Validación de Modelos de Tráfico sobre Enrutadores con QoS

Bocalón Franco, Ortíz Yanina, Salina Noelia,
Docentes: Corteggiano Fernando, Gioda Marcelo

Catedra: Tráfico, Departamento de Telecomunicaciones, Facultad de Ingeniería,
Universidad Nacional de Rio Cuarto

Resumen: Los modelos de tráfico aplicados a redes son importantes a la hora de desarrollar e implementar dispositivos o sistemas de red para una determinada aplicación y realizar la medición del desempeño de los mismos, comparando el comportamiento real con el predicho por los modelos teóricos. Este trabajo consiste en la implementación de una topología de red, configurando diferentes escenarios que corresponden a distintos modelos de tráfico, para una posterior medición del desempeño en base al retardo comparándolo con su modelo asociado. Para evitar la sincronización de los terminales se propone un algoritmo que permite medir el retardo bajo ciertas condiciones de trabajo.

Keywords: Modelo de tráfico de red, Medición de retardo, Teoría de Colas, Calidad de Servicio

1 Introducción

Los fenómenos reales, comenzando por los más básicos, han sido estudiados y comprendidos por el conjunto científico pudiendo así enumerar leyes para su explicación, logrando predecir el accionar de los mismos. En la actualidad los fenómenos a predecir y controlar son de una complejidad y abstracción mayores, sin quedar exentas las redes de comunicaciones.

Una red de comunicación, es un recurso finito por ende es imperiosa la aplicación de un criterio que regule de manera óptima las variables de control para hacer uso eficiente del sistema de telecomunicación.

Como en otras aplicaciones reales, estos sistemas han sido modelados estableciendo leyes de predicción de fenómenos observables; y como todo proceso de modelaje, se imponen considerandos truncando la magnitud de la observación del problema real.

Los modelos de tráfico son la herramienta inicial para el estudio de sistemas de redes de telecomunicaciones; permiten en un determinado escenario, establecer predicciones sobre rendimientos, factores de carga, retardos de transmisión, perdida de información, entre otros datos que garantizan la calidad del sistema [2].

En el presente trabajo se estudian modelos de tráfico aplicados en un escenario teórico, resolviéndose analíticamente cada modelo e implementándose en paralelo una topología de red entre dos terminales a través de un enrutador con capacidad de

implementar QoS (Calidad de Servicio)[6], es decir, manejo de colas con prioridades, reguladores de ancho de banda, entre otros; basados en las herramientas "Traffic Control" de Linux [10] . El tráfico se genera con el software D-ITG -Distributed Internet Traffic Generator- [1],[4],[5],[7]. A través de esta herramienta se miden los parámetros reales de desempeño en cada experiencia. Al sistema en conjunto se lo configura conforme al esquema planteado respectivo, comparándose en cada escenario las mediciones reales, con los cálculos teóricos de cada modelo.

Adicionalmente a los objetivos iniciales del presente, se propuso un algoritmo de corrección en la medición de retardos en la red, evitándose la complejidad técnica de la sincronización de los terminales con dispositivos externos adicionales [3].

2 Aspectos Teóricos

Los modelos de colas se utilizan para caracterizar el comportamiento de la cola de salida de un enrutador, el tráfico que arriba a dicha cola tendrá una distribución de tiempos de interarribo de paquetes exponencial y la longitud de los mismos también tendrá una distribución exponencial [8],[9],[11].

El primer análisis consiste en utilizar la cola de salida del enrutador tipo FIFO y la velocidad del servidor de salida (velocidad del enlace) regulada, el modelo que representa este escenario es el M/M/1 que está caracterizado por:

- Datos que arriban al sistema acorde a un proceso de Poisson (con tasa λ)
- Longitudes de datos que son independientes y están idénticamente distribuidas acorde a la distribución exponencial con media λ. Por lo tanto los tiempos de servicio tendrán también distribución exponencial con sus correspondientes primer y segundo momento E[x] y $E[x^2]$ iguales a $\overline{X}$ y $2\overline{X}^2$ respectivamente. En donde $\overline{X} = 1/\mu$
- Existe un único servidor s = 1.

