

# CARACTERIZACIÓN DE TRÁFICO EN UNA RED DE ÁREA LOCAL

---

**JORGE ALBERTO GÓMEZ MARTÍNEZ**  
*Aspirante a Magister en Ingeniería Electrónica*  
*Universidad Industrial de Santander*  
*Bucaramanga, Colombia*  
*gomezuis@gmail.com*

**OSCAR GUALDRÓN GONZÁLEZ**  
*Phd en Física*  
*Université Laval Québec*  
*Profesor asociado*  
*Universidad Industrial de Santander*  
*Bucaramanga, Colombia*  
*gualdron@uis.edu.co*

*Fecha de Recibido: 24/07/2009*  
*Fecha de Aprobación: 02/12/2009*

## RESUMEN

Este trabajo de investigación se orienta al análisis de un segmento de una LAN, con el fin de caracterizar el tráfico entrante y saliente a servidores para posteriormente contrastarlo mediante el uso de un simulador de tráfico. Para lograrlo, se registran datagramas de diferentes servidores correspondientes a protocolos comúnmente utilizados como el SMTP<sup>1</sup>, HTTP<sup>2</sup>, entre otros. Con dichos datagramas se determina el grado de auto-similitud y las características probabilísticas que presenta cada servidor, utilizando diferentes pruebas para cada caso. Una vez caracterizado el tráfico de los servidores, se utiliza una herramienta de simulación para conseguir un ambiente de tráfico similar al real.

**PALABRAS CLAVES:** Autosimilaridad, Dependencia a largo plazo, Parámetro de Hurst, Estimador de Whittle, periodograma.

## ABSTRACT

This research work presents an analysis of the traffic on a LAN segment heading to model its behavior and the generation of similar results by using a traffic simulation tool. To achieve this, datagrams are collected from servers supporting commonly used protocols such as SMTP<sup>1</sup> and HTTP<sup>2</sup>. The datagrams are then used to determine the degree of self-similarity of the network segment and its probabilistic characteristics. Different tests are used for each case. Once the network segment model is obtained, a simulation tool is used to get a similar traffic environment to the real network traffic.

**KEY WORDS:** traffic, self-similarity, long range dependence, Hurst parameter, Whittle estimator, periodogram.

---

1 Por sus siglas en inglés Simple Mail Transfer Protocol.

2 Por sus siglas en inglés Hypertext Transfer Protocol.

## 1. INTRODUCCIÓN

La caracterización de tráfico es un problema de vieja data, en el cual la distribución de los tiempos transcurridos entre la llegada de dos datagramas sucesivos (*interarrival time*) ha sido objeto de numerosos trabajos [1,2,10,11]. Si bien inicialmente la distribución de Poisson se consideró como válida para describir la estadística del proceso en una red de área local (*local area network, LAN*), la ocurrencia de ráfagas a diferentes escalas de tiempo condujo a contemplar modelos que no solamente comprendieran la dinámica a corto plazo (*short-range dependence, SRD*) sino también la de largo plazo (*long-range dependence, LRD*). En 1990 Will Leland y Daniel Wilson [1] comenzaron a proponer modelos estadísticos para describir el comportamiento de las redes LAN, concluyendo que los modelos tradicionales aplicados en comunicaciones no se adaptaban muy bien a ellas. Investigaciones posteriores encontraron que los modelos fractales aplicados no en el espacio si no en el tiempo, permitían modelar adecuadamente el proceso al involucrar la característica conocida como autosimilaridad [2].

El origen del comportamiento autosimilar ha sido atribuido a la distribución del tamaño de los paquetes, la interacción humana con el medio, la dinámica de la red y la dependencia a largo plazo (LRD).

## 2. HERRAMIENTAS DE SIMULACIÓN

El objetivo fundamental de los simuladores es comprender y/o evaluar el comportamiento del objeto real simulado, cuyos resultados deben ser lo suficientemente aproximados a los reales para extrapolarlos, y como consecuencia ahorrar tiempo y dinero al poder solucionar problemas en forma virtual.

Las herramientas de simulación presentan las siguientes ventajas [3]:

- Facilitan el análisis de tráfico.
- Permiten evaluar el desempeño de las redes frente a la implementación de nuevas tecnologías.
- Representan un ahorro sustancial de costos.
- Permiten la manipulación de la red, efectuar cambios y experimentar virtualmente.
- Complementan los conocimientos de análisis de redes.
- Mejoran la planeación y el diseño de nuevas redes y servicios de telecomunicaciones.

