

SIMULACIÓN DE ACTIVIDADES EN RED LOCAL

Cristóbal Raúl Santa María smaria@sion.com 1561454636 UNLaM.
Osvaldo Spósito. sposito@unlam.edu.ar 01144808952 UNLaM.
Sebastián Fontana sfontana@unlam.edu.ar 01144808952 UNLaM.
Alexis Romero romeroa@unlam.edu.ar 01144808952 UNLaM.

1-INTRODUCCIÓN

El presente trabajo es continuidad del Modelado de Actividades en Redes Locales presentado en WICC2007 y se inscribe dentro del proyecto que intenta desarrollar las bases teóricas de un simulador de la actividad de la red de computadoras de la UNLAM. El propósito es doble; por un lado se trata de elaborar una herramienta para evaluar en forma simulada el desempeño de la red ante un eventual aumento del tráfico de tramas y conocer así probables situaciones de congestión y por el otro, se pretende diseñar una herramienta propia que sirva de apoyo didáctico para la enseñanza de conocimientos sobre redes de computadoras. La simulación prevista considera cada trama desde que es emitida por una estación de trabajo hasta que arriba a su destino. Para ello se identificaron como relevantes las actividades de emisión de tramas y de procesamiento en el switch. A su vez la longitud, el origen, el destino y la prioridad de proceso resultan los atributos principales a simular en cada trama. Obtenidos desde un tráfico real, los patrones de comportamiento estadístico de estos atributos y de las actividades señaladas, se podrán utilizar como regla probabilística, y por medio de técnicas de Monte Carlo obtener un tráfico simulado a través de la red. El realizarlo implica resolver ciertas dificultades, pues la simulación debe actuar siguiendo un reloj con el que se actualiza el sistema cuando una trama es emitida, cuando entra al switch o cuando sale de él. Por lo tanto resulta necesario extraer, desde datos reales, el comportamiento probabilístico de los tiempos entre emisión de tramas y de los tiempos de proceso en el switch. En el trabajo anterior se confirmó la hipótesis de que la emisión de tramas hacia la red es un proceso aleatorio de carácter Poisson – exponencial. En la continuidad de la tarea, se trata ahora de establecer los patrones estadísticos de los atributos señalados de las tramas y de la actividad de procesamiento en el switch. De acuerdo al cronograma fijado para el proyecto se ha trabajado en los aspectos que se detallan a continuación.

2-ACTIVIDAD EN UN SWITCH

Una de las claves en el planteo de la simulación del tráfico de tramas a través de una red es el conocimiento de la actividad en el switch. Este dispositivo recibe cada trama y la envía hacia la dirección que esta lleva marcada en su header. Este proceso dada la velocidad con que arriban las tramas determina la formación de una cola de espera cuya longitud potencial es finita y depende de la tecnología que constituye el aparato. En cualquier caso y a los efectos de la simulación del paso de la trama por el switch, este puede ser considerado como una “caja negra” estableciendo una ley estadística para el tiempo que la trama permanece en el switch al medir ese tiempo para una muestra adecuada de tramas.

La idea requiere un análisis más detallado. Medir el tiempo que cada trama está en proceso en el switch implica en realidad, dadas sus características de “caja negra”, registrar el tiempo en que la trama entra en el switch y el momento en que sale del mismo. Por supuesto hay entonces que desarrollar aplicaciones con este fin que exceden la performance de los habituales programas tipo “sniffer” pues es preciso identificar cada trama a la entrada y a la salida del switch, realizar la diferencia de tiempos en cada caso, registrar el protocolo utilizado, la prioridad de la trama, las direcciones de origen y destino y la longitud medida en bytes.

Por otra parte, todo esto solo es posible si todas las estaciones de la red tienen sincronizados sus relojes. A efecto de estudiar el comportamiento en un principio se utilizó un switch Enterasys V2H124-24 S/W 2.5.2.9 pero con este dispositivo no se pudo lograr la sincronización de los equipos para usar un programa de testeo ni redireccionar tramas que entraban desde otros puertos.