Se define entonces el factor carga del sistema ρ como

$$\rho = \lambda/\mu \tag{1}$$

El número de mensajes en el sistema, se define como:

$$\overline{N} = \rho + \frac{\lambda^2 \overline{X}^2}{(1-\rho)} \tag{2}$$

A partir de la formula de Little (que relaciona el numero de mensajes en la cola, la tasa de arribos y el retardo que sufren los paquetes en la cola)

$$\overline{D} = \frac{N}{\lambda} \tag{3}$$

Se obtiene la fórmula de Pollaczek-Khinchin para el retardo que sufre un mensaje al atravesar el sistema:

$$\overline{D} = \frac{1}{(\mu - \lambda)} \tag{4}$$

El segundo análisis consiste en utilizar una cola con prioridad estricta a la salida del enrutador y la velocidad del servidor de salida (velocidad del enlace) regulada, el modelo que representa este escenario es el M/G/1 con prioridades y sin apropiación. Esta estrategia se caracteriza por contar con un simple servidor al cual arriban paquetes con P clases de prioridades, numeradas de 1 a P. Siendo la clase 1 la de mayor.

El tiempo medio de espera para la clase r, al igual que para el tiempo de servicio, quedará definido por E[Wr]. El tiempo medio de servicio, para la clase r, se define como el primer momento de la distribución del tiempo de servicio E [Sr]. Cuando un paquete de clase r arriba a la cola, el tiempo que debe esperar antes de ser servido depende de:

-El tiempo de espera de un paquete clase r
-El tiempo remanente de un paquete en servicio
-El tiempo que lleva servir las k clases que están presentes cuando arriba un cliente clase r
-El tiempo que lleva servir las k clases que arriban durante Wr y que necesitan ser servidos antes que el cliente clase r

Expresado en forma de ecuación el tiempo medio de espera:

$$E[W_r] = \frac{\frac{1}{2}\sum_{k=1}^{P} \lambda_k E[S_k^2]}{(1-\sigma_r)(1-\sigma_{r-1})} \tag{5}$$

Un paquete clase r está en servicio con probabilidad (utilización debida a los paquetes de clase r):

$$\rho r = \lambda E[Sr] \tag{6}$$

Debido a que la estrategia en las colas sin apropiación es cambiar el orden de servicio de los paquetes que arriban. Para expresar el efecto que causa la asignación

de prioridades sobre los paquetes de las diferentes clases, se cita la ley de conservación [6], la cual expresa que la cantidad de trabajo realizada por todo el sistema se mantiene a causa de que la sumatoria de los tiempos de espera promedio por clase, pesados por su utilización se mantiene, independiente de la prioridad asignada a las clases. Para un paquete con prioridad elevada, respecto a los de prioridades menores, habrá asociado un tiempo de espera menor, pero la utilización asociada a éste, será mayor.

$$\sum_{r=1}^{P} \rho_r E[W_r] = \rho E[W] \tag{7}$$

3 Algoritmo de corrección para la medición de retardos de red

El generador de tráfico al transmitir un paquete almacena el tiempo de transmisión basado en el reloj de la CPU transmisora, mientras que la CPU receptora de los paquetes almacena el tiempo de recepción de los mismos; es por ello que es de suma importancia el sincronismo de los relojes de ambas terminales. Se buscó una alternativa para evitar el sincronismo sin perder la capacidad de medición. La solución propuesta se basó en considerar que la diferencia de tiempo que existe entre que los datos parten desde el transmisor hasta que arriban al receptor, se compone de una porción correspondiente al tiempo empleado en atravesar el enlace completo y otra a la diferencia entre los relojes del transmisor y el receptor; es decir, la desincronización temporal entre ambos terminales (Des). A su vez el tiempo necesario para atravesar el enlace se encuentra compuesto por:
- La cola de entrada en el enrutador, esta cola tiene el tamaño de un paquete, por lo tanto se supondrá que es aproximadamente el tiempo se servicio de un paquete.
-Tiempo de servicio en el puerto de salida del enrutador
-Tiempo en la cola de salida del enrutador
-Tiempo de propagación (puede ser despreciado debido a la capacidad del Ethernet 100 Base-T).
- Se desprecia el tiempo de conmutación en el enrutador.