Las herramientas de simulación presentan las siguientes desventajas [3]:

- La veracidad de los resultados de la simulación depende directamente de la herramienta utilizada.
- Los resultados de una simulación son únicamente una aproximación del sistema real.
- Dificultad para considerar todos los parámetros.
- La validación del modelo se realiza con base en fundamentos teóricos.

Existe una gran variedad de aplicaciones para simular redes, algunas de ellas comerciales otras del tipo *Open Source*. Algunos de los simuladores más conocidos son: AdventNet, FLAN, Kiva, Network Simulator, OMNeT++, OpNet, IT Guru, Packet Tracer, y Shunra VE Desktop.

En este trabajo se utilizó Network Simulator 2, NS-2, una herramienta que acumula casi dos décadas de desarrollo, de código abierto y frecuentemente utilizada como base para investigaciones en tráfico de redes. NS-2 incluye soporte para redes terrestres, satelitales e inalámbricas; fuentes de tráfico como web, ftp, telnet, cbr, http, tráfico estocástico; fallas determinísticas, pérdidas probabilísticas, falla de conexiones, entre otras.

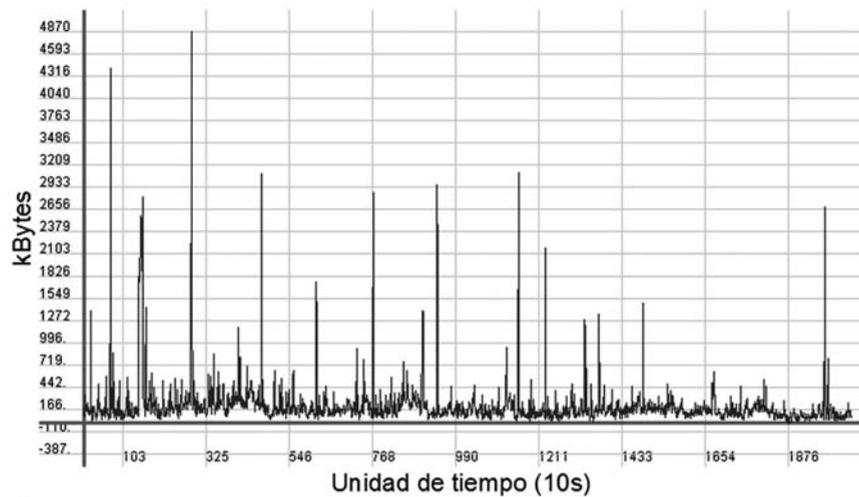
## 3. AUTOSIMILITUD

La autosimilaridad es una característica en donde existe similaridad estructural en diferentes escalas en una dimensión específica, como el tiempo o el espacio. En otras palabras, existe un patrón que se repite a sí mismo al cambiar su escala de medida, y las estadísticas del proceso no varían. Esta característica puede no tener límite de escalas en modelos matemáticos aunque en la realidad se observa solamente para un rango de escalas definido. La figura 1 ilustra mejor la autosimilaridad de un fenómeno en el tiempo, que corresponde al volumen de tráfico en kiloBytes para una interfaz del core switch de la Universidad Industrial de Santander. Dicha interfaz corresponde a la subred donde se ubican los servidores de interés para este trabajo.

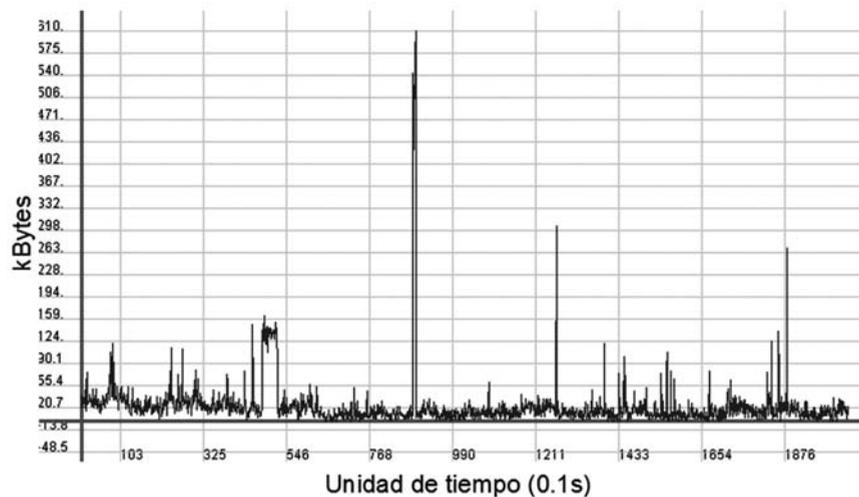
Se observa que al cambiar la escala temporal de 10 segundos (a) a 0.1 segundos (b) siguen existiendo ráfagas cortas o picos. Cada una de las ilustraciones se realizó con 2000 datos, lo que representa 5 horas de tráfico para (a) y solo un poco más de 3 minutos para (b). Los modelos de tráfico convencionales no reflejan esta situación adecuadamente, ya que al cambiar la escala dichas ráfagas se suavizan.

