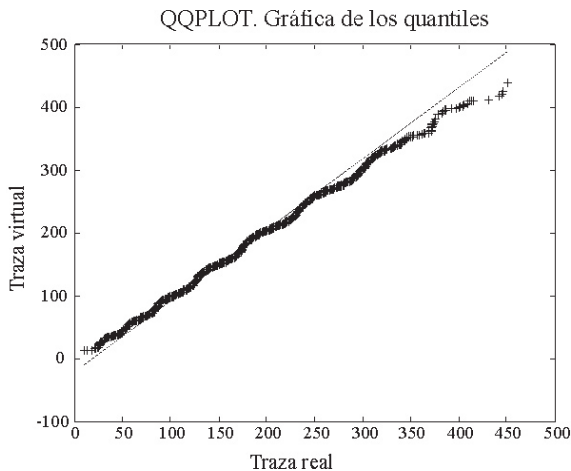


**Figura 6** Traza de tráfico original con datos capturados en Bellcore Labs, tomando un  $\Delta t = 500$  ms



**Figura 7** Traza de tráfico generada mediante MMPP, tomando un  $\Delta t = 500$  ms

de la prueba de bondad realizada a través de Q-Q se basa en que el modelo tendrá una mayor aproximación a la realidad, mientras menor sea el número de desviaciones de los datos de la gráfica de los quantiles, con respecto a la línea recta.



**Figura 8** Quantiles-quantiles para el tráfico de la traza BC\_pAug89, con  $\Delta t = 500$  ms

Con la obtención de un resultado satisfactorio de la gráfica de quantiles para la traza BC\_pAug89 de Bellcore Labs, se validan los algoritmos propuestos en [1] por Lucantoni; con esto se procede a realizar el análisis de tráfico para la red de la RDUPB seccional Bucaramanga y comprobar el ajuste del modelo para las diferentes trazas de los dos canales de acceso a Internet.

### Trazas de la RDUPBB

El modelo se aplica sobre la red de la Universidad Pontificia Bolivariana, seccional Bucaramanga.

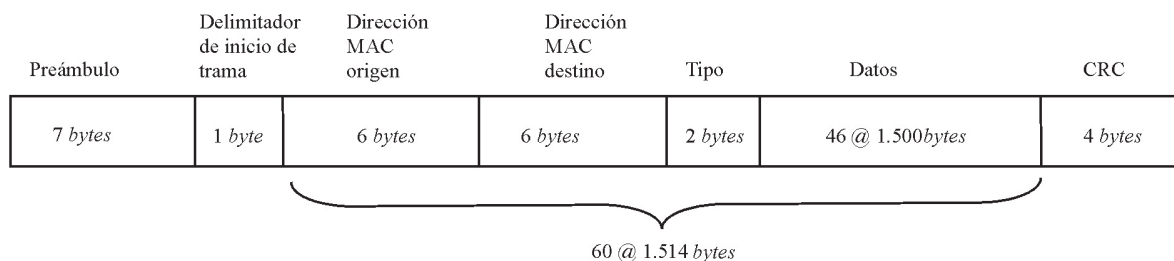
#### Características de la red Ethernet de la Universidad Pontificia Bolivariana, seccional Bucaramanga

La RDUPBB tiene una topología lógica en estrella y una topología física de estrella extendida. Está dividida en VLAN, las cuales son administradas por un *switch* capa tres ubicado en el (*core*) corazón de la red. Las VLAN se crearon por departamentos y la red se encuentra totalmente segmentada.

La topología física de la RDUPBB es una estrella extendida, completamente segmentada por *switches*.

La topología lógica está separada por VLAN que son controladas por un *switch* de capa tres (*core*) ubicado en el centro de la red.

Las trazas se encuentran en dos columnas, en formato ASCII. La primera columna muestra el tiempo de llegada por trama desde el inicio de la traza, la segunda columna muestra el tamaño de la trama en bytes, aquí no se incluye el código de redundancia cíclica (CRC) como se observa en la figura 9. Además, nótese que las tramas poseen un tamaño mínimo de 60 bytes y un máximo de 1.514 bytes.



**Figura 9** Características de las tramas capturadas en la UPB Bucaramanga

Las trazas de tráfico de la UPB Bucaramanga que se analizaron poseen un promedio de 200 s (aproximadamente 38.000 tramas fueron capturadas).

El *hardware* para capturar datos dejó perder un número pequeño de paquetes, llamados frecuentemente como ‘paquetes parásitos’, los cuales fueron incluidos en la muestra.

La RDUPBB posee dos canales que brindan servicio de Internet a todos los usuarios de la red: canal 1 y canal 2. Estos son suministrados por dos proveedores de servicio de Internet distintos, que en este trabajo por razones de confidencialidad se les ha llamado proveedor 1 y proveedor 2 respectivamente. El proveedor 1 brinda el servicio de Internet del canal 1 y el proveedor 2 brinda el servicio de Internet del canal 2. Ambos proveedores entregan una capacidad de 2 Mbps. El servicio prestado por el proveedor 1 es utilizado por los estudiantes, mientras que el servicio prestado por el proveedor 2 es aprovechado por el personal administrativo y docente.

Por el canal del proveedor 1 salen y entran de Internet solo los paquetes que provienen de la VLAN de estudiantes y llegan a la VLAN de estudiantes. A esta VLAN pertenecen todos los computadores en los que los alumnos pueden acceder a Internet. Por el canal del proveedor 2 salen y entran de Internet los paquetes provenientes de todas la VLAN, excepto de la VLAN de estudiantes.

A la fecha de la toma de datos (marzo de 2006) el canal 1 tenía el servicio de proxy en la salida como dispositivo de borde. El canal 2 tenía el servicio del proxy en un servidor dentro de la RDUPBB. En el momento de hacer comparaciones entre el rendimiento de los dos canales debe tenerse cuenta que el hecho de tener el servicio de proxy ubicado en sitios físicos distintos para los dos canales involucra que los dos canales tengan que ser evaluados de forma independiente.