A fines de analizar el comportamiento de los retardos temporales causados por la desincronización se transmitió un flujo de paquetes con longitudes constantes, tiempo de generación constante y la velocidad de salida sin limitación (100 Mbps). Se calculó el tiempo de desincronización empleando la siguiente expresión:

$$\text{Des} = \text{Dmáx} - \text{Tcola} \tag{11}$$

Siendo:
-Des: tiempo por desincronización
-Dmáx: Retardo obtenido utilizando un flujo constante de paquetes con longitud constante y sin limitaciones de canal

-Tcola : tiempo de permanencia en la cola de entrada al enrutador que se aproximará al tiempo de servicio (Tserv)

Para comprobar la validez de la expresión propuesta se repitieron las experiencias con el sistema M/M/1 y los retardos obtenidos de esta manera coincidieron con los calculados teóricamente.

4 Experiencias

Se implementaron dos experiencias, en el primer caso se verifica el comportamiento de un modelo M/M/1, y en otra el modelo de tráfico con prioridades. La implementación real de los modelos de tráfico, se realiza en un sistema Mikrotik Routerboard 433, generándose flujos de datos UDP en puertos determinados, configurándose cada escenario respectivo en el generador de tráfico y en el enrutador [4]. En todas las gráficas adjuntas se muestran las curvas del retardo D en función del rendimiento ρ.

Fig. 1. Escenario dela implementacion.

4.1. Modelo M/M/1: Cálculos teóricos e implementación.

Se analizan dos casos, en el primero, se calculan los parámetros de un modelo M/M/1, y se compara con los resultados de la medición real en la implementación.

Es importante tener en cuenta en los cálculos, la longitud real del paquete que será el valor de los datos propios enviados sumados con la longitud de la cabecera (336 bits) del paquete final a enviarse por la interfaz de red.

En un segundo caso, se analiza el modelo con una exigencia mayor, aumentándose en un factor de 10 la tasa de arribo y la capacidad del servidor de salida, de modo de mantener conservar el factor de carga del servidor.

4.1.1 Primer escenario

Se calcula el modelo M/M/1, correspondiente a los siguientes parámetros, y para diferentes valores de ancho de banda del canal.

Longitud media de datos por paquete: 8000 bits
Longitud media total incluyendo la cabecera UDP: 8336 bits
Tasa de arribo: 5 [paq/seg]

Tabla 1. Cálculos modelo M/M/1 Primer escenario.

C [Bits/seg]	μ [seg]	P	D [seg]	Q [paq]
50000	5,99808061	0,8336	1,00192308	4,17601538
70000	8,39731286	0,59542857	0,29435028	0,87632284
90000	10,7965451	0,46311111	0,17251656	0,39947167
110000	13,1957774	0,37890909	0,12201405	0,23116117
130000	15,5950096	0,32061538	0,09438406	0,15130491
150000	17,9942418	0,27786667	0,07695716	0,10691915
170000	20,3934741	0,24517647	0,06496259	0,0796365
190000	22,7927063	0,21936842	0,0562028	0,0616456

De la implementación real, se miden los parámetros anteriormente calculados. Se varía la velocidad de salida en el enrutador y se obtienen las siguientes mediciones:

Tabla 2. Mediciones de la implementación del modelo.

C [Bits/seg]	λ [paq/seg]	μ [seg]	ρ	D [seg]	Q [paq]
50000	5,2588	5,80720093	0,90556536	1,1805	5,30244804
70000	5,1087	8,33333333	0,613044	0,2455	0,64114185
90000	4,9739	11,0132159	0,45163012	0,1307	0,19845861
110000	5,0067	13,1061599	0,38201121	0,1003	0,1201608
130000	5,1404	15,4559505	0,33258388	0,0873	0,11617304
150000	4,9844	18,4842884	0,26965604	0,058	0,01943916
170000	4,8938	20,9205021	0,23392364	0,0584	0,05187428
190000	4,8167	22,4719101	0,21434315	0,0534	0,04286863

