

# Análisis de las características del tráfico de Escritorios Virtuales mediante el parámetro de Hurst

Pablo Andrés Pessolani  
Departamento de Sistemas de Información - Facultad Regional Santa Fe  
Universidad Tecnológica Nacional - Santa Fe – Argentina  
ppessolani@frsf.utn.edu.ar

**Resumen.** La Computación en la Nube, es un paradigma que permite ofrecer servicios de computación a través de las redes. Una de sus modalidades es el Escritorio como Servicio. Esta tecnología utiliza lo que se denomina Infraestructura de Escritorio Virtual. Dada la creciente adopción de este tipo de servicio, se hace necesario conocer las características del tráfico que genera en las redes y en sus equipos de tal modo de poder realizar su correcto dimensionamiento. Estudios previos sobre otros tipos de servicios demuestran las características autosimilares del tráfico que generan. Esta propiedad se pondera mediante el parámetro de Hurst. Este trabajo contribuye en la estimación de éste parámetro para el tráfico de los servicios de Escritorio Virtual para conocer sus características.

**Keywords:** Cloud Computing, DaaS, VDI, Hurst, LRD.

## 1 Introducción

La computación en la nube (en inglés "*Cloud Computing*"), es un paradigma que permite ofrecer servicios de computación a través de las redes. La "nube" es una metáfora de esa red ya que es la forma en que se la representa en los gráficos técnicos. Todo lo que puede ofrecer un sistema informático se ofrece como servicio[1], de modo que los usuarios puedan acceder a los servicios disponibles "*en la nube*" desde cualquier lugar mediante dispositivos terminales.

Una de las modalidades de la "*Computación en la Nube*" es el "*Escritorio como Servicio*" (en inglés Desktop as a Service o DaaS). DaaS requiere de una Infraestructura de Escritorio Virtual (en inglés Virtual Desktop Infrastructure o VDI) en un Datacenter (propio o tercerizado) que ofrece un escritorio de un sistema como un servicio sin perder el carácter personal y pudiendo ser accedido desde distintos dispositivos, ubicaciones y redes [2].

El concepto de VDI está relacionado con el de Escritorio Remoto (en inglés Remote Desktop o RD). En RD un usuario toma control remoto de una PC física donde se remotiza el mouse, teclado, el display y otros periféricos, pero la computación no se lleva a cabo en su equipo terminal sino en la PC remota controlada. El concepto de VDI es similar, el usuario opera remotamente pero sobre un hardware virtualizado. La tecnología RD no es nueva, pero recientemente ha resurgido reconvertida en VDI como consecuencia de los avances en:

1. *La virtualización de hardware*: Este concepto que deviene de los años 70, formalizado por Popek y Goldberg en [3] y que, con las mayores capacidades de

cómputo y almacenamiento de la actualidad, permite consolidar múltiples servicios y obtener una multiplicidad de beneficios, tales como ahorro de energía, menores costos de mantenimiento, períodos mas prolongados de disponibilidad de servicio, etc.

2. *Las redes de comunicaciones*: Los desarrollos en LANs, WANs, WLANs, etc. han permitido alcanzar mayores tasas de transferencia, mayor cantidad de puntos de presencia y dispersión, y mayor confiabilidad. Estas redes fueron diseñadas, para soportar aplicaciones de diferente naturaleza, algunas de las cuales requieren un servicio de transporte confiable, otras son exigentes en cuanto a otros parámetros de calidad de servicio tales como el retardo (delay), el caudal (throughput), la tasa de pérdidas de tramas o paquetes, etc. El desarrollo de nuevas tecnologías ha permitido la evolución de las redes, dotándolas de la capacidad de satisfacer simultáneamente los requisitos de tráfico de muy diversa naturaleza.

El tráfico actual de las redes de datos es consecuencia del uso de una multiplicidad de aplicaciones, entre ellas las más básicas tales como el correo electrónico, el chat, la transferencia de archivos, terminales remotas con telnet/ssh, la navegación web, etc.; de las aplicaciones corporativas como intranets, ERPs, CRMs, etc. que pueden estar basadas en diferentes modelos cliente/servidor, y las más recientes aplicaciones de videoconferencias, trabajo colaborativo, telepresencia, redes sociales, Video bajo Demanda (VoD), Voz Sobre IP (VoIP), etc.