La autosimilaridad se asocia con la geometría fractal, que describe objetos sin cambios de apariencia en diferentes escalas de medida. El concepto fractal incluye estadística y dinámica cuya descripción requiere espacios de dimensionalidad diversa. Algunos ejemplos de situaciones en las que se evidencia naturaleza fractal son el triángulo de Sierpinski (geometría), la ley de Pareto (economía), la ley de Zipf (lingüística), la ley de Lotka (sociología), la ley de Willis (biología), el espectro cardiaco (medicina), y los mapas determinísticos (modelamiento telegráfico).

La conducta autosimilar está ligada también a la teoría del caos. Un sistema caótico es altamente sensible a las condiciones iniciales del sistema, de manera que cambios pequeños en las condiciones iniciales pueden modificar el comportamiento del sistema a largo plazo. Una de las características fundamentales de estos sistemas es que su aleatoriedad crea formas determinísticas, en consecuencia algunos fenómenos en la naturaleza aparentemente aleatorios tienen un orden subyacente. Ejemplos de la teoría del caos se encuentran en la ecología, predicciones de población biológica, el corazón humano, ciencias computacionales, y las imágenes fractales.



(a) Unidad de tiempo de 10 segundos.



(b) Unidad de tiempo de 0.1 segundos.

**Figura 1.** Trazas de tráfico del core switch de la UIS.

### 3.1 El parámetro de Hurst

El parámetro de Hurst indica el grado de autosimilitud de una serie de datos y su valor oscila entre  $0.5 \leq H \leq 1.0$ . Cuando  $H$  tiende a 0.5 la autosimilaridad es baja, o ninguna en caso que sea igual a 0.5; cuando tiende a 1.0 es muy autosimilar, o exactamente autosimilar cuando es igual a 1.0.

El parámetro de Hurst es la base para modelar una red. El estudio del tráfico se desarrolla en dos tipos diferentes de fuentes: los modelos que logran caracterizar fuentes individuales de tráfico y los modelos para fuentes de tráfico agregado. Independientemente del tipo de estudio que se realice, ambos presentan normalmente un valor de Hurst entre  $0.7 < H < 0.9$  [5].

Algunas de las formas más utilizadas para medir el grado de autosimilaridad en un sistema son:

- Dependencia a largo plazo
- Decaimiento lento de la varianza
- Rango reescalado o R/S
- Periodograma
- Estimador de Whittle

En este trabajo, se abordó el cálculo del parámetro de Hurst mediante el Periodograma y el estimador de Whittle.

#### 3.1.1 Periodograma

El periodograma [5] es un método que ofrece un buen rigor estadístico. El periodograma (o función de intensidad)  $I_N(\omega)$  es la densidad espectral estimada de un proceso estocástico discreto  $X(t)$ . Se calcula mediante la ecuación 1, donde  $X_k$  es la serie de datos,  $N$  es la longitud de  $X_k$  y  $\omega$  es la frecuencia.

$$I_N(\omega) = \frac{1}{2\pi N} \left| \sum_{k=1}^N (X_k e^{jk\omega}) \right|^2 \quad (1)$$

Dado que la autosimilaridad afecta el comportamiento del espectro cuando  $\omega$  tiende a 0, con lo cual la expresión en (1) se aproxima a  $I_N(\omega) \sim |\omega|^{1-2H}$ . Al graficar  $\log[I_N(\omega)]$

contra  $\log(\omega)$  (únicamente en frecuencias bajas), y agregando una línea recta a la curva, la pendiente se aproxima a  $1-2H$ .

