

Medición de demanda y de capacidad en redes IP

Alejandro Daniel Popovsky* - Horacio José Peña**

Resumen

El objetivo de este trabajo es sentar las bases para medir y analizar el tipo de demanda de capacidad que una comunidad de interés genera sobre la Internet. Una vez conocida la demanda sobre la red es posible deducir la capacidad disponible para el tráfico generado por esa comunidad y su evolución en el tiempo. Se propone un esquema de clasificación del tipo de tráfico y se describen los criterios necesarios para su detección por medio de un analizador. Se describen las formas de aprovechar la información obtenida de este análisis. Se detallan las experiencias realizadas con un analizador construido en la Universidad de Palermo que implementa los conceptos desarrollados en este trabajo.

Introducción

Al observar una estadística de tráfico sobre una red no siempre salta a primera vista si la demanda de tráfico es menor o mayor que la capacidad disponible a cada momento. El tráfico aumenta y disminuye continuamente su caudal, haciendo variar a los distintos elementos de la red por los que pasa el tráfico, de un estado de saturación a un estado de capacidad disponible, y viceversa. Estos cambios pueden producirse a intervalos muy cortos, de menor duración a la resolución de las estadísticas disponibles a los administradores de red. Muchas veces además, no se dispone de acceso administrativo a todos los equipos de red por los que pasa el tráfico de interés, por lo que no se obtiene la información crítica como por ejemplo, información sobre descarte de datagramas.

Existen protocolos y aplicaciones que demandan una capacidad fija de la red, mientras que otros demandan toda la capacidad disponible. Esto hace pensar que si durante un intervalo de tiempo se está demandando toda la capacidad disponible, la cantidad de tráfico observada coincidirá con ella. De esta manera es posible medir o monitorear la capacidad disponible, de tanta utilidad, por ejemplo para clientes de proveedores de servicio. Y esto sin necesidad de acceder administrativamente a los equipos de red involucrados, y sin necesidad de depender de información de tráfico entregada por el proveedor.

* Ingeniero Electrónico de la Universidad de Buenos Aires, y Director del Departamento de Comunicaciones y de Tecnología, de la Universidad de Palermo. Áreas de Interés: comunicaciones, comunicación de datos, diseño de VLSI, procesamiento de señales.

** Estudiante de Ingeniería Informática orientación Ciencias de la Computación de la Universidad de Palermo.

Para medir la demanda sobre la red hemos utilizado como estrategia el seguimiento individualizado del estado de todas las sesiones individuales que componen el tráfico de interés generado por una comunidad. Esto permite no solamente medir la demanda y el caudal de tráfico, sino también otros datos de utilidad como la pérdida de datagramas, etc.

La primera parte de este artículo contiene una introducción a los conceptos de medición de performance, que puede ser saltada por quienes ya conozcan del tema. La segunda parte contiene el desarrollo de los conceptos de medición de demanda y limitación de velocidad de transmisión introducidos en este trabajo. La tercera parte contiene una descripción del desarrollo de un analizador que implementa los conceptos aquí definidos y las conclusiones.

Primera Parte

Utilización de la capacidad por las aplicaciones

La utilización de la capacidad disponible para el acceso a la internet por parte de un cliente, sea este una red corporativa o una red de un proveedor de servicios de acceso, varía según el tipo de tráfico en cuestión. Este tráfico podría ser del tipo **consulta-respuesta UDP**, podría ser tráfico **streaming UDP** (por ejemplo: voz sobre IP), podría ser tráfico **elástico TCP**, o podría ser tráfico de protocolos de control y administración de redes (por ejemplo: icmp).

El tráfico del tipo consulta-respuesta UDP ocupa generalmente poca capacidad. En caso de saturación de la capacidad, se producen pérdidas de datagramas, que obligan a repetir la consulta, hasta que llegue correctamente una respuesta.

El tráfico del tipo streaming UDP ocupa una capacidad que depende solo de la velocidad de generación de los datos en el emisor. Si falta capacidad de manera que esta es menor que la necesaria para el transporte de una sesión de streaming, esta se cancela. Es cierto que algunas aplicaciones seleccionan la calidad de su sonido o su imagen (lo que afecta la velocidad de generación de datos). Pero a una misma calidad, la cantidad de datos generados por unidad de tiempo es independiente de las condiciones de la red.