La tecnología VDI es exigente con los protocolos y con las redes de transporte. Se requiere de baja latencia y jitter (variabilidad de la latencia), suficiente ancho de banda y reducida tasa de pérdida de paquetes para que el usuario no experimente retrasos en sus tiempos de respuesta que resulten frustrantes mientras desarrolla su trabajo. En un estudio realizado por IBM [4, 5], quedó demostrado que cuando el tiempo de respuesta promedio se reduce de 2,2 segundos a 0,80 segundos por transacción, la productividad del usuario se duplica. Los tiempos de respuesta prolongados hacen que los usuarios pierdan el hilo de pensamiento sobre el trabajo que se encuentran realizando.

Los requerimientos de calidad de servicio de las aplicaciones no pueden ser satisfechos usando protocolos de alto nivel si las redes de transporte no ofrecen esas garantías. Las nuevas redes deben ser capaces de dar un servicio que garantice determinados parámetros de calidad de servicio para ser utilizada por una amplia variedad de nuevas aplicaciones. Una de esas aplicaciones es VDI que requiere de exigentes parámetros de calidad tanto para los servidores de virtualización como para las redes de datos.

Las empresas, gobiernos y otros organismos utilizan el VDI para el acceso a los sistemas corporativos desde diferentes localizaciones. Actualmente su adopción viene en constante crecimiento para su utilización en el teletrabajo para acceder a aplicaciones corporativas, en reemplazo del uso de Redes Privadas Virtuales (VPNs) para acceder a las redes corporativas.

Uno de los principales problemas que se plantean a la hora de realizar una evaluación rigurosa de las prestaciones de una red de datos es el modelado del tráfico de entrada a la red [6]. El modelado del tráfico es el más crítico de los problemas relacionados con la evaluación de prestaciones de redes de datos, ya que el éxito del análisis depende en gran medida de lo representativos de la realidad que sean los modelos de tráfico utilizados.

El modelado del tráfico tiene sus orígenes en la telefonía convencional, y se ha basado casi exclusivamente en suponer independencia en los tiempos entre llegadas (modelos de Poisson), y distribuciones exponenciales de los tiempos de uso de los recursos. La razón más importante de tales suposiciones se haya en poder obtener modelos relativamente simples desde el punto de vista analítico. Se ha logrado determinar en base a estudios y observación de exhaustivas mediciones realizadas sobre una red Ethernet [7] que el tráfico (número de tramas por unidad de tiempo) era de naturaleza fractal, “autosimilar” o “autosemejante” (self-similar), que implica la existencia de una correlación a largo plazo (Long-Range Dependence - LRD) [8]. Es importante aclarar que el término autosimilitud refiere a los parámetros estadísticos y se lo suele referir con las siglas SSS (del inglés Statistical Self-Similarity).

Con el descubrimiento de fenómenos de autosimilitud en el tráfico de las redes actuales, se han reportado evidencias de autosimilitud tanto en LANs como WANs, bajo protocolo IP o ATM; con enlaces de cobre [9], fibra óptica o inalámbrica; en la navegación por la Web [10] o en la transferencia de archivos. La autosimilitud y la fractalidad describen el fenómeno en el que cierta propiedad de un objeto se preserva con respecto a las diferentes escalas en el tiempo o en el espacio. En un objeto autosimilar o fractal, sus partes magnificadas se asemejan a la forma del objeto completo.

Conforme mejoran los modelos, se pueden desarrollar soluciones más eficientes para reducir la pérdida de paquetes en los sistemas de encolamiento y es considerado como una herramienta de evaluación de performance para la gestión de buffers en los equipos de conmutación (routers, switches, etc.).

En este trabajo se estudia la caracterización del tráfico autosimilar y la estimación del parámetro fundamental de Hurst (H) para la tecnología VDI mediante la captura de tráfico en la red de una corporación que la utiliza para teletrabajo y donde las aplicaciones más frecuentemente utilizadas son las de oficina, tales como edición de documentos, planillas electrónicas, presentaciones, etc., correo electrónico corporativo, aplicaciones CRM, ERP, navegación web en internet e intranets.

Debe considerarse que se realizaron diferentes capturas de paquetes en la red durante períodos de *aproximadamente 30 minutos de un día laboral normal*, y que la misma fue limitada a ese tiempo debido a la inmensa cantidad de datos generados en el proceso de captura como se demuestra en este artículo.

Existen varias tecnologías para brindar el servicio VDI, típicamente Citrix XEN Desktop, RealVNC, pcAnywhere, etc. Este trabajo se realizó sobre una infraestructura basada en Microsoft Terminal Services que utiliza protocolo RDP.

Los resultados de este trabajo no deben considerarse concluyentes respecto a las características autosimilares del tráfico VDI. El valor de este trabajo radica en el procesamiento y análisis de los resultados para esta captura en particular, y las herramientas y metodologías utilizadas para realizar este análisis.