### **Captura de datos de de la RDUPBB**

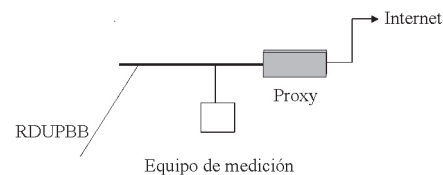
En general, en el estudio preliminar de la red se pudo evidenciar a través del MRTG que existe

mayor intensidad de tráfico en el canal 1 que en el canal 2. Esto debido a las aplicaciones utilizadas por los estudiantes. Por tanto la utilización de canal es más alta en el canal 1. Sin embargo, como se verá más adelante, se escogieron los momentos de mayor tráfico en donde el canal 2 presentó un pico alto de tráfico. Esto se hace para garantizar que el estudio se realice en condiciones extremas de tráfico.

### **Proveedor 1**

Las trazas de tráfico de entrada, desde Internet hacia la RDUPBB, se capturan en el nodo en el que se encuentra la interfaz privada del proxy que es la salida a Internet por este canal.

En la figura 10 se puede observar la ubicación del equipo de medición en el nodo del proveedor 1.



**Figura 10** Ubicación física del equipo de medición en la RDUPBB-canal 1

Se capturaron los datos de tráfico de entrada desde Internet hacia la RDUPBB. Durante el rango de tiempo de captura de datos, únicamente se encontraba el equipo de medición (figura 15) tomando tráfico en la RDUPBB.

En la tabla 1 se muestran el número de *bytes* enviados y recibidos con su respectivo *throughput* (velocidad de transmisión) para la captura de datos en el canal 1 que se realizó en las horas de la tarde (4:00:34 p. m.) del día viernes 19 de mayo del año 2006.

### **Proveedor 2**

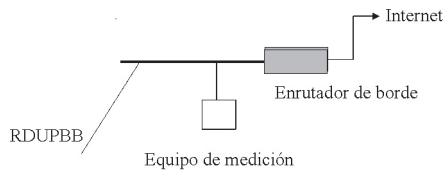
Las trazas de tráfico de entrada, desde Internet hacia la RDUPBB, se capturan en el nodo en el

**Tabla 1** Datos capturados en el canal 1

<i>Muestra de tráfico</i>	<i>Kb de entrada</i>	<i>Kb salida</i>	<i>Tiempo de captura</i>	<i>Velocidad de entrada</i>	<i>Velocidad de salida</i>
106	27.190	23.853	162,2s	1,3 Mbps	1,17 Mbps

que se encuentra la interfaz privada del enrutador de borde y por este canal es la salida a Internet.

En la figura 11 se puede observar la ubicación del equipo de medición para capturar las trazas de tráfico en el canal del proveedor 2.

**Figura 11** Ubicación física del equipo de medición en la RDUPBB-canal 2

Se capturaron los datos de tráfico de entrada de tramas desde Internet hacia la red RDUPB, allí se encontraba el equipo de medición tomando datos en el nodo del proveedor 2 y al mismo tiempo se encontraba tomando datos una aplicación instalada en el proxy de ese canal, el cual se encuentra dentro de la RDUPBB.

Es importante resaltar que cuando se toman los datos simultáneamente en dos puntos como es este caso, que se toman los datos del tráfico en el

nodo de entrada y salida de Internet y al mismo tiempo en el proxy dispuesto para el proveedor 2, se incrementa el uso de la CPU (unidad central de proceso) del equipo proxy, encargado de la captura de tráfico interno en la red. Esto hace que se disminuya el rendimiento tanto en la función del proxy como en la captura de los datos.

La traza de tráfico que se analizó corresponde a los datos de entrada desde Internet hacia la RDUPBB. Primero es importante mostrar la tasa promedio de llegada de tramas por unidad de tiempo para esta traza de tráfico, la cual posee un valor de 165,2549 (tramas por unidad de tiempo).

En la tabla 2 se muestran el número de *bytes* enviados y recibidos con su respectivo *throughput* (velocidad de transmisión) para la captura de datos en el canal 2 que se realizó en las horas de la tarde (12:49:32 p.m.) del día martes 30 de mayo del año 2006.

Las gráficas de tráfico original y generado mediante MMPP, están dadas en bits por unidad de tiempo. La unidad de tiempo es 500 ms.

Se estudia el comportamiento de las trazas de tráfico en los dos canales de la RDUPBB. Se

**Tabla 2** Datos capturados en el canal 2

<i>Muestra de tráfico</i>	<i>Kb de salida</i>	<i>Kb de entrada</i>	<i>Tiempo de captura</i>	<i>Velocidad de entrada</i>	<i>Velocidad de salida</i>
95	1.117,6	26.007,5	219,6 s	947,2 kbps	407,06 kbps

utilizan trazas de tráfico de un tiempo de captura promedio de 200 s.

### ***Generación de traza de tráfico virtual mediante MMPP***

#### ***Proveedor 1***

La traza original fue procesada y discretizada tomando un intervalo de tiempo ( $\Delta t$ ) de 500 ms. Al igual que la traza original, la traza virtual fue generada tomando un intervalo de tiempo ( $\Delta t$ ) de 500 ms. Esto debido a que fue el intervalo de tiempo con el que se logró un mejor ajuste en el modelo.

En el estudio se trabajó tanto el tráfico de entrada como de salida. En este artículo se muestran solamente los resultados para el tráfico de entrada ya

que, como se estableció en el estudio preliminar de los datos, el tráfico de entrada es el más representativo de la RDUPBB a largo plazo.

Al igual que para el canal 2, después de varias pruebas se llegó a la conclusión de que el modelo para la traza de tráfico correspondiente al canal 1, también se ajusta mejor tomando una unidad de tiempo ( $\Delta t$ ) de 500 ms.

En las tablas 3 y 4 se muestra el tráfico de entrada/salida real y simulado (generado por MMPP). En ellas se puede apreciar que para la muestra de tráfico escogida, el tráfico de entrada es más representativo tanto en la traza real como la virtual. Además se muestran los parámetros más relevantes que comparan la tasa de transferencia de la traza real con la virtual, presentando valores similares.