#### 3.1.2 Estimador de Whittle

El estimador de Whittle [5] está basado en el periodograma y utiliza máxima verosimilitud. Se calcula mediante la ecuación 2 para  $d = H - 0.5$

$$Q(d) = \int_{-\pi}^{\pi} \frac{I_N(\omega)}{S(\omega)} \quad (2)$$

En la práctica la integral de la ecuación 2 es una sumatoria. Considerando la frecuencia del sistema  $\omega_k = k \frac{2\pi}{N}$  para los valores de  $k$  desde 0 hasta  $N-1$ , se llega a la ecuación 3.

$$Q(d) = \sum_{k=0}^{N-1} \frac{I_N(\omega)}{S(\omega)} \quad (3)$$

Donde la ecuación de densidad espectral se expresa mediante la ecuación 4, con una función que corresponde a  $R(k)$ .

$$S(\omega) = \sum_{k=-\infty}^{\infty} R(k) e^{-j2k\omega} \quad (4)$$

El estimador de Whittle tiene la limitante de ser paramétrico y por lo tanto requiere conocer la densidad espectral con anterioridad, lo cual se traduce en un alto costo computacional.

## 4. RESULTADOS EN LA LAN UIS

Se seleccionaron servidores dentro del mismo segmento de red, donde el tráfico de estos se recibe por un mismo puerto del core switch institucional. Además, se seleccionaron por utilizar diferentes protocolos importantes como FTP, HTTP, SMTP y DNS. Con esto se obtiene una porción de la red limitada en tamaño, pero que facilita la experimentación con datos reales de los protocolos de mayor uso en la mayoría de redes.

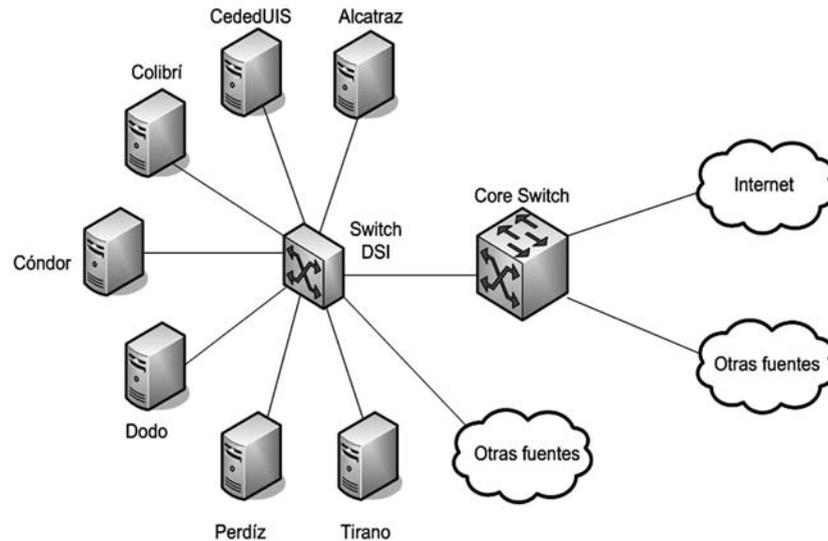


Figura 2. Esquema de evaluación.

La figura 2 muestra el esquema utilizado para la toma de datos sobre la red, en el cual se muestreó el tráfico entrante y saliente del puerto asociado en el core switch al segmento de red seleccionado. Por supuesto en este puerto se registra tráfico de otras fuentes, que debe ser filtrada.

#### 4.1 Pruebas de autosimilitud sobre la red

Sobre el Core Switch se realizó un análisis del tráfico mediante el analizador de protocolos Wireshark (<http://www.wireshark.org>), el cual registra los datagramas entrantes y salientes de los servidores para varios días de tráfico.

Se desarrollan dos pruebas diferentes para identificar posibles patrones de autosimilitud en el segmento de red.

La primera de estas pruebas consiste en comparar el comportamiento del tráfico agregado contra un segmento del mismo, para explorar la hipótesis de autosimilitud, según la cual se debe presentar un valor de parámetro de Hurst constante a medida que se cambia la escala de tiempo, como se propone en [3]. El tamaño para los granos utilizados en esta prueba es de 0,1s, 1s y 10s. La tabla 1 presenta resultados de la estimación del parámetro de Hurst para los 3 tamaños de grano, mientras P y W corresponden al uso del periodograma y del estimador de Whittle respectivamente

La segunda prueba consiste en comparar un registro consigo mismo para tamaños de grano diferentes y observar sus efectos en los resultados. El tamaño de los granos utilizados es de 10 segundos y 0.625 segundos. Para este tipo de prueba se separa individualmente el tráfico de cada servidor para comparar las características de tráfico de cada uno y comparar sus resultados contra el core switch. En esta prueba se encuentra que varios servidores aportan menos del 1% del tráfico, por lo cual su efecto es marginal y no serán estudiados profundamente debido al precio computacional que implica su análisis. Sin embargo, se hace una profundización a dos servidores (S1 tráfico SMTP y S2 tráfico HTTP) que aportan el 95% del tráfico del segmento de red seleccionado. La tabla 2 muestra un resumen de los resultados promediados de los varios días de tráfico con los que se hizo la prueba.