El presente artículo está organizado de la siguiente forma; en la [Sección 2](#) se realiza una introducción a los procesos autosimilares y a los diferentes métodos para estimar el valor del parámetro de Hurst que es el indicativo del grado de autosimilitud. En la [Sección 3](#) se presenta información de la topología de la red objeto de las capturas y de los datos capturados. En la [Sección 4](#) se realiza el análisis de autosimilitud de las aplicaciones VDI y de los resultados obtenidos. Finalmente, en la [Sección 5](#) se resumen las conclusiones del trabajo realizado.

## 2 Procesos Autosimilares

Las redes de datos pueden analizarse utilizando el modelo de teoría de colas y redes de colas [11]. Los parámetros que caracterizan el comportamiento de las colas son:

- *Tasa de Arribo o Tasa de Llegadas*: La tasa media de llegadas es el número esperado de llegadas por unidad de tiempo que están ingresando al sistema.
- *Intensidad de Tráfico o Factor de Utilización*: Es una variable que relaciona la tasa de llegadas con la tasa de servicio, ó sea es el número de paquetes o tramas que completan su servicio por unidad de tiempo. Dada la tasa media de llegadas y la tasa media de servicio, se define el factor de utilización del sistema.
- *Caudal (Throughput) o Tasa de Transferencia*: Es la Tasa promedio de atención de paquetes o tramas en el sistema de colas. Esta variable mide la proporción con que son atendidos los elementos en el sistema.
- *Longitud del Sistema*: Es la longitud promedio de una cola e incluye todos los elementos (tramas, paquetes, etc.) 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 elemento en el sistema. Este corresponde al tiempo en que el elemento está en la cola esperando ser servido más el tiempo de servicio.

Un proceso autosimilar es aquel que tiene el mismo comportamiento cuando se visualiza a diferentes escalas de una dimensión específica ya sea longitud, ancho o tiempo.

Los modelos clásicos no logran capturar el comportamiento real del tráfico de las redes de datos y son en general, optimistas. En múltiples estudios se determinó que el proceso de Poisson no aparece en las redes de datos de la actualidad, por lo que dió a lugar a la aparición del proceso autosimilar. En estos estudios se muestra la naturaleza autosimilar del tráfico de las redes y de la mayoría de los servicios que se utilizan sobre ellas [7, 8, 9, 10]. La autosimilitud está presente en tráfico de Ethernet, WWW, SS7, TCP, FTP, TELNET, Video en VBR.

### 2.1 Propiedades de los Procesos Fractales

Las propiedades de los Procesos Fractales según [6]:

Autosimilitud: La autosimilitud exacta se define como la apreciación de una estructura estadísticamente similar a lo largo de un amplio margen de escalas temporales, es decir, la señal mantiene su distribución estadística aunque se varíe la escala de observación. El movimiento browniano fraccionario es el ejemplo más conocido de procesos autosimilares. El parámetro de Hurst (H) mide el grado de autosimilitud de los procesos. Su valor oscila entre 0,5 y 1, siendo el proceso más autosimilar cuanto más se acerque H a 1. Un proceso exactamente autosimilar es no estacionario por definición.

Dependencia de Largo Rango: Una de las propiedades más significativas de los procesos autosimilares es referida a la dependencia de rango largo. Esta propiedad es definida en términos de la conducta de la autocovarianza cuando  $t$  (generalmente el tiempo) se incrementa. La LRD refleja el fenómeno de persistencia en un proceso

autosimilar y es principalmente definido en términos del comportamiento de la función de autocorrelación.

Un proceso estocástico presenta LRD si su autocorrelación decae a un ritmo lento (inferior al de la exponencial). El tráfico medido en un cierto instante  $t_0$  y el medido en un instante posterior  $(t_0 + \tau)$  mantienen una correlación importante aunque la separación temporal  $\tau$  sea muy elevada; fenómeno que no se da en los procesos de llegadas de Poisson. Por ello se dice que los procesos LRD tienen memoria.

La definición de los procesos LRD establece que su densidad espectral satisface:

$$S(f) = \frac{cf}{|f|^\alpha} \quad \text{Para } f \rightarrow 0$$

Donde el parámetro de escalado es una medida cualitativa de la intensidad del fenómeno LRD y se expresa habitualmente reescribiendo el parámetro de Hurst  $H = (1+\alpha)/2$ , mientras que  $cf$  tiene dimensiones de segundo momento o varianza y puede ser interpretado como una medida cuantitativa de la LRD.

