

IETFDay 2015, 1º Taller del Grupo de Trabajo de Ingeniería de Internet/Argentina.

Demora en cola aplicada al control de congestión TCP

Guillermo Rigotti

ISISTAN, Fac. de Ciencias Exactas

Universidad Nacional del Centro de la Pcia. de Bs. As.

grigotti@exa.unicen.edu.ar

Palabras clave: control de congestión, TCP

Introducción y Trabajo relacionado

Los métodos de control de congestión utilizados por TCP se basan en modificar la tasa de envío en función de la carga de la red. Tradicionalmente la estimación se hacía sólo en base a las pérdidas de paquetes [1]. Dos problemas conocidos de este mecanismo son la necesaria producción de pérdidas de paquetes para estimar la carga y la sincronización de flujos. En 1989 surge la idea de utilizar la demora ida y vuelta experimentada por los paquetes (RTT) para detectar la congestión anticipadamente [2]. El emisor disminuye entonces su tasa de envío antes de que se produzca una pérdida, mejorando la eficiencia del protocolo y el throughput de la red (p. ej. TCP Vegas y FAST).

Si bien medir el RTT es sencillo, el valor obtenido no basta para una correcta estimación de la carga, ya que se debe tener en cuenta, entre otros factores, la asimetría en los caminos de ida y de vuelta. Por eso se ha intentado estimar la carga en la red de manera más precisa a partir de la demora emisor-receptor (OWD: one way delay). Sin embargo, medir esta demora no es inmediato, ya que los relojes del emisor y del receptor deben estar sincronizados, para lo que se recurre a servidores NTP o uso de GPS. Esto es adecuado para algunas aplicaciones (p.ej. VoIP), pero no para el caso de flujos TCP, ya que por su número y ubicación pueden generar un overhead considerable (NTP) o no tener capacidad de GPS.

Propuesta

A diferencia de las alternativas anteriores, en este trabajo se propone determinar la carga en la red a través de la estimación directa de la demora que sufren los paquetes en las colas de transmisión de los routers. Este parámetro está directamente relacionado con esa carga. Para determinar la demora de queueing emisor-receptor y posibilitar que sea conocida por el emisor para adecuar su tasa de envío, se implementa 1-una nueva opción IP (queueing delay), 2-un mecanismo a nivel IP que procesa esta opción en los routers, y 3-uno a nivel TCP, procesado en emisor y receptor, que permite ponerse de acuerdo en el uso de la opción y luego comunicar de receptor a emisor las demoras percibidas.

Para estimar la carga, se optó por medir demoras en cola y no otro parámetro relacionado, tal como ocupación en cola. Esto se basa en que la demora es una

medida genérica aplicable a todos los routers, que permite agregar la carga individual de cada uno de ellos; por otro lado, es un parámetro que se utiliza normalmente en y por lo tanto puede aplicarse fácilmente a los algoritmos de control de congestión en uso.

La opción `queueing delay` se implementa en IPv4 como una opción del header [3] y como opción del header `hop by hop` en IPv6 [4]. El proceso de esta opción en los routers es simple, y consiste en sumar a la demora registrada en la opción, la demora estimada agregada por el router.

La implementación de la opción IP se realizó en Linux (kernel versión 3.13.0). El reconocimiento de la opción se realiza en el "slow path" y el registro de la demora implica un único punto de modificación en el kernel y cálculos simples sin invocación de funciones ni almacenamiento de estado para el paquete en proceso: 1-se calcula la demora inmediatamente antes de que el contenido del `sk_buffer` sea enviado a la cola de transmisión, 2-la demora no se calcula en base a diferencia de tiempos de llegada y de salida, sino en base a la cantidad de paquetes en cola en la interfaz de salida, a la velocidad de transmisión y al MTU del link. La cantidad de paquetes en cola se obtiene, cuando es necesario, invocando a los posibles mecanismos GSO o TSO en uso.

A nivel TCP, en la etapa de establecimiento de conexión ambas partes intercambian configuraciones específicas en los bits de flags del header TCP para determinar si son capaces y están configuradas para usar la opción `queueing delay`. En la etapa de transferencia de datos cada parte en su rol de emisor determina según algún criterio basado en el algoritmo de control de congestión en uso, en qué momento medir la demora (en todos los segmentos con datos emitidos, cada cierto intervalo de tiempo, etc.). Para hacerlo, indica al nivel IP que en el correspondiente datagram genere la opción `queueing delay`, con un campo demora en cero.

Cuando el datagram llega a destino, la opción conteniendo la demora de `queueing` es accedida por TCP, que la envía al emisor en el siguiente segmento, para que el emisor conozca la demora en un tiempo igual al RTT.

Conclusiones

Para comprobar los siguientes aspectos de importancia de la propuesta presentada se realizaron simulaciones en Network Simulator 2:

1-Valor obtenido en redes asimétricas: Como era de esperar, el valor obtenido fue independiente de la asimetría, no ocurriendo lo mismo con el tiempo de retorno del mismo al emisor.

2-Implementación parcial: Debido a que es muy probable que en una red haya un cierto porcentaje de routers que no implementen la opción, se realizaron pruebas que demostraron que en general se mantiene un valor proporcional (menor) al real; esta proporción se va distorsionando cuando el porcentaje de implementación es menor al 40% de los routers.

3-Applicabilidad del valor obtenido en algoritmos de control de congestión: consiste en reemplazar la medida utilizada en diferentes algoritmos (generalmente o RTT o un valor derivado mediante una heurística) por el valor obtenido de la demora de `queueing`. Esta comprobación está en vías de realización.

Referencias

- [1] Allman, M; Paxson, V.; Blanton, E; "TCP Congestion Control", RFC 5681, IETF, September 2009.
- [2] Jain, R., "A Delay Based Approach for Congestion Avoidance in Interconnected Heterogeneous Computer Networks", Computer Communications Review, ACM SIGCOMM, pp. 56-71, October 1989.
- [3] "Internet Protocol - DARPA Internet Program Protocol Specification", RFC 793, DARPA, September 1981.
- [4] S. Deering R. Hinden, "Internet Protocol, Version 6 (IPv6):Specification", Network Working Group, IETF, RFC 2460, December 1998