Se decidió entonces realizar una experiencia sincronizando una pequeña red experimental formada por cuatro equipos vinculados a un 3Com Switch 5500 Family. Se generó en ella tráfico que se procesó con una aplicación desarrollada para tal finalidad que identifica por un número de código cada trama, realiza el cómputo de la diferencia de tiempos de entrada y salida del switch y registra las direcciones de origen y destino, el tamaño en bytes, la prioridad y el protocolo utilizado y vuelca todo en una planilla de cálculo.

3-SINCRONISMO

Se utiliza el comando NET TIME para sincronizar las estaciones de trabajo en el Sistema Operativo Windows XP. Se describe ahora el método

En una red que tiene software diverso de cliente de red, grupos de trabajo diferentes y dominios diferentes, el mantenimiento de un archivo de comandos de inicio de sesión que emita un comando net time puede resultar complejo. La solución para mantener sincronizada la hora de todas las estaciones de trabajo con un servidor determinado consiste en ejecutar un archivo de comandos de inicio de sesión en cada estación de trabajo o hacer que cada estación de trabajo ejecute su propio archivo por lotes después de iniciar sesión en la red.

Por ejemplo, si se supone que se utiliza un servidor llamado SrvHora que tiene un reloj de sistema fiable el siguiente es un archivo por lotes de ejemplo que puede ejecutarse desde la estación de trabajo o puede llamarse desde (o formar parte de) un archivo de comandos de inicio de sesión: net time \\SrvHora /set /yes

Si utiliza sólo el nombre de servidor en lugar del parámetro /domain o /workgroup, evita la complejidad de mantener diferentes archivos de comandos de inicio de sesión o archivos por lotes para los distintos tipos de software de cliente de red. Si todas las estaciones de trabajo ejecutan el mismo tipo de software de cliente de red, se puede utilizar el parámetro /domain o /workgroup. Sin embargo, en una red que tenga software diverso de cliente de red, resulta más fácil utilizar el parámetro \\computername. El comando net time está habilitado de forma predeterminada en todas las estaciones de trabajo Windows XP.

4-DESARROLLO DE LA APLICACIÓN

Plataforma: Aplicación de Windows realizada con Visual C#

El programa recorre un archivo con extensión “.txt” y realiza una extracción de diversos campos de interés, almacenándolos en una base de datos, para realizar un informe luego. El archivo .txt contiene información de diversos paquetes que viajan por la red. La información es procesada por el programa, para poder extraer los campos siguientes:

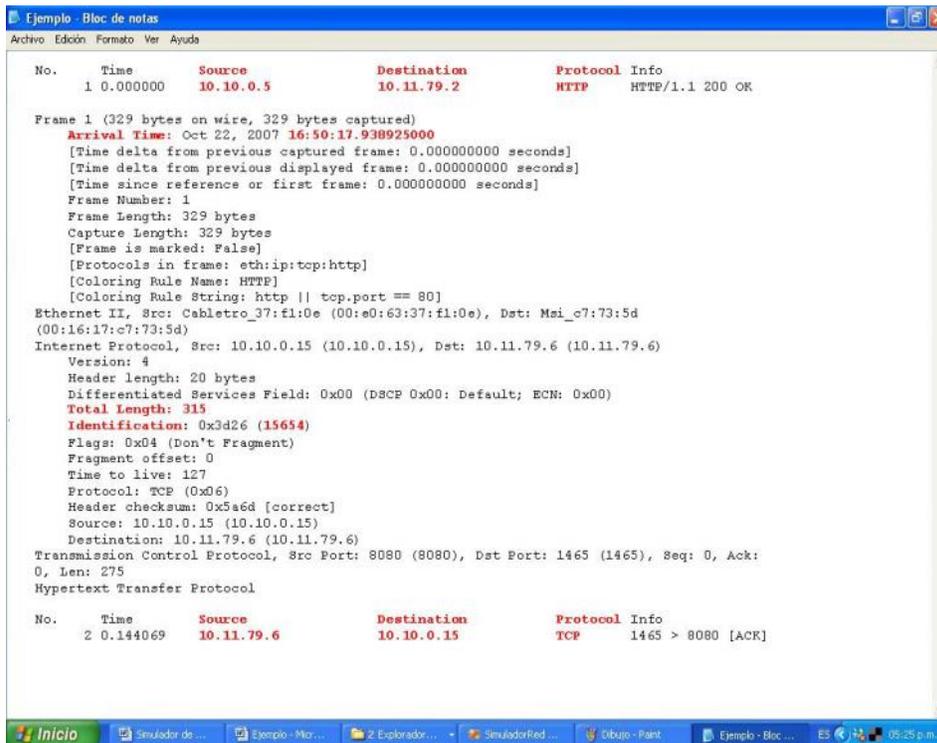
- Origen - Source
- Destino - Destination
- Protocolo - Protocol
- TiempoLlegada – Arrival Time
- LongitudTotal – Total Length
- Identificación – Identification