Decaimiento Lento de la Varianza: La varianza de las muestras decaen más lentamente que el recíproco del tamaño de las muestras. Por ejemplo:

$$Var [x^{(m)}] = \frac{Var(x)}{m^\beta} \quad \text{Si decimos que } Var [x] \approx a_2$$

Entonces:

$$Var [x^{(m)}] \approx a_2 m^{2H-2} \quad \text{como } m \rightarrow \infty$$

Para  $H = 0.5$ , tenemos que  $\beta = 1$  lo cual es un proceso (SDR) dependencia de corto rango, tendremos:

$$Var [x^{(m)}] \approx a_2 m^{-1}$$

Distribuciones De Colas Pesadas: Estas distribuciones son usadas para caracterizar densidades de probabilidad que describen el proceso de tráfico como paquetes a intervalos de tiempo y longitud de ráfagas.

La distribución de una variable aleatoria  $X$  se dice que es de cola pesada, si:

$$1 - F(x) = \Pr[X > x] \approx \frac{1}{x^\alpha} \quad x \rightarrow \infty ; 0 < \alpha$$

Exhibe una alta varianza, que podría ser infinita.

La distribución de cola pesada más simple es la distribución de Pareto con parámetros  $k$  y  $\alpha$  ( $k, \alpha > 0$ ), con una densidad y una función de distribución:

$$f(x) = F(x) = 0 \quad (x \leq k)$$

Siendo:

$$f(x) = \frac{\alpha}{k} \left( \frac{k}{x} \right)^{\alpha-1}$$

Con valor de media:

$$E[x] = \frac{\alpha}{(\alpha-1)} k \quad (\alpha > 1)$$

El parámetro  $k$  especifica el mínimo valor que la variable aleatoria puede tener. El parámetro  $\alpha$  determina la media y la varianza de la variable aleatoria: si  $\alpha \leq 2$ ,

entonces la distribución tiene varianza infinita; y si  $\alpha \leq 1$ , tiene media y varianza infinitas. La propiedad de distribución de cola pesada no es una condición necesaria para la autosimilitud. Sin embargo, la naturaleza autosimilar de muchas trazas de tráfico resulta directamente de distribuciones de cola pesada.

Densidad Espectral: La LRD también se puede dar en el dominio de la frecuencia. La densidad espectral de potencia obedece a una ley de potencia cerca del origen:

$$S(\omega) \approx \frac{1}{|\omega|^\gamma} \quad \text{con } \omega \rightarrow 0 \text{ y } 0 < \gamma < 1$$

Donde  $\gamma = 1 - \beta = 2H - 1$ . Esta divergencia de la densidad espectral de potencia cerca al origen indica una importante contribución de los componentes de baja frecuencia. La densidad espectral para un proceso estocástico discreto en el tiempo se define de la siguiente manera:

$$S(\omega) = \sum_{K=-\infty}^{\infty} R(k)e^{-j2k\omega} \quad S(0) = \sum_{K=-\infty}^{\infty} R(k)$$

Dimensión Fractal: Un fractal es una figura geométrica con una estructura compleja y pormenorizada a cualquier escala. Normalmente los fractales son autosemejantes, es decir, tienen la propiedad de que una pequeña sección de un fractal puede ser vista como una réplica a menor escala de todo el fractal. Donde el número de objetos autosimilares es igual a:

$$N = r^{-D}$$

Teniendo en cuenta que  $r$  es el tamaño de la escala lineal de los objetos cubiertos y  $D$  la dimensión Fractal, visto de otra manera; La dimensión fractal es:

$$D = \frac{\log N}{\log \frac{1}{r}}$$

Los procesos autosemejantes o autosimilares pueden ser modelados por procesos con dimensiones fractales. Esta es una consecuencia de la lenta y muy lenta convergencia del proceso a valores medios.

## 2.2 Métodos de Estimación del Parámetro de Autosimilitud H

En esta sección describen los métodos para la estimación del parámetro de Hurst implementados en la herramienta SELFIS [12]:

1. *Método del Valor Absoluto*: Se realiza una grafica log-log del nivel de agregación o sumarización versus el valor absoluto del primer momento de las series sumarizadas  $X(m)$ . Si los datos tienen una dependencia de largo plazo (LRD), esta gráfica debería ser una línea recta con una pendiente  $H-1$ .
2. *Método de la Varianza*: Se realiza una grafica log-log de la varianza de la muestra versus el tamaño del bloque de cada agregación. Si los datos tienen una dependencia de largo plazo (LRD), esta gráfica debería ser una línea recta con una pendiente  $\beta > 1$ . El párametro H estimado está dado por  $H = (1 + \beta/2)$ .