De la primera prueba se observa que tanto el periodograma como el estimador de Whittle disminuyen a medida que el grano aumenta de tamaño, situación que era de esperar debido a que al aumentar el tiempo de la prueba, la autosimilitud entre los datos se degrada. Esto es un problema ya que se pierde la posibilidad de caracterizar la dependencia a largo plazo.

De la segunda prueba se observa que un grano más fino conduce a mejores resultados en autosimilitud aunque no es muy notoria en el periodograma.

Tabla 1. Resultados prueba 1.

(a) Recepción			
Grano →	0,1	1	10
P	0,907	0,681	0,571
W	0,881	0,667	0,573
P	0,786	0,522	0,52
W	0,836	0,639	0,567
P	0,883	0,634	0,503
W	0,842	0,702	0,635
P	0,79	0,548	0,572
W	0,823	0,549	0,598
(b) Transmisión			
Grano →	0,1	1	10
P	0,671	0,665	0,651
W	0,965	0,756	0,659
P	1,181	0,693	0,687
W	0,884	0,83	0,72
P	0,893	0,799	0,704
W	0,999	0,892	0,663
P	0,731	0,619	0,628
W	0,734	0,781	0,643

Tabla 2. Resultados prueba 2.

(a) Recepción				
Grano ↓		S1	S2	Core
0.625s	P	0,606	0,550	0,598
	W	0,859	0,920	0,875
10s	P	0,574	0,552	0,572
	W	0,584	0,572	0,599
(a) Transmisión				
Grano ↓		S1	S2	Core
0.625s	P	0,638	0,636	0,645
	W	0,788	0,675	0,780
10s	P	0,669	0,679	0,683
	W	0,648	0,698	0,673

## 4.2 Aleatoriedad en la red

Se desarrolla además un análisis probabilístico de la red, para determinar la función de densidad de probabilidad (*PDF*) de los servidores y el core switch.

Se deben realizar pruebas de bondad de ajuste, siendo las más conocidas son la Kolmogorov-Smirnov y la Chi cuadrado [6].

Se determina con las pruebas que la función de distribución de probabilidad para el servidor 1 se ajusta mejor a una del tipo log-logística, mientras que el servidor 2 y el Core Switch se ajustan mejor una distribución tipo Burr. La tabla 3 muestra el promedio de los resultados de las pruebas de bondad de ajuste.

Tabla 3. Resultados prueba probabilística.

Servidor 1	Recepción	Transmisión
$\alpha$	1,7	18,8
$\beta$	400,9	84,2
Distribución Log-logística		
Servidor 2	Recepción	Transmisión
k	28,7	18,5
$\alpha$	24,5	2,6
$\beta$	1618,3	361,0
Distribución Burr		
Core switch	Recepción	Transmisión
k	10,1	0,7
$\alpha$	30,6	37,4
$\beta$	1522,7	329,5
Distribución Burr		

## 5. SIMULACIÓN DE LA RED

### 5.1 Escenario de simulación y ajustes

La figura 3 muestra el escenario de simulación en formato nam de ns2, en la cual los nodos incluidos son:

- Nodo 0: Switch DSI, conecta los servidores a la Intranet.
- Nodo 1: Switch clientes Internet.
- Nodo 2: Core switch, conecta los servidores y los clientes de la Intranet.
- Nodo 3: Servidor Condor (SMTP).
- Nodo 4: Servidor Dodo (HTTP).
- Nodos 5 - 10: Clientes.
- Nodo 11: Router, conecta la Intranet a la Internet (cuello de botella).

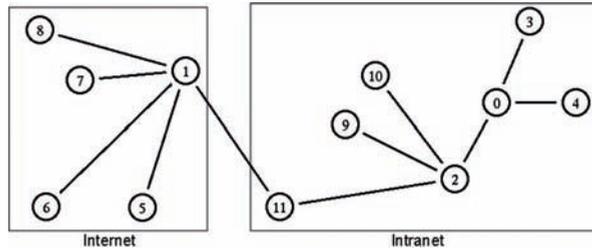


Figura 3. Escenario de simulación en ns2.