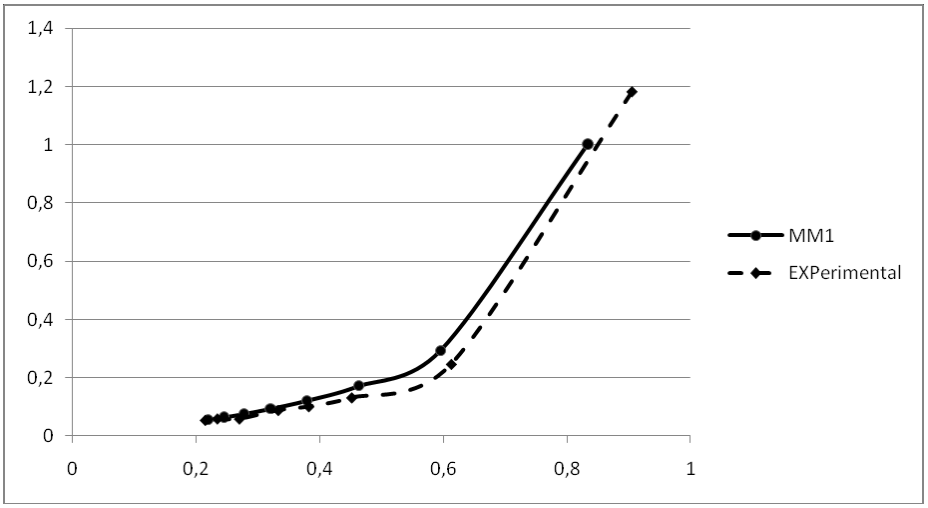

Fig. 2. Desempeño del sistema en el primer escenario. Comparación de resultados del modelo teórico y la implementacion.

4.1.2 Segundo escenario

Se calcula el modelo M/M/1, correspondiente a los siguientes parámetros, y para diferentes valores de ancho de banda del canal.

Longitud media de datos por paquete: 8000 bits
Longitud media total incluyendo la cabecera UDP: 8336 bits
Tasa de arribo: 50 [paq/seg]

Tabla 3. Cálculos modelo M/M/1 Primer escenario.

C [Bits]	μ [seg]	ρ	D [seg]	Q [paq]
500000	59,9808061	0,8336	0,10019231	4,17601538
700000	83,9731286	0,59542857	0,02943503	0,87632284
900000	107,965451	0,46311111	0,01725166	0,39947167
1100000	131,957774	0,37890909	0,01220141	0,23116117
1300000	155,950096	0,32061538	0,00943841	0,15130491
1500000	179,942418	0,27786667	0,00769572	0,10691915
1700000	203,934741	0,24517647	0,00649626	0,0796365
1900000	227,927063	0,21936842	0,00562028	0,0616456

De la implementación real, se miden los parámetros anteriormente calculados:

Tabla 4. Mediciones de la implementación del modelo.

C [Bits]	*λ [paq/seg]*	*u [seg]*	*Rho*	*D [seg]*	*Q [paq]*
500000	48,3927	59,5238095	0,81299736	0,0921	3,64397031
700000	49,6453	84,0336134	0,59077907	0,0294	0,86879275
900000	50,3228	109,89011	0,45793748	0,0224	0,66929324
1100000	49,4794	136,986301	0,36119962	0,016	0,43047078
1300000	51,59	156,25	0,330176	0,0144	0,41272
1500000	48,4652	178,571429	0,27140512	0,0117	0,29563772
1700000	48,9887	204,081633	0,24004463	0,0082	0,16166271
1900000	49,1894	232,55814	0,21151442	0,0057	0,06886516

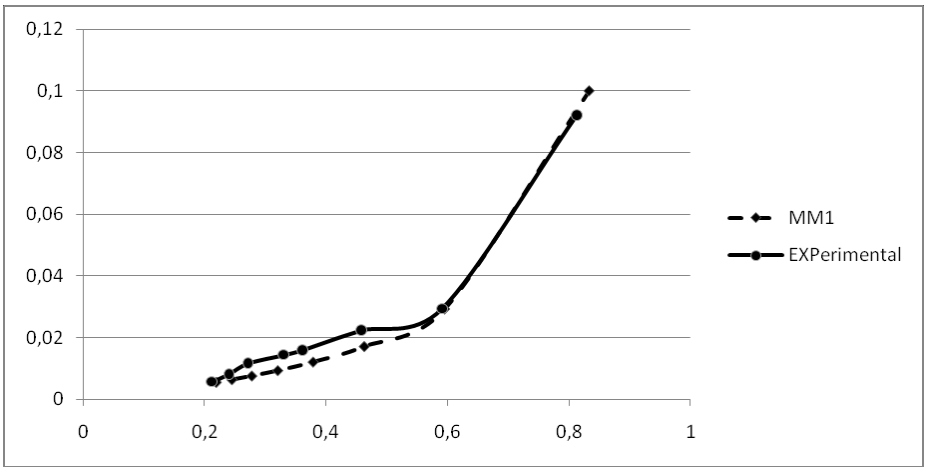

Fig. 3. Desempeño del sistema en el segundo escenario. Comparación de resultados del modelo teórico y la implementacion.