3. *Método R/S*: Este método usa un rescalado del rango estadístico (estadística R/S). La estadística R/S es el rango de las sumas parciales de las desviaciones de su media de una serie de tiempo, reescalada por su desviación estándar. Se realiza una grafica log-log de la estadística R/S versus el número de puntos de la serie sumariada. Si los datos tienen una dependencia de largo plazo (LRD), esta gráfica debería ser una línea recta con una pendiente H.
4. *Método del Periodograma*: Este método realiza una grafica log-log de la densidad espectral de una serie de tiempo versus las frecuencias. La pendiente es una estimación del parámetro H.
5. *Método de Whittle*: Está basado en la minimización de la función de probabilidad, la cual es aplicada al periodograma de una serie de tiempo. Brinda una estimación del parámetro H y un intervalo de confianza. No produce un resultado gráfico.
6. *Método de la Varianza de los Residuos*: Este método realiza una grafica log-log del nivel de sumariación versus el promedio de la varianza de los residuos de las series. Debería ser una recta con pendiente H/2.
7. *Método Abry-Vitch*: Se utiliza la transformada onditas (wavelets) para estimar el exponente de Hurst. Se utiliza la energía de las series en varias escalas para estimar el parámetro H.

### 2.3 Protocolo RDP

El protocolo RDP (Remote Desktop Protocol) [13] fue diseñado para facilitar la interacción del usuario con una computadora remota transfiriendo información de pantallas gráficas desde la computadora remota hacia la computadora del usuario y transportando los comandos del usuario hacia la computadora remota.

El RDP también dispone de mecanismos que le permiten transportar comunicaciones entre componentes de la computadora en la que se encuentra el usuario y la computadora remota (ej: sonidos, videos, el Clipboard, impresoras [14]). En las conexiones de reducido velocidad, estas aplicaciones compiten por el ancho de banda disponible. Como resultado de esto compiten los gráficos importantes para la percepción del usuario con datos que circulan en el background (trasfondo) tales como los trabajos de impresión o la copia de un archivo. Esto produce un deterioro significativo el rendering de los gráficos y por lo tanto en el disgusto por parte del usuario. Actualmente el protocolo soporta Circuitos Virtuales que permiten priorizar las transferencias de aquellos gráficos que mejoran la percepción del usuario.

## 3 Mediciones de Tráfico

Para realizar la captura de paquetes de la red se utilizó TCPDUMP con un filtro, de tal forma que solo capturen paquetes con puertos orígenes o destino iguales a RDP (3398) e ignorar otro tipo de tráfico y así reducir el tamaño de los archivos de captura.

Se realizó un análisis del tráfico de una única sesión VDI entre un cliente y un servidor dentro del tráfico capturado (Sesión Simple) y con todas las sesiones VDI que se capturaron (Sesiones Múltiples).

### 3.1 Red Objeto de Captura

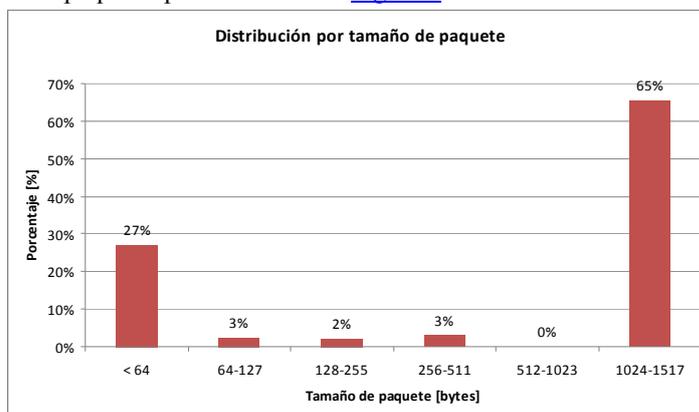
Para desarrollar el presente trabajo, se realizaron mediciones de tráfico en una compañía dedicada a servicios TICs de aproximadamente 15.000 empleados. Esta compañía tiene aproximadamente 500 empleados en la modalidad de teletrabajo utilizando la tecnología VDI. Los teletrabajadores conectan las PCs clientes de sus domicilios particulares a una Red Privada Virtual (VPN) con tecnología MPLS mediante conexiones hogareñas con ADSL.

### 3.2 Características del Tráfico Capturado

Se utilizó la herramienta TCPTRACE para realizar un resumen de las características del tráfico capturado. Los siguientes son los resultados arrojados por la herramienta para uno de los conjuntos de muestras.

```
Port 3389 bytes:195223199 pkts:4581923 conns:18 tput:111492 B/s
Overall Statistics over 1751 seconds (0:29:11.748641):
4490190495 ttl bytes sent, 2564357.702 bytes/second
4489949809 ttl non-rexmit bytes sent, 2564220.272 bytes/second
240686 ttl rexmit bytes sent, 137.456 bytes/second
4581923 packets sent, 2616.746 packets/second
18 connections opened, 0.010 conns/second
149 dupacks sent, 0.085 dupacks/second
average RTT: 8.027 msec
```

También se utilizó la herramienta CAPSA para realizar un resumen de las características del tráfico capturado con la que se construyó el gráfico de la distribución de paquetes por tamaño de la [Figura 1](#).