Para la simulación se utilizan las mismas pruebas realizadas al tráfico real con el propósito de comparar los resultados posteriormente.

Las características de ancho de banda y retardos utilizados en la simulación se aproximan a los reales de la Intranet. La sección de Internet utiliza retardo y ancho de banda aleatorios para cada nodo. Se utilizan muchos nodos para asegurar diversidad.

El tráfico correspondiente a SMTP en el nodo 3 se simuló mediante la generación, en los nodos cliente, una combinación de tráfico de protocolo Telnet y FTP, donde Telnet se asocia al envío del texto del correo electrónico y FTP a los adjuntos. Las peticiones se hacen en forma aleatoria. El nodo 3 simula el tráfico del servidor 1, el cual utiliza 5 variables aleatorias; 2 se utilizan para controlar los adjuntos que se envían al servidor de correos y 2 para solicitar el adjunto; cada par maneja el tamaño del archivo (distribución tipo Pareto) y la frecuencia (distribución tipo exponencial) a la cual se adjunta o se solicita un archivo; la última variable aleatoria controla la frecuencia (distribución tipo exponencial) con que se realizan las conexiones Telnet. Estas distribuciones típicas para la frecuencia y el tamaño de los archivos se encuentran descritas en [7]. Aquí la distribución de Pareto es importante ya que utiliza dos parámetros para obtener el tamaño de la transmisión, uno es su promedio y el otro es la forma, éste último ajusta el parámetro de Hurst del servidor mediante la ecuación 5.

$$H = 1 - \frac{\beta}{2} \quad (5)$$

El nodo 4 (HTTP) contiene un grupo de recursos con características similares a las del portal web UIS real. El

nodo está programado para recibir peticiones por parte de los nodos cliente. Las peticiones se generan en forma aleatoria. El nodo 4 simula el tráfico del servidor 2.

Este servidor utiliza 10 variables aleatorias, 6 son utilizadas para generar documentos de tamaño aleatorio y las 5 restantes para obtener las características de la página principal. La configuración de la simulación se basa en el tamaño de las páginas UIS reales. En [8] se explica detalladamente el comportamiento de cada una de las variables del tráfico HTTP.

Las variables aleatorias incluyen:

1. Tiempo entre sesiones de clientes, en segundos (exponencial).
2. Consultas realizadas por sesión (uniforme).
3. Frecuencia con que el cliente solicita un nuevo recurso en la sesión, en segundos (exponencial).
4. Cantidad de objetos incrustados en el recurso (constante).
5. Tiempo entre envío de objetos, en segundos (exponencial).
6. Número de paquetes por objeto, paquetes de 1kB (Pareto).

La página principal utiliza las mismas variables excepto la 2 y la 3.

Es importante para la configuración de los nodos servidor utilizar la tasa promedio de transferencia de los servidores reales, para lo cual se utiliza la ecuación 6, donde  $\alpha$  es el tiempo promedio de la distribución exponencial que controla la frecuencia y  $C$  el número de clientes de la red. Estos aspectos se explican con detalle en [9]

$$\bar{P} = \frac{1}{\alpha} (\text{tamaño archivo en bytes}) (C) \quad (6)$$

## 5.2 Autosimilitud en la simulación

A continuación se presentan (tablas 4 y 5) los resultados de la estimación del parámetro de Hurst mediante el periodograma y el estimador de Whittle para determinar la autosimilitud del tráfico agregado en el nodo 0 (N0) y cada uno de los nodos servidor (N3 y N4), de acuerdo con las pruebas mencionadas en 4.1.

Tabla 4. Resultados prueba 1, simulación.

(a) Recepción			
Grano →	0,1	1	10
P	0,875	0,629	0,528
W	0,879	0,722	0,638
P	0,716	0,580	0,545
W	0,805	0,654	0,637
P	0,827	0,654	0,528
W	0,751	0,689	0,638
P	0,867	0,660	0,549
W	0,851	0,715	0,618
(b) Transmisión			
Grano	0,1	1	10
P	1,103	0,655	0,650
W	0,988	0,690	0,678
P	0,811	0,657	0,647
W	0,947	0,730	0,688
P	0,706	0,632	0,601
W	0,784	0,682	0,640
P	0,742	0,611	0,589
W	0,881	0,692	0,674

En la tabla 5 se comparan el promedio de los resultados empíricos y los resultados simulados para la prueba 2.