**Tabla 3** Tasas de transferencia y promedio de bits de llegada

<b><i>Parámetro</i></b>	<b><i>Valor</i></b>
Lambda promedio virtual	233,0989 (tramas por unidad de tiempo)
Tasa de transferencia virtual	1,3294 e + 006 (bits por unidad de tiempo)
Lambda promedio real	234,7143 (tramas por unidad de tiempo)
Tasa de transferencia real	1,3386 e + 006 (bits por unidad de tiempo)
Promedio de bits de llegada	5703,1037 (bits)

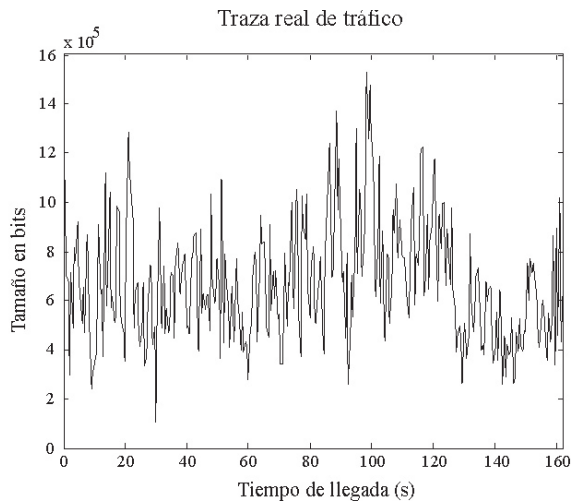
**Tabla 4** Tasas de transferencia y promedio de bits de salida

<b><i>Parámetro</i></b>	<b><i>Valor</i></b>
Lambda promedio virtual	228,6719 (tramas por unidad de tiempo)
Tasa de transferencia virtual	1,1756 e + 006 (bits por unidad de tiempo)
Lambda promedio real	228,3914 (tramas por unidad de tiempo)
Tasa de transferencia real	1,1741 e + 006 (bits por unidad de tiempo)
Promedio de los bits de llegada	5,1409 e + 003 (bits)

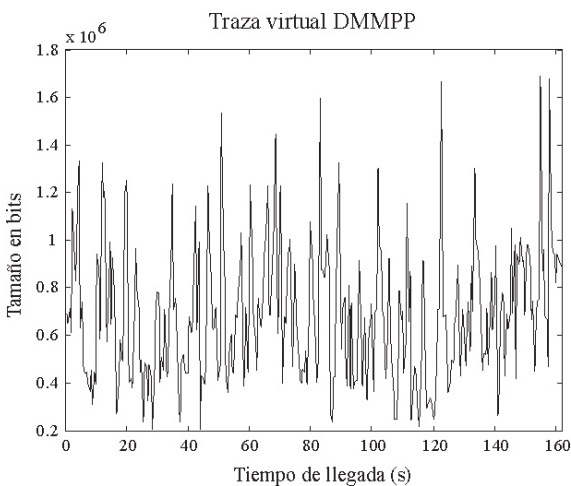


Como un criterio de validación de los datos mostrados en las tablas 3 y 4 se pueden comparar las tasas de transferencia real medida en un analizador de tráfico como el Ethereal [18] y el Dice [19] mostradas en la tabla 1.

La traza de tráfico original y la traza generada mediante el modelo MMPP, son mostradas en las figuras 12 y 13 respectivamente.



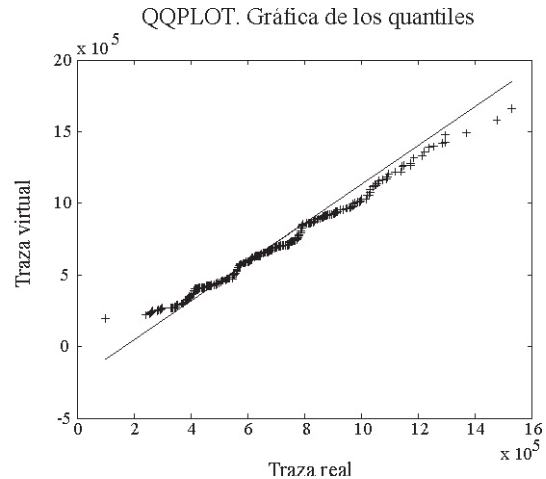
**Figura 12** Traza de tráfico original tomando un  $\Delta t = 500$  ms



**Figura 13** Traza de tráfico generada mediante MMPP tomando un  $\Delta t = 500$  ms

Se puede observar la aproximación del modelo para esta traza de tráfico, por medio de la grá-

fica de los cuantiles entre la traza original y la generada mediante MMPP, como se muestra en la figura 14.



**Figura 14** Gráfica de cuantiles (Q-Q plot) tomando una unidad de tiempo  $\Delta t = 500$  ms.

En la figura 14 se observa que el modelo posee una gran aproximación para esta traza de tráfico, esto se debe a que la subred con proveedor 1 posee un número grande de usuarios dentro de la RDUPBB, lo que hace que el tráfico tenga gran variabilidad en el tiempo y con esto se obtiene mayor cantidad de estados en la cadena de Markov.

### Proveedor 2

Para este canal también se analiza el tráfico de entrada para tener un criterio de comparación con el otro canal de Internet, aunque en el estudio se trabajó tanto el tráfico de entrada como de salida de la RDUPBB. Además, en el ámbito del diseño de redes es primordial la tasa de entrada frente a la tasa de salida, ya que un usuario siempre descargará más información que la que cargará hacia Internet. Este es el principio de funcionamiento de las redes cliente/servidor.

En las tablas 5 y 6 se muestra el tráfico de entrada/salida real y simulado (generado por MMPP). En ellas se puede apreciar que para la muestra de tráfico escogida, el tráfico de entrada es más representativo tanto en la traza real como la

virtual. Además se muestran los parámetros más relevantes que comparan la tasa de transferencia de la traza real con la virtual, presentando valores similares. Aquí, al igual que para el canal 1, los

datos obtenidos en el procesamiento, caracterización y simulación mostrados en las tablas 5 y 6 se pueden comparar con los resultados arrojados por el analizador de tráfico mostrados en la tabla 2.