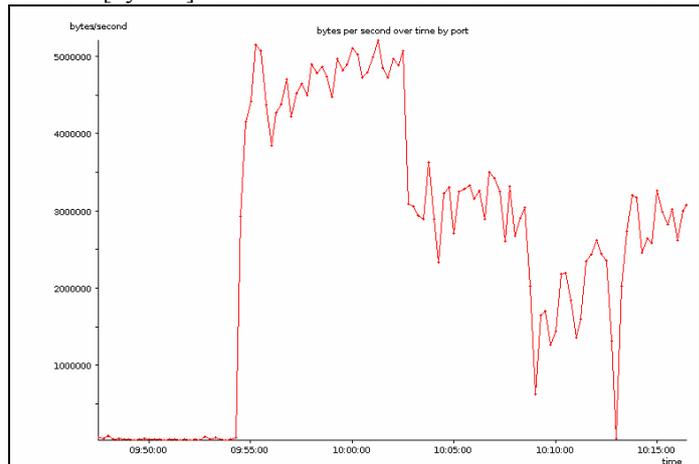


**Figura 1:** Distribución por Tamaño de Paquete.

Se puede apreciar que el número de paquetes de tamaño menor a 64 bytes representa el 27% de los paquetes cursados por la red que son consecuencia, mayoritariamente, de los reconocimientos (Acknowledges) del protocolo TCP, de los movimientos del mouse, y de las teclas pulsadas por el usuario con el algoritmo de Nagle deshabilitado [15]. Los paquetes con tamaños mayores a 1023 bytes, se deben

principalmente a las actualizaciones de los gráficos de los escritorios remotos y representan el 65% de los paquetes cursados por la red.

Utilizando las herramientas TCPTRACE y XPLOTT se realizaron las gráficas del tráfico. Se filtraron los datos solo para el puerto RDP. En la [Figura 2](#) se presenta el ancho de banda en [bytes/s] en la LAN de los servidores VDI de la Red de Captura.



**Figura 2:** Ancho de Banda en la LAN de los Servidores VDI [bytes/s].

### 3.3 Estimación del Parámetro de Hurst

En la [Tabla 1](#) se presentan los parámetros estadísticos de los dos conjuntos de muestras (0406 y 0709) para una Sesión Simple (SS) y para todo el conjunto de Sesiones (SM).

**Tabla 1:** Características del Tráfico Capturado (Tamaño de Paquete) obtenido con CAPSA

<b>Parámetro Estadístico</b>	<b>SS0406</b>	<b>SS0709</b>	<b>SM0406</b>	<b>SM0709</b>
<b>Media</b>	997,5	992,8	980,5	976,4
<b>Varianza</b>	42821	41967	43635	42734
<b>Desv. Estándar</b>	654,3	647,8	660,5	653,7
<b>Sesgo</b>	-0,661	-0,639	-0,607	-0,587
<b>Kurtosis</b>	1,552	1,533	1,477	1,463

Para realizar el análisis de autosimilitud mediante el parámetro de Hurst se utilizaron dos herramientas diferentes: SELFIS y R+LRDcode.

Los resultados de las estimaciones de R+LRDcode para la variable aleatoria  $T(i)$  donde  $T$  es el Tamaño de Paquete e  $i$  es el orden o número de secuencia de arribo se presentan a continuación:

```
RS plot H= 0.6234302
Agg var H= 0.5593731
Period. H= 0.3255724
Wavelet H= 0.4592094 +/- 0.07929455
Loc. W. H= 0.8015381
```

Los resultados de las estimaciones de SELFIS para la variable aleatoria  $T(t)$  donde  $T$  es el Tamaño de Paquete y  $t$  es el tiempo de arribo se presentan en la [Figura 3](#).