Campos que deben ser ingresados manualmente:

- CódigoLote – Código de Lote
- FechaProcesamiento – Fecha de procesamiento

- MaquinaOrigen – Maquina de origen
- Campos calculados por el programa:
- Prioridad – En realidad para estas primeras pruebas se le asignó una prioridad igual a 0 a todos los paquetes
- TiempoSegundos – Calcula el Delta tiempo en segundos que tarda un paquete en ser recibido.

A continuación se muestra un ejemplo de un archivo, en el mismo se resaltan los campos que se extraen en rojo.



5-DISEÑO DE EXPERIMENTOS Y ANALISIS ESTADÍSTICO DE REGISTROS

Hasta el presente se han realizado las siguientes experiencias utilizando para ello solo 2 computadoras conectadas en red:

Primera prueba:

Se transfirieron archivos de entre 5 y 10 Mb (algo no muy pesado para testear el comportamiento del switch).

Observación: El tiempo empleado en la transferencia de archivos desde una maquina hacia la otra era relativamente parecido.

Segunda prueba:

Se decidió transferir archivos un poco más pesados que en la anterior, estos son de 50 Mb aproximadamente).

Observación: A diferencia de la primer prueba, esta ha generado muchos más paquetes en la transferencia

Nota: Estos procesos no duraron más de 1 minuto debido a la carga que genera el Log para procesarlo.

La información obtenida debe ser procesada estadísticamente a efecto de determinar los patrones de aleatoriedad de los atributos y actividades de interés.

La primera prueba se refiere a 6130 tramas de las que se cuenta con la siguiente información: direcciones de origen y destino, tamaño en bytes y código identificador, tiempo de permanencia en el switch. Una parte de la planilla Excel respectiva se reproduce a continuación:

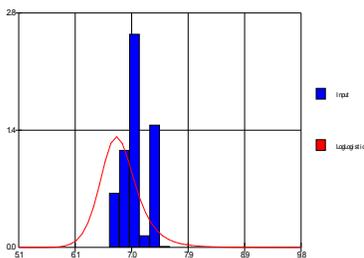
Ip Origen	Ip Destino	Protocolo	Tam Paquete	Identificador	Prioridad	Delta
10.11.79.2	10.11.79.5	DCERPC	172	2160	0	0,006973
10.11.79.2	10.11.79.5	DCERPC	172	2282	0	0,007027
10.11.79.2	10.11.79.5	DCERPC	172	5252	0	0,007293
10.11.79.2	10.11.79.5	DCERPC	172	5265	0	0,007312
10.11.79.2	10.11.79.5	DCERPC	180	5231	0	0,007174
10.11.79.2	10.11.79.5	SMB	100	2149	0	0,006976
10.11.79.2	10.11.79.5	SMB	100	2201	0	0,006988
10.11.79.2	10.11.79.5	SMB	100	5247	0	0,007277
10.11.79.2	10.11.79.5	SMB	100	5262	0	0,007316
10.11.79.2	10.11.79.5	SMB	1024	2220	0	0,007143
10.11.79.2	10.11.79.5	SMB	103	5232	0	0,007158
10.11.79.2	10.11.79.5	SMB	103	5293	0	0,00731
10.11.79.2	10.11.79.5	SMB	103	5322	0	0,007322
10.11.79.2	10.11.79.5	SMB	103	5389	0	0,007317
10.11.79.2	10.11.79.5	SMB	103	5454	0	0,007373
10.11.79.2	10.11.79.5	SMB	103	5455	0	0,007313

Estos datos se procesaron con el software BestFit de Palisade.