**Tabla 5** Tasas de transferencia y bits promedio de llegada

<i>Parámetro</i>	<i>Valor</i>
Lambda promedio virtual	165,2549 (tramas por unidad de tiempo)
Tasa de transferencia virtual	9,7798 e + 005 (bits por unidad de tiempo)
Lambda promedio real	159,9359 (tramas por unidad de tiempo)
Tasa de transferencia real	9,4651e+005 (bits por unidad de tiempo)
Promedio de bits de llegada	5,9180 e + 003 (bits)

**Tabla 6** Tasas de transferencia y bits promedio de salida

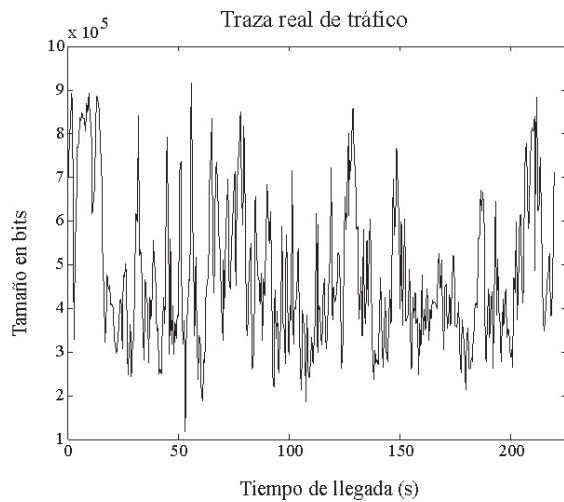
<i>Parámetro</i>	<i>Valor</i>
Lambda promedio virtual	168,8048 (tramas por unidad de tiempo)
Tasa de transferencia virtual	4,1141 e + 005 (bits por unidad de tiempo)
Lambda promedio real	166,8879 (tramas por unidad de tiempo)
Tasa de transferencia real	4,0674 e + 005 (bits por unidad de tiempo)
Promedio de bits de llegada	2,4372 e + 003 (bits)

La traza de tráfico original (real) es mostrada en la figura 15, mientras la traza de tráfico generada mediante MMPP es mostrada en la figura 16. Gráficamente se puede apreciar la semejanza de las dos trazas, corroborando los resultados mostrados en la tabla 5.

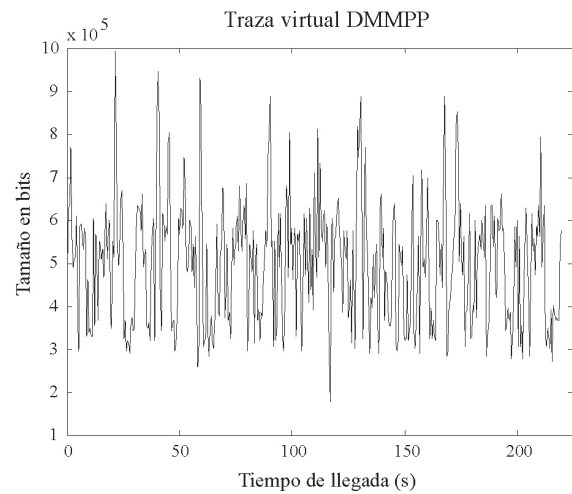
La aproximación del modelo es medida mediante la prueba de bondad de la gráfica de los cuantiles entre la traza original y la traza de tráfico generada mediante el modelo MMPP. En la figura 17 se muestra la aproximación del modelo para esta traza de tráfico. Se debe partir de la premisa de que el modelo se ajustará mejor, entre más cerca

estén los puntos de la gráfica de los cuantiles de la recta con pendiente positiva que especifica una correlación exacta de los datos.

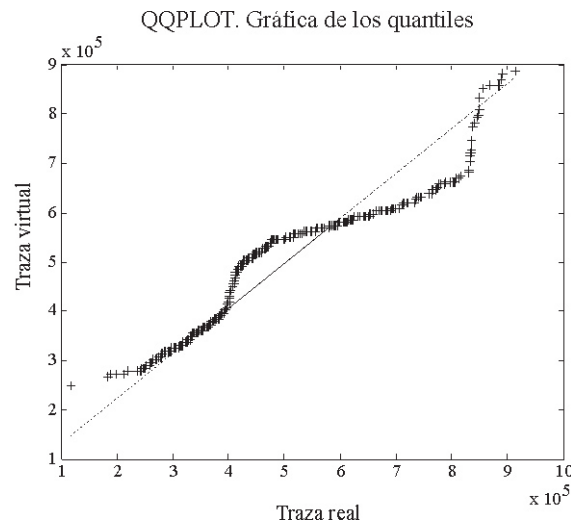
En la figura 17 se puede observar cierta desviación del modelo. Esto se debe a que el tamaño de cada paquete virtual, es generado con el valor promedio del tamaño de los paquetes de la traza original, lo que altera el modelo. Otro de los factores que alteran el modelo es la cantidad de datos capturados de la red, ya que para esta muestra se realizó una toma de datos con una duración de tiempo pequeña. Se evidencio que el modelo se ajusta mejor o presenta una mejor



**Figura 15** Traza de tráfico Original tomando un  $\Delta t = 500$  ms



**Figura 16** Traza de tráfico generada mediante MMPP tomando un  $\Delta t = 500$  ms



**Figura 17** Gráfica de quantiles (Q-Q plot) tomando una unidad de tiempo  $\Delta t = 500$  ms

prueba de bondad cuando el número de datos es mayor como es de esperarse en cualquier proceso estocástico.

Además, la cantidad de usuarios que están dentro de las VLAN (diferentes dominios de broadcast), que utilizan como canal de entrada/salida al canal de salida 2, es pequeña. Esto hace que el tráfico varíe poco en el tiempo y por ende genera menor número de estados en la cadena de Markov en el momento de caracterizar el tráfico.