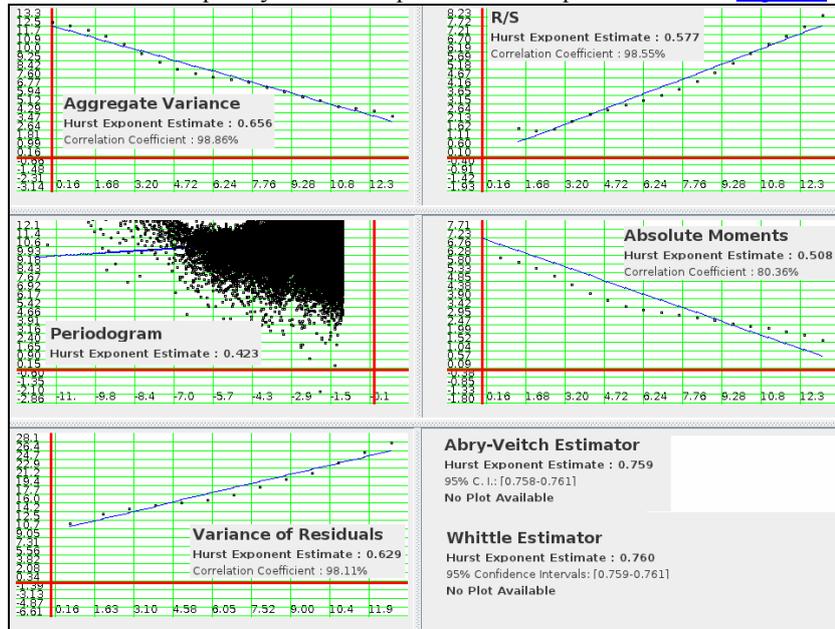


Figura 3: Estimación de Parámetro de HURST realizado con SELFIS

#### 4 Análisis y Resultados

Los resultados arrojados por SELFIS para diferentes muestras se resumen en la [Tabla 2](#), donde se pueden comparar los valores estimados del parámetro de Hurst.

Tabla 2: Resumen de los Valores Estimados del Parámetro de Hurst para VDI por SELFIS

Muestra	AggVar	Period.	VarRes	Whittle	R/S	AbsMom	A/V
SS0406	0,751	0,392	0,572	0,777	0,610	0,565	0,761
SS0709	0,621	0,384	0,573	0,750	0,528	0,470	0,742
SM0406	0,800	0,421	0,642	0,780	0,653	0,620	0,770
SM0709	0,656	0,423	0,629	0,760	0,577	0,508	0,759

SS0x0y: Sesión Simple.

SM0x0y: Sesiones Múltiples.

AggVar: Método de la Varianza Sumarizada.

Period.: Método de Periodograma.

VarRes: Método de la Varianza de los Residuos

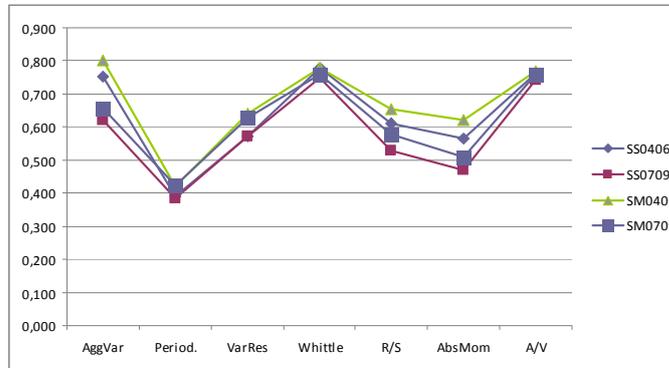
Whittle: Método de Whittle.

R/S : Método R/S.

AbsMom: Método de Momentos Absolutos

A/V: Método Abry-Vitch

Llevando los valores a la gráfica de la [Figura 4](#) se puede apreciar la marcada diferencia entre las estimaciones de los distintos métodos.



**Figura 4:** Comparación de las estimaciones del parámetro de HURST por SELFIS.

Trasladando los resultados arrojados por R+LRDcode a la [Tabla 3](#), se pueden comparar los valores estimados del parámetro de Hurst.

**Tabla 3:** Resumen de los Valores Estimados del Parámetro de Hurst para VDI por R+LRDcode

Muestra	AggVar	Period.	VarRes	Whittle	R/S	AbsMom	A/V
<b>M04-06</b>	0.609	0.295	---	0.812	0.632	---	0.444
<b>M07-09</b>	0.463	0.353	---	0.788	0.592	---	0.478

La diferencia de resultados no debe sorprender dado que en [16] y [17] los autores arribaron a resultados de similares características. En [17] se realizó un estudio de la calidad de los estimadores, definiendo a un “*buen estimador*” no solo a aquel que produce un valor estimado lo más próximo al real (sobre muestras conocidas con valor exacto de H), sino también aquel que tiene la menor varianza [18, 19].

## 5 Conclusiones

La tecnología VDI es exigente con las redes, particularmente las WANs, por lo que presentará importantes desafíos para los diseñadores de equipamiento, para los ingenieros de tráfico, para los proveedores de servicios, etc.