4.2 Modelo de tráfico con prioridades: Cálculos teóricos e implementación.

En esta experiencia se calculan los parámetros de un modelo de tráfico con prioridades y se comparan con los resultados de la medición real en la implementación. El escenario se plantea con tres flujos diferentes, donde se aplican colas con prioridades. Los flujos son generados en tres puertos diferentes, 5001, 5002 y 5003. Las prioridades son asignadas de mayor prioridad a menor en el orden mencionado. Nuevamente se analiza para diferentes anchos de banda.

Tabla 5. Cálculos del modelo teórico y mediciones en la implementación.

	MODELO TEÓRICO			MEDICIONES IMPLEMENTACIÓN		
C [bits/seg]	130000					
FLUJOS	**FL 5001**	**FL 5002**	**FL 5003**	**FL 5001**	**FL 5002**	**FL 5003**
Lambda	5			4,98	5,0403	4,8439
Tiempo de servicio [seg]	0,064123077			0,0636	0,065	0,0638
P	0,320615385			0,31673	0,32762	0,309
Retardos [seg]	0,0908	0,253	4,5057	0,1542	0,334	2,9846
C [bits/seg]	150000					
FLUJOS	**FL 5001**	**FL 5002**	**FL 5003**	**FL 5001**	**FL 5002**	**FL 5003**
Lambda	5			5,1066	4,8837	5,0479
Tiempo de servicio [seg]	0,055573333			0,0582	0,054	0,0555
P	0,277866667			0,2972	0,26372	0,2802
Retardos [seg]	0,0642	0,1444	0,6267	0,1247	0,1688	0,5525
C [bits/seg]	170000					
FLUJOS	**FL 5001**	**FL 5002**	**FL 5003**	**FL 5001**	**FL 5002**	**FL 5003**
Lambda	5			5,0518	4,8875	4,8889
Tiempo de servicio [seg]	0,049035294			0,0491	0,0486	0,0498
P	0,245176471			0,24804	0,23753	0,2435
Retardos [seg]	0,0478	0,0938	0,2676	0,0853	0,1159	0,2746
C [bits/seg]	190000					
FLUJOS	**FL 5001**	**FL 5002**	**FL 5003**	**FL 5001**	**FL 5002**	**FL 5003**
Lambda	5			4,8724	4,7449	4,6944
Tiempo de servicio [seg]	0,043873684			0,0435	0,0453	0,046
P	0,219368421			0,21195	0,21494	0,2159
Retardos [seg]	0,037	0,0659	0,1505	0,0714	0,0979	0,1731
C [bits/seg]	220000					
FLUJOS	**FL 5001**	**FL 5002**	**FL 5003**	**FL 5001**	**FL 5002**	**FL 5003**
Lambda	5			5,0897	4,858	4,9785
Tiempo de servicio [seg]	0,037890909			0,0399	0,0383	0,0373
Rho	0,189454545			0,20308	0,18606	0,1857
Retardos [seg]	0,0266	0,0428	0,0803	0,0613	0,0687	0,1021
C [bits]	250000					
FLUJOS	**FL 5001**	**FL 5002**	**FL 5003**	**FL 5001**	**FL 5002**	**FL 5003**
Lambda	5			4,9784	4,9707	5,1052
Tiempo de servicio [seg]	0,033344			0,0316	0,0327	0,0335
P	0,16672			0,15732	0,16254	0,171
Retardos [seg]	0,02	0,03	0,0501	0,0311	0,039	0,0417

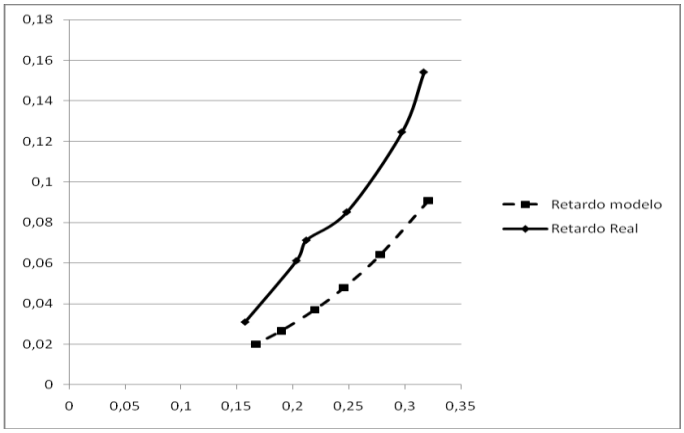

Fig. 4. Comparación de desempeño entre el modelo teórico y la implementacion del flujo en el puerto 5001.

