

# Predicción en el tiempo de respuesta de páginas web

Martha Azucena Fernández Saucedo  
Rogelio Ferreira Escutia  
Instituto Tecnológico de Morelia  
Ingeniería en Sistemas Computacionales

## Resumen

*El aumento del tráfico electrónico debido a la creciente demanda del uso de internet ha provocado que la eficiencia, rapidez y calidad de los servidores sea ahora un tema importante tanto para los empresarios como para los consumidores. Por lo cual es necesario evaluar los servidores o sitios web de tal manera que la información obtenida facilite la toma de decisiones. Una forma de obtener esta información es con la combinación del comando traceroute y el método de Monte Carlo que son parte fundamental de la presente propuesta.*

## 1. Introducción

Al intentar acceder a una dirección de internet en un navegador, esperamos cierto tiempo a que la página sea cargada, pero ¿Cómo se puede predecir cuánto tiempo tardará en cargarse una página y por qué? Para poder dar una respuesta confiable a esta pregunta, se debe hacer un cuidadoso estudio que nos permita conocer las probabilidades, que a su vez nos darán un resultado.

Hay muchos factores que intervienen al acceder a una página: el tipo de conexión, el tráfico, el tipo de servidor, etc. Uno de dichos factores es la ruta que sigue la petición al salir del ordenador para llegar al servidor que hospeda la página en cuestión.

En el presente artículo se propone el uso de la técnica de Monte Carlo para pronosticar el tiempo de respuesta de un servidor, y los posibles bloqueos durante la ruta.

## 2. Rutas

Al utilizar internet hay envío y recepción de información, esta información se divide en paquetes; cada paquete contiene datos codificados. Cuando se intenta ingresar a una página, se hace una petición al servidor que la contiene y éste envía al usuario, paquetes que contienen la información solicitada.

Una analogía sería: cierta persona manda una solicitud de pedido por correo; el correo manda al cartero a entregar la

correspondencia. El cartero analiza el lugar de destino de cada uno de los envíos, así sigue una ruta de tal manera que no tenga que regresar, sino que todos queden "en el camino". Al llegar la solicitud al destinatario, este manda paquete por paquete el pedido, uno a la vez. Si el paquete no llega, el proveedor lo envía de nuevo. Podemos pensar en la correspondencia como los paquetes y el cartero como la internet, entonces las estaciones de correo por donde pasa serán la ruta.



Fig.1 Ruta vista geográficamente

Los paquetes viajan a través de servidores desde el equipo emisor hasta el receptor, trazando un camino o ruta. Las rutas son administradas y asignadas por equipos especiales llamados ruteadores.

## 3. Traceroute

*Traceroute* es un comando usado en UNIX (el equivalente en Windows es *tracert*). Este comando traza la ruta que toma un paquete de datos al hacer una petición a un IP o dominio a través de la red. Cuando un sitio Web es particularmente lento, un Traceroute determina que parte de la ruta de Internet es responsable por la latencia.

Utiliza el campo TTL (Time To Live) del protocolo IP, inicializándolo con valores bajos y aumentándolo conforme va recibiendo tramas ICMP. Cada vez que un paquete pasa por un ruteador, este decrementa el campo TTL en una unidad; en el caso de que se alcance un valor 0, se devuelve un paquete y se descarta la trama. Así, traceroute inicializa a 1 este campo, lo que ocasiona que el primer ruteador encontrado,

devuelva el mensaje de error; al recibirlo, lo inicializa a 2, y ahora es el segundo ruteador el que descarta el paquete y envía otro mensaje de error, y así sucesivamente. De esta forma se va construyendo la ruta hasta un determinado host remoto [3].

```
traceroute to antares.itmorelia.edu.mx (200.33.171.2), 30 hops max, 40
byte packets
 1 10.10.1.1 (10.10.1.1) 12.913 ms 37.152 ms 10.155 ms
 2 192.168.14.254 (192.168.14.254) 13.92 ms 12.281 ms 9.593 ms
 3 host-207-248-243-57.block.alestra.net.mx (207.248.243.57) 22.194
ms 90.6 ms 43.355 ms
 4 rdmex2.alestra.net.mx (207.248.240.130) 45.871 ms 16.303 ms
110.938 ms
 5 rcMex2.alestra.net.mx (207.248.240.225) 22.556 ms 20.27 ms
18.552 ms
 6 inetmexvallejo6-pos1-3.uninet-ide.com.mx (148.233.248.126)
132.759 ms 28.186 ms 39.068 ms
 7 bb-mexvallejo-1-g8-0.uninet.net.mx (200.38.197.33) 193.777 ms
200.642 ms 30.263 ms
 8 bb-jal-tlaquepaque-1-pos3-0.uninet.net.mx (200.38.192.245) 119.809
ms !H * 30.129 ms !H
```