El modelado del tráfico es una herramienta fundamental para el diseño y la evaluación de comportamiento de redes y equipamiento. Actualmente la comunidad científica y tecnológica ha adoptado el modelo de tráfico autosimilar como aquel que mejor refleja el comportamiento del tráfico. La caracterización de la autosimilitud mediante el parámetro de Hurst del tráfico VDI es indispensable entonces para la confianza en el modelo. En [18] se analiza el impacto el parámetro de Hurst la tasa pérdida de paquetes para distintas tasas de utilización de una red.

Vistos los resultados divergentes arrojados por los distintos métodos de estimación del parámetro de Hurst, no se está en condiciones de afirmar cuál de esos métodos es el más confiable [16]. Según el criterio antes mencionado el estimador que arroja resultados más confiables es el método de Whittle [17, 19].

Considerando los valores resultantes del estimador de Whittle, una conclusión sobre el valor del parámetro de Hurst para las muestras de tráfico utilizando SELFIS

(variable aleatoria  $T(t)$ ) es  $0,750 \leq H \leq 0,780$  y utilizando R+LDRcode (variable aleatoria  $T(i)$ ) es  $0,788 < H < 0,812$ . Se puede concluir entonces, que el tráfico VDI tiene propiedades autosimilares a diferentes escalas.

En términos prácticos este trabajo contribuye a realizar una correcta estimación de las características del tráfico para dimensionar redes y equipamientos necesarios para el servicio VDI independientemente de la escala.

## Referencias

1. *Las Capas de Servicios Arquitectónicos del Cómputo Cloud*.  
<http://www.itnews.ec/marco/000037.aspx>
2. Jeff Fisher's "Desktops as a Service" Blog. Desktops as a Service (DaaS) and Cloud Computing – An Overview.
3. Gerald J. Popek and Robert P. Goldberg (1974). «*Formal Requirements for Virtualizable Third Generation Architectures*». Communications of the ACM 17 (7): pp. 412–421.
4. Lambert, G. N.; *A comparative study of system response time on program developer productivity*; IBM Systems Journal, 1984.
5. A. J. Thadhani, *Factors affecting programmer productivity during application development*; IBM Systems Journal, 1984.
6. Porras Castro, Stalin Efrain, Rivera Moreno, Paulo Cesar, *Análisis del algoritmo algebraico de asignación dinámica del buffer para sistemas de encolamiento bajo condiciones de tráfico autosimilar*, 2008.
7. Will E. Leland, Murad S. Taqqu, Walter Willinger, Daniel V. Wilson, *On the Self-Similar Nature of Ethernet Traffic*, March 3, 1993.
8. Vern Paxson, Sally Floyd, *Wide-Area Traffic: The Failure of Poisson Modeling*.
9. Sandra Liliana Cajamarca Guevara, Camilo Torres Zambrano; *Análisis y modelado de tráfico de datos con aplicación a redes de acceso ADSL*; Pontificia Universidad Javeriana, Septiembre de 2002.
10. Gagandeep Kaur, Dr. Vikas Saxena, Prof. J. P. Gupta; *Characteristics Analysis of Web Traffic with Hurst Index*; Proceedings of the World Congress on Engineering and Computer Science 2010 Vol I WCECS 2010, October 20-22, 2010, San Francisco, USA.
11. John N. Daigle; *Queueing Theory With Applications To Packet Telecommunication*; 2005 Springer Science; ISBN 0387228578
12. Thomas Karagiannis, Michalis Faloutsos; *SELFIS: A Tool For SelfSimilarity and LongRange Dependence Analysis*.
13. MS-RDPBCGR, *Remote Desktop Protocol: Basic Connectivity and Graphics Remoting Specification*, Microsoft Corporation, March 18, 2011.
14. *Remote Desktop Protocol Performance - Presentation and Hosted Desktop*; Microsoft; October 2008.
15. Kevin Burns; *TCP/IP Analysis and Troubleshooting Toolkit*; 2003; Wiley Publishing, Inc.; ISBN: 0-471-42975-9.
16. Richard G. Clegg; *A Practical Guide to Measuring the Hurst Parameter*; 2005.
17. Hae-Duck Joshua Jeong, *Modelling of self-similar teletraffic for simulation*, PhD thesis in Computer Science in the University of Canterbury, 2002.
18. Richard G. Clegg; *The Statistics of Dynamic Networks*; PhD Thesis; University of York; 2004.
19. Christian Grimm., Georg Schluchtermann; *IP Traffic Theory and Performance*; Springer-Verlag; ISBN 978-3-540-70603-8; 2008.