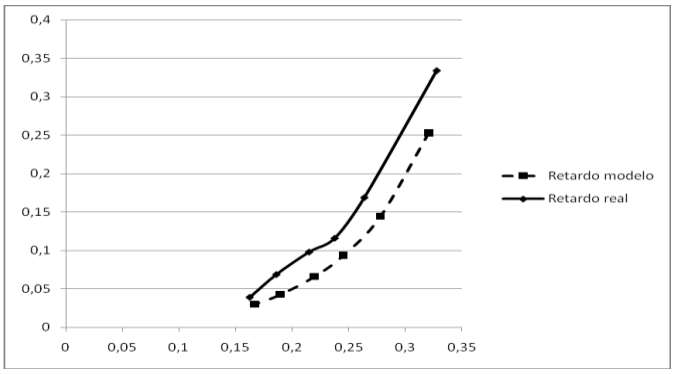

Fig. 5. Comparación de desempeño entre el modelo teórico y la implementacion del flujo en el puerto 5002.

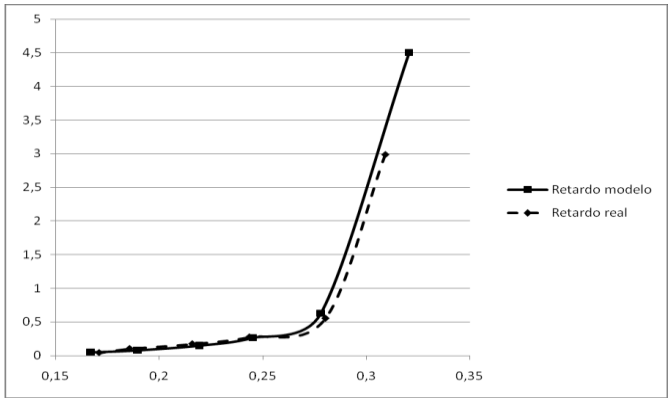

Fig. 6. Comparación de desempeño entre el modelo teórico y la implementacion del flujo en el puerto 5003.

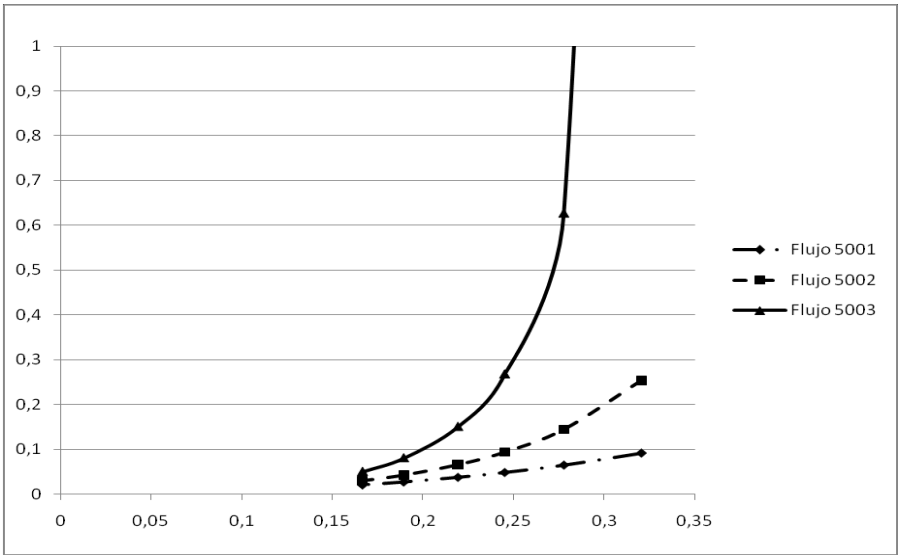

Fig. 7. Comparación del desempeño de los tres flujos en el modelo teórico.