El tráfico elástico TCP en cambio, se adapta a la capacidad de la red y sus variaciones durante la transmisión. TCP intenta obtener todo lo que puede de la capacidad disponible de manera de disminuir el tiempo de transmisión. En caso de tener que compartir la capacidad con otras conexiones, esta se dividirá en forma más o menos pareja entre ellas. Existen varias implementaciones de TCP (Tahoe, Reno, SACK, etc) pero todas ellas persiguen este objetivo.

En todo momento la capacidad disponible está siendo utilizada por tráfico de distintos tipos, y por múltiples sesiones. La limitación a cada una de ellas puede estar en un lugar diferente de la Internet, y también estar causada por un problema distinto. Es decir que la performance observada en una sesión no tiene necesariamente relación con la observada en otras sesiones.

Métricas y performance

La performance puede analizarse a través de diferentes **métricas**. Estas métricas pueden estar más o menos cerca de la percepción subjetiva de la performance que tiene un usuario. Por ejemplo el **tiempo medio de bajada de una página web** es una métrica más cercana a la percepción, que la **probabilidad de pérdida de datagramas** hacia o desde un destino genérico. La segunda se relaciona con la primera en una forma indirecta.

Las características de las redes (bit rate de circuitos y circuitos virtuales, retardos, capacidad de procesamiento de routers, etc) por donde pasa el tráfico a analizar, afectan a cada métrica en forma diferente. Por ejemplo el delay afecta fuertemente el tiempo medio de bajada de páginas web, ya que estas se bajan con conexiones relativamente cortas, en donde el tiempo de establecimiento y adaptación es un porcentaje importante de la duración de la sesión. Pero el delay no afecta el throughput agregado máximo obtenible que se puede lograr con todas las sesiones en paralelo de la red del cliente.

En cambio el bit rate de un enlace podría en ciertos casos (cuando el delay y bit rate son grandes) no afectar demasiado el tiempo medio de bajada de páginas web. Sin embargo el **throughput agregado máximo obtenible** sí se verá fuertemente afectado por el bit rate del enlace.

Tipos de mediciones

Una forma de clasificar las mediciones es según si el dispositivo de medición genera o no tráfico que afecte la medición misma.

Cuando estamos generando tráfico para estimar una determinada métrica, podríamos estar afectando las condiciones de la red, y por lo tanto estimar un valor que sería distinto si no estuviéramos realizando la medición. Asimismo se podría estar afectando el servicio a los clientes ya que la capacidad disponible estaría disminuida por su utilización para la medición. A este tipo las podemos llamar **mediciones activas**.

Si las mediciones activas no generan mucho tráfico (por ejemplo: traceroute, ping) se pueden realizar en forma constante. Pero si generan mucho tráfico (por ejemplo: establecer decenas de conexiones ftp simultáneas a diferentes servidores y medir sus duraciones) solo pueden realizarse en forma ocasional ante la sospecha de problemas.

Cuando el dispositivo de medición observa el tráfico pero no lo afecta ni genera tráfico que lo pueda afectar, estamos ante **mediciones pasivas**. Estas tienen la ventaja que pueden realizarse en forma continua y permanente y no solo cuando hay problemas a detectar.

Su desventaja es que **algunas métricas solo pueden ser medidas cuando están presentes ciertas condiciones de tráfico**, y si usamos mediciones pasivas no siempre podremos obtenerlas. Es decir, podrían haber problemas que no se estén detectando dado que no se alcanzan las condiciones de stress necesarias para que afecten el tráfico. Pero si el problema no está afectando el tráfico... podemos decir que hay un problema? Muchos proveedores subsisten gracias a que dimensionan sus redes para que la probabilidad de saturación sea muy baja, pero no nula (por ejemplo que el tiempo acumulado de saturación anual sea menor que el downtime máximo comprometido).

Las mediciones pueden clasificarse como **con acceso administrativo** a los equipos que afectan el tráfico, o **sin acceso administrativo** a los mismos. El tráfico de interés pasa generalmente por muchos dispositivos a los que no se tiene acceso administrativo, ya que podrían por ejemplo pertenecer a otras redes. Por eso son de mucho valor los métodos de medición que permiten obtener información útil, sin acceso administrativo a los equipos donde pueden estar habiendo problemas.

Las mediciones también pueden dividirse entre mediciones **distribuidas** y **concentradas**. En el primer caso se puede en general obtener más información, ya que son afectadas por más dispositivos de red y porque pueden cubrir adecuadamente la topología a estudiar.

Algunas mediciones distribuidas se pueden realizar con herramientas que fueron originalmente concebidas para mediciones concentradas, que luego reportan su información a un sitio central para su análisis.

Este trabajo estudia el análisis de tráfico con un **anализador pasivo, sin acceso administrativo** a los dispositivos de red que afectan el tráfico de interés.