Se analizan entonces los resultados obtenidos para el ajuste de los tiempos de permanencia en el switch que en la planilla corresponden a la columna titulada delta teniendo en cuenta solo los correspondientes a tramas de protocolo SMB.

Para empezar y a efecto de mayor comodidad en el trabajo los datos se escalaron multiplicándolos por el factor 1000. Se ajustaron 21 distribuciones usuales y se analizó en cada caso la bondad del ajuste por medio del test Chi Cuadrado. Si bien el programa empleado produce un ranking de mejor ajuste de acuerdo a esta prueba resultó que, en todos los casos, el ajuste debía ser rechazado. La gráfica y el informe a continuación intercalados evidencian esta situación.

Comparison of Input Distribution and LogLogistic(4,87,1,92;10,05)



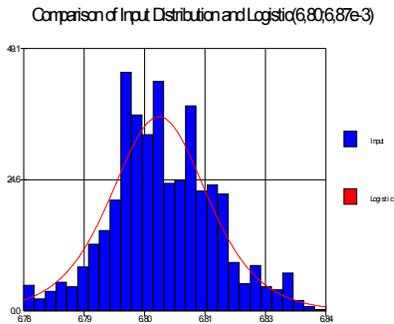
Rank/Distribution Chi-Square

1. LogLogistic 1,510409e+4
2. Weibull 1,957391e+4
3. Triang 2,240191e+4
4. Uniform 4,511309e+4

Chi-Square

Test Value	1,510409e+4	1,957391e+4	2,240191e+4	4,511309e+4
Confidence	Rejected	Rejected	Rejected	Rejected
Rank	1	2	3	4

A continuación se filtraron los datos de entrada para intentar un mejor ajuste quitando valores atípicos. Sin embargo y como se ve en la gráfica y el informe reproducidos a bajo los resultados fueron igualmente pobres y debieron rechazarse los ajustes realizados.



Rank/Distribution	Chi-Square
1. Logistic	643,628036
2. Lognorm2	750,187121
3. Lognorm	750,187121
4. InverseGaussian	750,187577

Chi-Square

Test Value	643,628036	750,187121
	750,187121	750,187577

Confidence	Rejected	Rejected	Rejected	Rejected
Rank	1	2	3	4

Hasta aquí se observa que el carácter polimodal de la distribución empírica impide hallar una fórmula entre las distribuciones usuales para representar el fenómeno. Esto en principio obligaría a simular la actividad utilizando directamente la tabla de la distribución acumulada obtenida de los datos pero tal conclusión debe ser apoyada en observaciones más generales que evidencien este patrón estadístico con mayor certeza. Con esa finalidad se han obtenido nuevos datos que al momento de redacción de este informe se encuentran en proceso.

6-CONCLUSIONES

En la etapa actual de construcción de un simulador de red local se ha logrado avanzar en la identificación de las actividades que sobre la propia red constituyen los aspectos principales del tráfico de tramas, se ha resuelto con suficiente precisión el problema de sincronización de relojes entre los distintos equipos conectados, se ha desarrollado una aplicación que permite extraer del tráfico los datos necesarios para la determinación de los patrones de comportamiento y se han diseñado dos experimentos de captura de tráfico procesándose estadísticamente los datos obtenidos. Se encuentra en curso el análisis estadístico del segundo lote de datos a efecto de ampliar el conocimiento sobre el patrón estadístico de los tiempos de proceso en el switch. Con este patrón y con el ya determinado para los tiempos entre emisión de tramas se procederá durante los meses próximos a desarrollar el programa de simulación en su parte esencial de emisión de tramas y de procesamiento en el switch, tal como está pautado en el cronograma del proyecto. Para estas tareas se han incorporado además al grupo de investigación un graduado reciente y un estudiante avanzado de la carrera de ingeniería informática de la universidad.