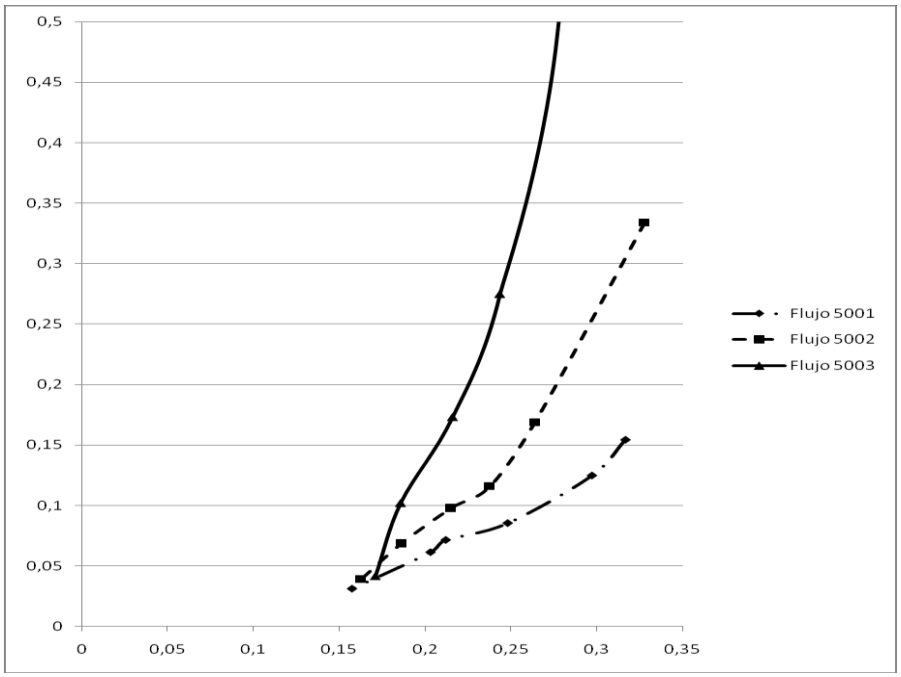

Fig. 8. Comparación del desempeño de los tres flujos en la implementación real.

5 Conclusión

Luego de analizar los resultados obtenidos se logro validar los modelos de tráfico teóricos, con las mediciones obtenidas de una implementación real mediante la comparación de las curvas teóricas con las experimentales. Cabe destacar que sólo es posible medir los retardos del orden de los milisegundos o superiores debido a que el algoritmo utilizado para lograr la sincronización introduce un error significativo para órdenes superiores. Se observa además que el tiempo de retardo del paquete en el buffer de entrada para el caso de colas con prioridades, genera mayores retardos que el tiempo de servicio supuesto, lo que explica el error que se produce para cargas elevadas ($\rho \rightarrow 0.33$).

Referencias

1. Botta A., Dainotti A., Pescape A., "Multi-protocol and multi-platform traffic generation and measurement", INFOCOM 2007 DEMO Session, May 2007, Anchorage (Alaska, USA).
2. Frost V., Melamed B., "Traffic Modeling for Telecommunications Networks", IEEE Telecommunication Magazine, 1994.
3. Cosart, L., "Precision Clock Synchronization for Measurement, Control and Communication", 2007. ISPCS 2007. IEEE International Symposium.
4. Karam M., Tobagi F., "Analysis of the delay and jitter of voice traffic over the Internet", Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2001.
5. De Vito, L.; Rapuano, S.; Tomaciello, L.,"One-Way Delay Measurement: State of the Art", Instrumentation and Measurement, IEEE Transactions, 2008.
6. Siachalou S., Georgiadis L., "Efficient QoS Routing", IEEE Infocom 2003
7. Emma D., Pescape A., Ventre G., "Analysis and experimentation of an open distributed platform for synthetic traffic generation", IEEE Distributed Computing Systems, 2004.
8. L. Kleinrock, "Queueing System", volume 1, John Wiley, 1975
9. L. Kleinrock, "Queueing System", volume 2, John Wiley, 1976
10. http://lartc.org/
11. Hayes J., Ganesh Babu V.J.. "Modeling and Analysis of Telecommunications Networks", John Wiley, 2004.