Segunda Parte

Limitaciones a la velocidad de transmisión

Todas las conexiones TCP tienen alguna limitación de velocidad de transmisión, si no la tuvieran, aumentarían su velocidad indefinidamente. Clasificamos las limitaciones en tres tipos básicos:

- **DG: Limitación por generación de datos.** En este caso la velocidad de transmisión está limitada por la estación emisora, debida a la velocidad de su procesador, o de su disco, de su placa de red, o de su módem. También puede estar limitada por el hecho que la aplicación que estableció la conexión TCP no tenga datos para transmitir (por ejemplo: un telnet.)

- **FC: limitación por control de flujo.** Esta limitación se da cuando el receptor no está en condiciones de aceptar todos los datos que le envía el emisor. Puede darse en el caso de procesadores lentos o de aplicaciones que no leen de su socket a suficiente velocidad, provocando saturación de su buffer de recepción.

- **NC: limitación por congestión de red.** Esto se da cuando el factor limitante es la capacidad de la red. Tanto el emisor como el receptor están en condiciones de aumentar la velocidad, pero la red no puede soportarlo.

Solo en el caso NC es que la conexión está demandando el máximo de capacidad disponible. Es cuando TCP hace gala de su elasticidad y capacidad de adaptación a las condiciones de tráfico, tratando de minimizar la pérdida de datagramas por saturación, a la vez que minimiza el tiempo de duración de la transmisión.

Vale la pena notar que el proceso de adaptación no es perfecto sino que hay un tiempo transitorio de adaptación y que además no se logra utilizar el 100% de la capacidad disponible sino un valor cercano a ella. Una buena introducción a los mecanismos de

adaptación de TCP se puede encontrar en [2]. Propuestas de mecanismos más avanzados de adaptación se pueden ver en [1].

La mayoría de las veces (y en todos los casos de interés) la capacidad disponible debe compartirse entre muchas conexiones TCP, las cuales se reparten más o menos homogéneamente esta capacidad. La eficiencia en el aprovechamiento de la capacidad disponible también se ve afectada por el número de conexiones que la están compartiendo. Esto se debe a los transitorios de adaptación a los cambios de capacidad disponible que se producen cuando se agrega o se termina una conexión del conjunto. Estos casos se ven en la figura siguiente.

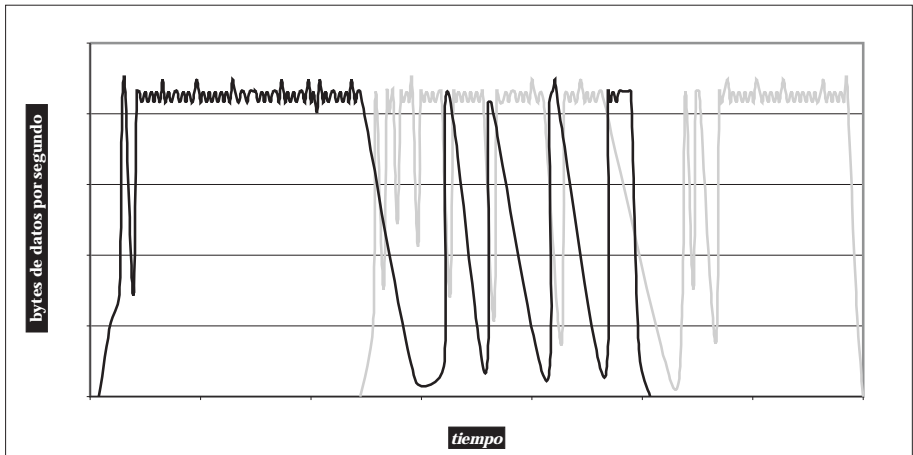


Figura 1: ritmo de transmisión instantáneo de 2 conexiones TCP que comparten un canal. El ritmo instantáneo está calculado como la longitud del segmento TCP dividido el intervalo de tiempo entre el segmento y el segmento anterior de la misma conexión

Además, la capacidad disponible para TCP se ve afectada por el tráfico no TCP que generalmente ocupa una capacidad que no depende de las condiciones de la red, es decir que no es elástica. Por ejemplo RTP, etc.

Tráfico de interés y ubicación del analizador

Muchas veces interesa medir solo una parte del tráfico observado por la estación del analizador. Por ejemplo el tráfico entre una universidad y países del extranjero, entre dos clientes específicos de un proveedor de servicios, o entre un cliente y el propio país en donde se encuentra, etc.