Tabla 5. Resultados prueba 2, simulación.

(a) Recepción		
	N3	S1
P - 10 s	0,574	0,574
W - 10s	0,712	0,584
P - 0.625s	0,794	0,606
W - 0.625s	0,999	0,859
(b) Transmisión		
	N4	S2
P - 10 s	0,546	0,552
W - 10s	0,560	0,572
P - 0.625s	0,617	0,550
W - 0.625s	0,792	0,920
(c) Transmisión		
	N0	C
P - 10 s	0,538	0,572
W - 10s	0,627	0,599
P - 0.625s	0,692	0,598
W - 0.625s	0,860	0,875

## (b) Transmisión

	N3	S1
P - 10 s	0,625	0,669
W - 10s	0,773	0,648
P - 0.625s	0,849	0,638
W - 0.625s	0,999	0,788
(b) Transmisión		
	N4	S2
P - 10 s	0,620	0,679
W - 10s	0,650	0,698
P - 0.625s	0,747	0,636
W - 0.625s	0,879	0,675
(c) Transmisión		
	N0	C
P - 10 s	0,631	0,683
W - 10s	0,669	0,673
P - 0.625s	0,801	0,645
W - 0.625s	0,946	0,780

La primera prueba presenta características similares a las obtenidas con los datos empíricos, donde el valor del parámetro de Hurst calculado mediante el periodograma y el estimador de Whittle disminuyen con el aumento del tamaño del grano.

En el caso de la segunda prueba, se logran mejores resultados en autosimilitud con la disminución del tamaño del grano. Se observa además que para el grano de 10 segundos los resultados del parámetro de Hurst son próximos entre sí.

### 5.3 Aleatoriedad de la simulación

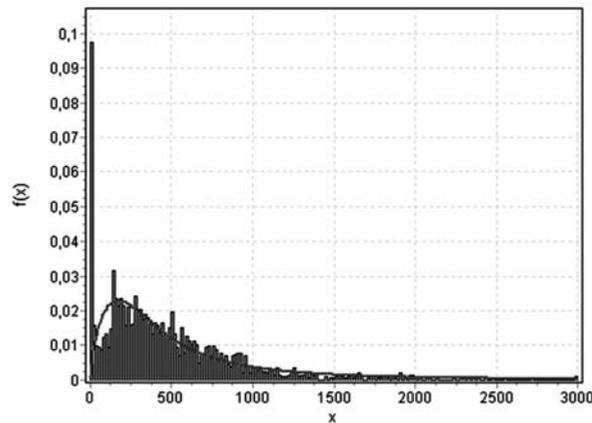
Se realizan las pruebas de bondad de ajuste a los resultados de la simulación con respecto a las distribuciones que presentaron mejor ajuste para los datos de tráfico registrados físicamente. La tabla 6 muestra el error medio cuadrático promedio entre el tráfico simulado y la distribución ajustada, R y T para recepción y transmisión respectivamente.

El resultado final muestra que el nodo 0 presenta un error cuadrático medio menor al 5% con respecto a la distribución empírica. Si se tiene en cuenta que el simulador tiene 15 variables aleatorias, se comprueba que a través de una simulación se pueden obtener resultados muy próximos a los reales.

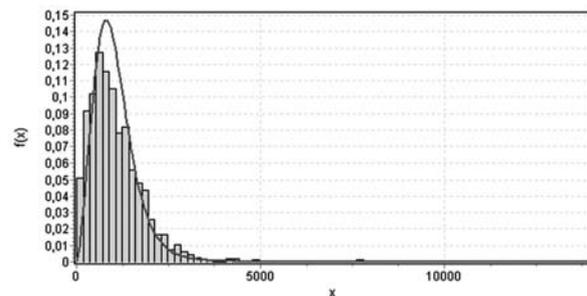
La figura 4 muestra la distribución de una simulación del nodo 3 y el nodo 4, contra la PDF ajustada para los datos de tráfico registrados.

**Tabla 6.** Error cuadrático medio entre los datos simulados y la distribución empírica.

	N3R	N3T	N4R	N4T	N0R	N0T
%e	4.86	7.14	8.75	8.37	5.82	6.14



(a) PDF servidor 1, simulación nodo 3



(b) PDF servidor 2, simulación nodo 4

**Figura 4.** Función de densidad de probabilidad de los servidores contra las simulaciones de los nodos 3 y 4.

## 6. CONCLUSIONES

El parámetro de Hurst puede ser obtenido mediante una variedad de estimadores. En este trabajo en particular el periodograma y el estimador de Whittle permitieron calcular resultados que se ubican dentro de los límites teóricos (0,5 a 1). Otros estimadores suelen sobrepasar estos límites con frecuencia.