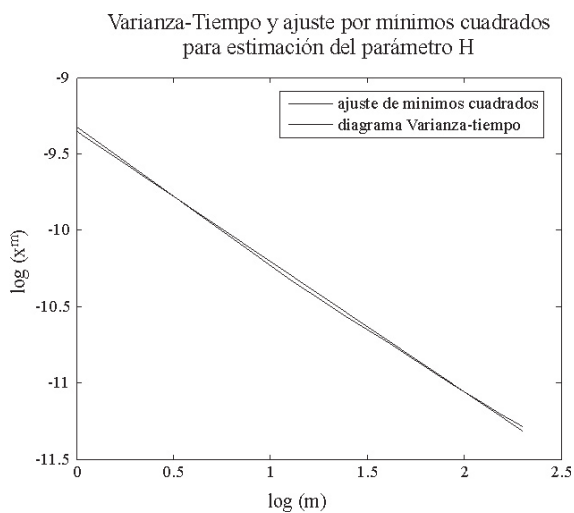
### **Estimación del parámetro Hurst**

El parámetro Hurst es muy importante estimarlo debido a que por medio de él se puede concluir si una traza de tráfico es o no autosimilar. Si el parámetro Hurst ( $H$ ) posee un valor entre el rango ( $0.5 < H < 1$ ), quiere decir que el tráfico es autosimilar. De no cumplirse esta condición se concluye lo contrario. Es necesario determinar, si las trazas de tráfico en estudio cumplen con el criterio de autosimilaridad, ya que el modelo utilizado (MMPP) supone este tipo de característica.

Para la estimación del parámetro Hurst ( $H$ ) existen varios métodos, ya mostrados en el marco teórico del artículo. El método escogido fue varianza-tiempo (varianza de los agregados) debido a que este método es muy sencillo de realizar y arroja resultados confiables. Este método ha sido utilizado en trabajos anteriores como en el de Alzate Monroy, teniendo resultados satisfactorios. Si el tráfico es o no autosimilar, se deben tener en cuenta dos aspectos: el primero es observar el valor de la pendiente del diagrama varianza-tiempo; si este se encuentra entre 0 y -1, entonces el tráfico es autosimilar. Esto, debido a que la pendiente de esta recta es  $\beta$ , y  $\beta = 2(H-1)$ . El segundo aspecto es determinar el valor de  $H$  a partir de la definición anteriormente mencionada. Este valor es visualizado en la interfaz de la herramienta desarrollada en Matlab 7.1 para tal fin.

#### Proveedor 1

En la figura 18 se puede observar el diagrama varianza-tiempo para la estimación del parámetro  $H$  para el canal 1.



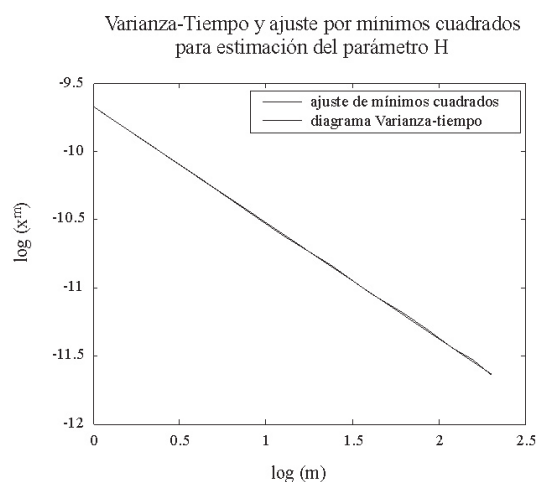
**Figura 18** Diagrama varianza-tiempo para la estimación del parámetro  $H$

En la figura 18 la recta correspondiente al diagrama varianza-tiempo, posee una pendiente  $\beta = -0,92$  y el algoritmo propuesto para la estimación del parámetro Hurst arroja un valor de  $H = 0,5581$ . Esto es suficiente para concluir que el tráfico es autosimilar.

En el estudio se realizó la caracterización y la simulación con unidades de tiempo de 500 ms, 100 ms y otras, tanto para el canal 1 como para el canal 2. El algoritmo para calcular el parámetro Hurst real es independiente del algoritmo de discretización de la traza, por lo cual se determina que el parámetro Hurst real no depende de la unidad de tiempo ( $\Delta t$ ) que se haya escogido para discretizar la traza original; el parámetro Hurst real depende únicamente de los tiempos de llegada de cada una de las tramas. Sin embargo, debido que en la generación de la traza virtual, interviene este proceso de discretización, es importante observar que el parámetro Hurst sí dependerá de la unidad de tiempo que se haya escogido para el análisis. De lo anterior se pudo determinar que la comparación entre el parámetro Hurst real y virtual puede ser un criterio de validación del modelo.

#### Proveedor 2

En la figura 19 se observa el diagrama varianza-tiempo que se obtiene para el canal 2. Para esta traza de tráfico el algoritmo arrojó un valor de  $H = 0,5753$ , lo que indica que el tráfico es autosimilar en el rango de tiempo de captura.



**Figura 19** Diagrama varianza-tiempo para la estimación del parámetro  $H$

#### **Estimación del tiempo de espera promedio y probabilidad de pérdidas**

El tiempo de espera promedio se refiere al tiempo que debe esperar cada paquete para ser atendido, a

medida que se aumenta la intensidad del tráfico. El tiempo de espera está definido como la diferencia entre el tiempo en que es atendido el paquete y el tiempo en que este llega a la cola del sistema.

El sistema de cola del presente trabajo fue MMPP/D/1, esto debido a que la fuente es generada mediante el modelo MMPP, la distribución de los tiempos de servicio Discreta. Esto es, se asume que el tiempo que el enrutador demora en decidir la ruta de los paquetes es constante debido a que es un enrutador de borde y solo tiene una entrada y una salida.

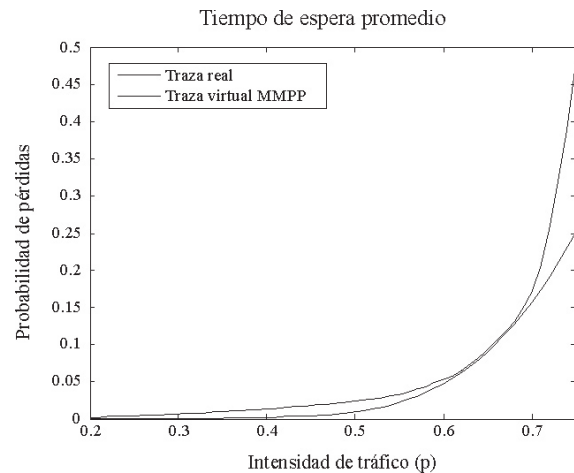
La probabilidad de pérdidas indica qué probabilidad existe de que se pierdan paquetes debido a un buffer lleno, es decir, cuando la capacidad de la cola está ocupada totalmente, entonces los paquetes que lleguen en el instante en que el buffer esté lleno, serán rechazados y se consideran paquetes perdidos. Este análisis se desarrolla a medida que se aumenta la intensidad de tráfico.