Fig. 2 Comando traceroute para antares.itmorelia.edu.mx desde Morelia

#### 4. Método de Monte Carlo

El método de Monte Carlo es un método numérico que se utiliza para resolver problemas mediante el uso de variables aleatorias. Este método general tiene múltiples aplicaciones, incluso puede usarse en la resolución de integrales múltiples. [2]

La idea básica del método consiste en que se genera un número aleatorio uniforme; después, este número se busca dentro de los rangos en los que se subdivide el intervalo de cero a uno, el rango en el que haya sido encontrado, determinará el valor de la variable.[1]

#### 5. Propuesta

Conocer el tiempo de respuesta de un servidor no debe ser una tarea difícil si aplicamos los conceptos anteriores adecuadamente.

La primera tarea es definir, si no es que se tiene ya definido, el servidor o página web a estudiar. A continuación se hace un muestreo ejecutando n veces el comando traceroute. Se puede ejecutar el comando desde diferentes equipos, ubicados en lugares distintos, conexiones diferentes y con diferentes sistemas operativos.

Es muy importante saber que si las muestras tienen valores no estables o muy dispersos el tamaño de la muestra debe ser mayor. Una vez hecho el muestreo, se agrupan los tiempos de respuesta y los servidores en los que hubo bloqueo, en clases. Después, se calcula la probabilidad observada que se tomará como la probabilidad real.

Cuando ya se han calculado las probabilidades, entonces se aplica el método;

es decir, se toma un número aleatorio y se busca el valor en los rangos. El rango al que pertenezca dirá será la predicción de tiempo de respuesta o el servidor en el que el bloqueo ocurrirá. Cabe mencionar que para uno de los aspectos evaluados se debe generar un número aleatorio diferente.

#### 6. Ejemplo de implementación

Se estudió el sitio web del Instituto Tecnológico de Morelia. Se tomaron 500 muestras con una conexión a 64kbps por cable y 500 con conexión a 56kbps por línea telefónica.

Al aplicar la propuesta a los resultados, se decidió hacer un estudio para las muestras de línea telefónica y otro para las de cable.

Los resultados fueron los siguientes: El tiempo de respuesta proyectado para una conexión con cable es de 450ms, sin bloqueos; para la conexión con línea telefónica es de 730ms., sin bloqueos.

#### 7. Conclusiones

Durante la implementación del ejemplo, al obtener las muestras se encontró una de ellas particularmente singular y que sin duda llamó la atención: al utilizar el comando antes mencionado en Madrid, España, el tiempo de respuesta fue mucho menor a los obtenidos en la ciudad de Morelia aún cuando el servidor se encuentra en esta última; incluso, el número de servidores por los que viajaron los paquetes fue menor.

Los resultados obtenidos nos llevan a pensar que la ubicación física de un servidor y del usuario que solicita un servicio no es un factor determinante en el tiempo de respuesta y que se deben crear métodos que hagan más eficiente el tráfico de paquetes.

Si estudiamos determinado servidor tomando en cuenta todos los aspectos que influyen en el proceso de comunicación, los resultados serán más concretos, claros y precisos.

#### 8. Referencias

[1] Friedman, Sasieni, et al. Investigación de operaciones. 968-18-0592-5. México, D.F. 1982. Editorial Limusa.

[2] García, Valle J. Luis. Matemáticas especiales para computación. 84-7615-267-1. Madrid, España. 1994. Mc Graw Hill.

[3] Linux Forum  
www.linuxforum.com/shell/traceroute