Sobre el cálculo y uso del parámetro de Hurst se encontraron varios aspectos valiosos. En la prueba 1, a medida que la escala de tiempo aumenta el valor del parámetro disminuye. Esta situación es limitante para el estudio de la dependencia a largo plazo y no permite caracterizar el tráfico con ese sólo parámetro; en la prueba 2, el tamaño del grano presenta notorias diferencias en los resultados obtenidos con el estimador de Whittle, sin embargo, los resultados del periodograma son más próximos entre sí, teniendo en cuenta esta propiedad del periodograma no se hace necesario utilizar grano pequeño para determinar el valor del parámetro de Hurst. Esta propiedad es muy importante ya que permite hacer el estudio de un nodo con un menor costo computacional.

Debido a que muchos sistemas pueden presentar resultados cercanos en autosimilitud, es necesario hacer un estudio para obtener las características probabilísticas del tráfico, el cual permite conocer su distribución, su media y desviación, datos muy importantes para ajustar las variables aleatorias del simulador.

La correcta simulación del modelo depende de la herramienta de simulación seleccionada. NS-2 tiene excelentes opciones que facilitan el diseño de simulaciones a pesar de utilizar línea de código y no una interfaz gráfica. Con el NS-2 se logró generar tráfico cuyas características de autosimilitud resultaron muy próximas a aquellas de los registros de tráfico empíricos. Adicionalmente, el valor del parámetro de Hurst calculado a partir de las trazas de tráfico simuladas, tiene relación directa con las modificaciones realizadas al del tamaño de los paquetes en cualquiera de los dos tipos de tráfico simulados.

Las simulaciones mostraron para la prueba 1 que a medida que el tamaño del grano aumenta la autosimilitud disminuye, y para la prueba 2 que al disminuir el tamaño del grano el estimador de Whittle genera valores más altos de autosimilitud y que el periodograma es más consistente en sus resultados a pesar del tamaño del grano. Las dos pruebas conservan las mismas propiedades tanto en el tráfico empírico como en el tráfico simulado.

Mediante el cálculo del error cuadrático medio se estimó la diferencia entre las trazas de tráfico generadas por simulación y las distribuciones que mejor ajustaban los datos experimentales. La prueba mostró un error menor al 5% para el servidor SMTP (nodo 3), cercano al 8% para el servidor HTTP (nodo 4) y cercano al 5%

para el nodo 0. Teniendo en cuenta que el sistema utiliza 15 variables aleatorias se considera que los resultados obtenidos con la simulación son excelentes.

## 7. REFERENCIAS

- [1] W. Leland and D. Wilson. *High time-resolution measurement and analysis of lan traffic: Implications for lan interconnection*. Bellcore Labs, 1990.
- [2] W. Willinger, W. Leland, M. Taquu and D. Wilson. *On the self-similar nature of ethernet traffic*. Proceedings of ACM SIGCOMM 93, 1993.
- [3] W. Castellanos, M. Chacón. *Fundamentos teóricos y herramientas software para el modelado y la simulación de redes de comunicaciones*. Reporte Interno. Unidades tecnológicas de Santander. 2005.
- [4] Network Simulator, [citado el 30 de junio, disponible en <http://www.isi.edu/nsnam/>]
- [5] A. Popescu. *Traffic self-similarity*. in Proc. IEEE Int. Conf. Telecommunications, Bucharest, Romania, 1993.
- [6] EasyFit – Distribution Fitting Software, [citado el 30 junio de 2009, disponible en <http://www.mathwave.com/>]
- [7] E. Altman, T. Jiménez. *NS Simulator for Beginners*. Universidad de los Andes, Mérida, Venezuela, and ESSI Sophia-Antipolis, France. 2003.
- [8] M. Molina, P. Castelli, G. Foddis, *Web traffic modeling exploiting TCP connections temporal clustering through HTML-REDUCE*, IEEE Network Vol.14 No.3, 2002. pp.46-55
- [9] E. Altman, T. Jiménez. *NS Simulator for Beginners*. Universidad de los Andes, Mérida, Venezuela, and ESSI Sophia-Antipolis, France. 2003. pp. 46