Es importante resaltar que la variación de los tiempos de servicio dependen de la tasa promedio de llegadas de paquetes por unidad de tiempo, ya que el tiempo de servicio está definido como el cociente entre la intensidad de tráfico y el promedio de las tramas de llegada por unidad de tiempo.

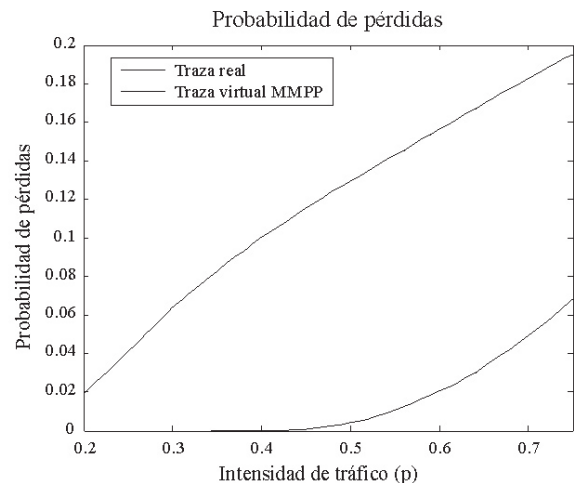
#### Proveedor 1

La traza de tráfico que se analizó corresponde a los datos de entrada desde Internet hacia la RDUPBB. Primero es importante mostrar la tasa promedio de llegada de tramas por unidad de tiempo para esta traza de tráfico, la cual posee un valor de 233.0989 (tramas por unidad de tiempo). En las figuras 20 y 21 se muestran las gráficas del tiempo de espera promedio y probabilidad de pérdidas, para esta traza de tráfico respectivamente.

Se puede observar, en la figura 20, que las dos trazas siguen la misma curva exponencial, sus discrepancias son mínimas, aquí se puede observar también que la traza generada mediante MMPP es un poco más optimista que la traza original.



**Figura 20** Tiempo de espera promedio para un búfer infinito tomando un  $\Delta t = 500$  ms



**Figura 21** Probabilidad de pérdidas para un búfer de 10 tomando un  $\Delta t = 500$  ms

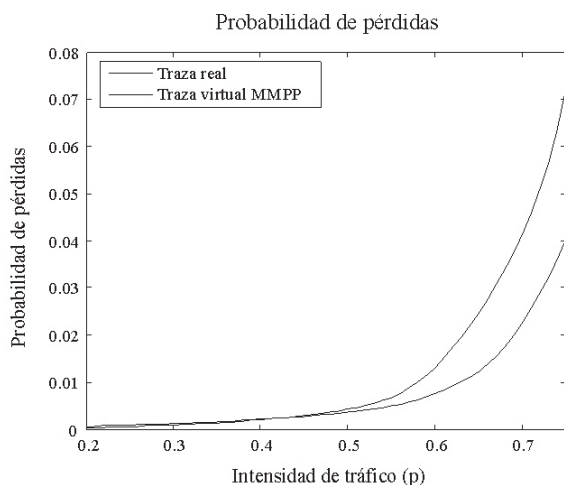
Se observa, en la figura 21, que la traza generada mediante MMPP es muy optimista comparada con la traza real. La probabilidad de pérdidas se incrementa en mayor proporción para la traza real, a medida que se aumenta la intensidad de tráfico. Si se compara la probabilidad para un solo valor de intensidad de tráfico con alta discrepancia; por ejemplo, para  $\rho = 0,6$ , se aprecia una diferencia entre las dos trazas de casi 0,15. Para una intensidad de tráfico de 0,6 se perderían aproximadamente 20 paquetes por cada 1.000 transmitidos para el caso de la traza generada. Para la misma intensidad de tráfico de 0,6 se perderían 150 paquetes por cada 1.000 transmi-

tidos para el caso de la traza real. Se aprecia que en ese punto de intensidad de tráfico es un punto de gran discrepancia entre el comportamiento de la probabilidad para la traza real y la virtual. Sin embargo, la aproximación es buena si se tiene en cuenta que el tamaño del buffer es bastante pequeño, lo que no ocurre en un dispositivo activo real y actual.

### Proveedor 2

El análisis del tiempo de espera promedio se desarrolla, teniendo un buffer infinito, es decir, sin pérdidas de paquetes, mientras que el análisis de la probabilidad de pérdidas se desarrolla teniendo un búfer de valor finito.

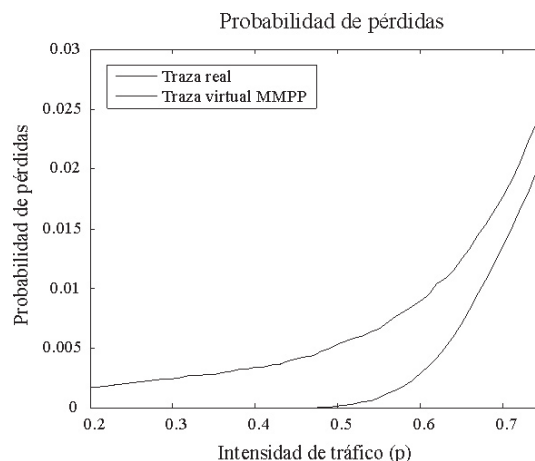
En las figuras 22 y 23 se muestra las gráficas del tiempo de espera promedio y probabilidad de pérdidas, para esta traza de tráfico respectivamente.



**Figura 22** Tiempo de espera promedio para un búfer infinito tomando un  $\Delta t = 500$  ms

En la figura 22 se puede observar que la traza de tráfico generada mediante MMPP es un poco más optimista que la traza real, sus discrepancias se encuentran dentro del mínimo permisible que es 600 ms, es importante observar que las dos trazas siguen la misma curva exponencial.

Se puede observar, en la figura 23, que las dos trazas siguen la misma curva exponencial, sus



**Figura 23** Probabilidad de pérdidas para un búfer de 10 tomando un  $\Delta t = 500$  ms

discrepancias son pequeñas. A partir de una intensidad de tráfico de 0,5 se comienzan a incrementar en mayor proporción las pérdidas por paquetes, esto es lo esperado debido a que al aumentar la intensidad de tráfico, aumenta también el tiempo de servicio, entonces los paquetes deberán esperar mayor tiempo en la cola y esto hace que se incremente las pérdidas.

### Síntesis de resultados

En la tabla 7 se muestran los principales resultados de las simulaciones, como son:  $\lambda$  promedio para la traza original y para la generada mediante MMPP, las tasas de transferencia (bps) para la traza original y para la traza generada mediante MMPP, el parámetro Hurst para el tráfico original y para el tráfico generado mediante MMPP, los *bytes* promedio característicos de cada traza que se desea analizar y el total de los paquetes que se capturaron con los analizadores de tráfico.

### Conclusiones

- Nuestro estudio demuestra que esta metodología puede ser utilizada en redes LAN bajo condiciones reales. Para tal fin se utiliza el tráfico capturado en los enlaces de entrada/salida de Internet de la Red LAN de la Universidad Pontificia Bolivariana, seccional



**Tabla 7** Comparación de resultados de simulación para canal 1 y canal 2

<b>Parámetros</b>	<b>Canal 1</b>		<b>Canal 2</b>	
	<b>Entrada</b>	<b>Salida</b>	<b>Entrada</b>	<b>Salida</b>
Lambda promedio real (tramas/ $\Delta t$ )	234,214000	288,037000	188,137000	191,769000
Lambda promedio DMMPP (tramas/ $\Delta t$ )	240,545000	240,832000	187,854000	193,352000
Tasa de transferencia real (bits/ $\Delta t$ )	1.338.630,000000	1.175.900,000000	369.980,000000	1.290.360,000000
Tasa de transferencia D-MMPP (bits/ $\Delta t$ )	1.223.140,000000	1.348.710,000000	380.181,000000	1.302.030,000000
Parámetro Hurst real	0,574749	0,657801	0,635703	0,639018
Parámetro Hurst MMPP	0,670572	0,731812	0,674526	0,687283
Bytes promedio	712,907000	642,609000	245,818000	841,089000
Total tramas de llegada	38.083,000000	37.119,000000	835.982,000000	852.139,000000

Bucaramanga. Los resultados demuestran que esta metodología genera una traza simulada bastante aproximada a la traza medida. La exactitud del modelo se ve reflejada en la gráfica de los quantiles (Q-Q plot) entre la traza medida (real) y la traza generada mediante D-MMPP.

- El modelo D-MMPP es más aproximado cuando se analiza el tráfico IP (tramas o paquetes de llegada por unidad de tiempo), ya que la cantidad de tramas o paquetes que entran desde Internet o salen hacia Internet, tienen un comportamiento autosimilar. Cuando se ajusta el modelo a bits por unidad de tiempo se introduce a un porcentaje de error, generado a causa de trabajar con el promedio del tamaño (en bits) de los paquetes virtuales de llegada de la traza original de tráfico. Esto produce cierta desviación entre la traza real y la traza generada mediante D-MMPP, lo cual puede observarse en la gráfica de los quantiles. Se propone entonces una solución de esto para

futuros trabajos mediante el modelamiento del tamaño de las tramas para cada estado de la cadena de Markov caracterizado por su tasa de generación específica. Es importante resaltar que para obtener una buena aproximación del modelo D-MMPP, es necesario realizar una toma de tráfico amplia; tomar datos por un periodo de tiempo grande, de esta forma se obtiene una mejor aproximación en el modelamiento del comportamiento autosimilar. Para las trazas de tráfico capturadas en la RDUPBB, se realizó la toma de datos por un periodo de tiempo pequeño, con un promedio de 200 s, lo que origina ciertas desviaciones entre el tráfico real y el generado mediante el modelo D-MMPP.

- Para la implementación de las simulaciones se desarrollaron y codificaron los algoritmos correspondientes para generar la traza discreta y para analizar su comportamiento en la cola MMPP/D/1. Se obtuvo una herramienta de simulación en Matlab 7.1 con una GUI (interfaz

gráfica para el usuario). Esta herramienta permite que el usuario pueda simular y analizar cualquier traza de tráfico IP de forma práctica y guiada.

- La herramienta desarrollada puede utilizarse para cualquier traza con comportamiento autosimilar; esto involucra que se haga una aproximación al tamaño de las tramas que se generan. Para el efecto, se trabajaron los tamaños de las tramas virtuales generadas mediante MMPP como el promedio de los tamaños de las tramas de llegada de la traza original. Esta aproximación hace que el modelo sea muy optimista para ciertas trazas de tráfico en las cuales el tráfico tiene mayor variabilidad en la cantidad de tramas que llegan por unidad de tiempo.
- El modelo del canal para los nodos de entrada y salida de Internet en la RDUPBB fue analizado por medio de una cola D-MMPP/D/. La fuente fue generada mediante el modelo D-MMPP, los tiempos de servicio son determinísticos, los cuales dependen del promedio de las tramas de llegada de la traza original. Se asume que el tamaño de los paquetes es 1, por esta razón el tiempo de servicio también se hace 1 con lo cual se tiene la variable tiempo de servicio de tipo determinístico. Este tipo de cola posee un solo servidor, el cual representa la RDUPBB o el router de borde, dependiendo de si se habla del tráfico de entrada o de salida respectivamente. Los tiempos de espera promedio por trama dependen exclusivamente de qué tan congestionada se encuentra la red. Si se encuentra con una red que alcanza grandes cantidades de tramas por unidad de tiempo, este retraso promedio tendrá valores pequeños, debido a que las tramas llegan más cercanas una de otra, es decir, el tiempo entre tramas de llegada es menor y con esto deben esperar menos tiempo en ser atendidos. De forma contraria ocurre cuando el tráfico posee una pequeña cantidad de tramas por unidad de tiempo. El tiempo de servicio depende del promedio de tramas de llegada por unidad de tiempo. Para trazas que poseen un alto promedio de tramas de llegada por unidad de tiempo, se tienen tiempos de servicio muy pequeños, pero para trazas con bajo promedio de llegada de tramas por unidad de tiempo, se tienen tiempos de servicio grandes.
- Las tasas de entrada de paquetes o tramas desde Internet son mayores que las tasas de salida de paquetes o tramas hacia Internet para ambos proveedores de servicio de la RDUPBB (proveedor 1 y proveedor 2). Esto se debe a que en la mayoría de los casos en una red LAN solo se envía una trama de petición de cierta información y en respuesta se reciben las tramas que conforman la información solicitada. Las tasas de transferencia en (bits por unidad de tiempo) para el tráfico generado con el proveedor 1 son mayores que las tasas de transferencia para el tráfico del proveedor 2. Esto se debe a que el proveedor 1 dentro de la RDUPBB contiene mayor número de usuarios, además de transportar aplicaciones P2P utilizadas a menudo por los estudiantes. Este tipo de aplicaciones envían y se reciben constantemente paquetes o tramas hacia y desde Internet, mientras que el proveedor 2 es utilizado por los docentes y administrativos que poco utilizan aplicaciones que generan alto tráfico.
- Las probabilidades de pérdidas de tramas para el proveedor 1 son mayores que las probabilidades para el proveedor 2. Este resultado es el esperado debido a que el proveedor 1 mantiene la red mucho más congestionada que el proveedor 2, ya que este último es solo utilizado por un pequeño número de usuarios con características de un usuario de bajo consumo de canal. Al aumentar el tráfico en cualquiera de las trazas trabajadas se aumenta la probabilidad de pérdidas de tramas y viceversa.
- Las trazas de tráfico de entrada y salida de Internet para el proveedor 2 contienen el menor número de estados en la cadena de Markov. Debido a que el número de estados es bajo, el modelo se puede definir con pocos estados. Esto implica una mayor invariabilidad del

tráfico en el modelo. Además se pudo verificar que para trazas de tráfico con pocos estados en la cadena de Markov el modelo no es muy optimista. Las trazas de tráfico de entrada y salida de Internet para el proveedor 1 contienen mayor cantidad de estados en la cadena de Markov. Esto hace que el tráfico sea muy variante en el tiempo en el modelo. Con esto se logra una mayor aproximación del modelo debido a que hay menores discrepancias entre intervalos de tiempo.

## Referencias

1. D. P. Heyman, D. Lucantoni, "Modeling IP traffic streams with rate limits. IEEE/ACM". *Transactions on Networking*, Vol. 11, 2003, pp. 948-958.
2. <http://ita.ee.lbl.gov/html/contrib/BC.html>. Febrero, 2006
3. K. Shoji, "Internet Tráfico Modelling: Markovian Approach to Self-Similar Traffic and Prediction of Loss Probability For Finite Queues". *IEICE TRANS. COMMUN.* Vol. E84-B, August, 2001, pp. 2134-2141.
4. M. A. Alzate Monroy, *Introducción al tráfico autosimilar en redes de comunicación*. Universidad Distrital Francisco José de Caldas. Revista Ingeniería. Año 2001.
5. Home page RUPBB [www.upbbga.edu.co](http://www.upbbga.edu.co). Consultado marzo 2007.
6. W. Leland, M. Taqqu, W. Willinger, D. Wilson, "On the self-similar Nature of Ethernet Traffic(extended version)" *IEEE/ACM Transaction on Networking*, Vol 2, 1994, pp. 1-15.
7. D. P. Heyman, T.V. Laskman, A. Neidhardt, "A New Method for Analysing Feedback-Based Protocols with Applications to Engineering Web Traffic Over Internet" *SIGMETRICS 97*, Seattle, WA, USA. Disponible en <http://networks.ecse.rpi.edu/~rsatsh/sigmetrics.html>. Consultado mayo de 2006.
8. QoS & Traffic Management in IP & ATM, McDYSAN, New York. Ed. McGraw-Hill. 2000.
9. A. Erramilli, P. Pruthi, W. Willinger, "Recent developments in fractal traffic modeling" *Proc. of ITC Regional Seminar*, St. Petersburg, 26 June-2 July, 1995, pp. 1-12.
10. V. Paxson, S. Floyd, "Wide Area Traffic. The Failure of Poisson Modelling". *IEEE/ACM Transactions on Networking*, Vol. 3, 1995, pp. 226-244.
11. N. C. Hock, *Queueing Modelling Fundamentals*, New York. Ed. Wiley, 1997, pp. 214.
12. M. Taqqu, W. Wilinger, V. Teverosky, "Estimators for Long- Range Dependence: An Empirical Study. Fractals" Vol 3, 1995, pp. 785-798.
13. S. Giordano, S. Miduri, M. Pagano, F. Russo, S. Tartarelli, "Wavelet-based aproach to the estimation of the Hurst parameter for self-similar data". Universidad de Pisa, Italia.
14. G. A. Borray, J. P. Páez, J. M. Wilches, *Modelado matemático del tráfico de redes de telecomunicaciones de acceso fijo*. Pontificia Universidad Javeriana. 2001.
15. D. M. Lucantoni, "The BMAP/G/1 Queue: A Tutorial", *Models and techniques for performance Evaluations of computer and communications systems*, L. Donatiello, R. Nelson, eds. New York: Springer-Verlang, 1993, pp. 330-358.
16. R. Vecchio, *Modelización de tráfico autosimilar y evaluación de sus efectos en el tamaño de las colas*. Instituto Tecnológico de Buenos Aires. 2003. [www.itba.edu.ar/capis/epg-tesis-y-tf/vecchio-tesisdemagister.pdf](http://www.itba.edu.ar/capis/epg-tesis-y-tf/vecchio-tesisdemagister.pdf). Mayo 2006.
17. C. V. Castellanos López, *Modelo de tráfico IP*. Universidad de los andes. 2005.
18. Home page Ethereal [www.ethereal.com](http://www.ethereal.com). Consultado mayo de 2006.
19. <http://www.ngthomas.co.uk/dice.htm>. Consultado marzo de 2006.