Es por ello que el analizador debe clasificar el tráfico observado como perteneciente o no a una **comunidad de interés**. Por comunidad de interés se entiende el tráfico entre un grupo de bloques de direcciones de IP y otro grupo de bloques de direcciones de IP.

El analizador debe ubicarse de manera que esté en condiciones de escuchar **todo el tráfico** entre estos grupos. Es decir, que el hecho que un datagrama llegue a destino

implique que el analizador lo haya visto y clasificado. Esto es necesario para poder seguir el estado de las conexiones.

Capacidad disponible

En las redes IP pueden existir muchos caminos entre dos grupos de bloques de direcciones IP. En la figura 2 se muestran algunos ejemplos.

La capacidad disponible total no siempre se ve fácilmente en un gráfico, sino que depende de la topología y de la capacidad de cada segmento, que puede o no convertirse en cuello de botella para una porción del tráfico. A veces tampoco se conoce la topología completa.

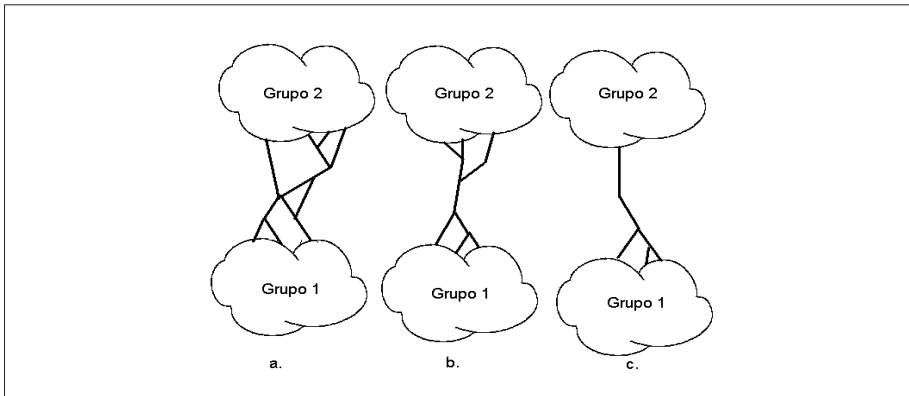


Figura 2: ejemplos para el camino que pueden seguir los datagramas de una comunidad de interés. En los casos b y c puede encontrarse un lugar en donde ubicar el analizador y observar todo el tráfico de interés.

Sin embargo se ve en las figuras 2.b y 2.c que hay casos (y son la mayoría) en donde se puede ubicar el analizador en un segmento por el que se observe todo el tráfico entre los grupos de interés, y de esta manera estimar la capacidad disponible utilizando los métodos tratados aquí.

Muchas veces la capacidad disponible se ve determinada por la existencia de tráfico ajeno al tráfico de interés, pero que comparte alguno de los segmentos por donde circula este. Esto hace que la capacidad disponible varíe en el tiempo más o menos rápidamente. (La medición y auditoría de estas variaciones en el servicio de un ISP fue una de las motivaciones para comenzar este trabajo.)

Máxima demanda

Como se mencionó anteriormente, la capacidad de adaptación a la capacidad disponible y a las condiciones de tráfico de TCP no es perfecta sino que hay tiempos de adaptación a la misma. Estos tiempos dependen del retardo entre extremos y de la

capacidad disponible. Además la aparición de nuevas conexiones y la desaparición de otras hacen variar la capacidad disponible para cada conexión, como puede observarse en la figura 1.

Sin embargo si hay una suficiente cantidad de conexiones TCP simultáneamente en estado NC durante un intervalo de tiempo mínimo determinado, entonces podemos estar seguros que se está demandando toda la capacidad disponible. Llamaremos a estos intervalos: **intervalos de máxima demanda**.

En algunos casos la comunidad de interés puede ser la formada por las comunicaciones entre un grupo particular (por ejemplo una universidad o una empresa) y el resto de la Internet. En estos casos la capacidad disponible estará determinada por la capacidad de acceso del grupo en cuestión a la Internet. Muchas conexiones no estarán limitadas por la capacidad de este acceso, sino por la capacidad de acceso de la red donde se encuentre el otro host en la Internet. Esto da una pauta que hace falta un buen número de conexiones TCP en estado NC para disminuir a un valor aceptable el riesgo de estimar erróneamente la capacidad disponible.

Normalmente en los gráficos de tráfico comunes se observa la evolución del promedio del tráfico tomado de a intervalos de duración fija. En la figura 3 hay un ejemplo con intervalos de 5 minutos.

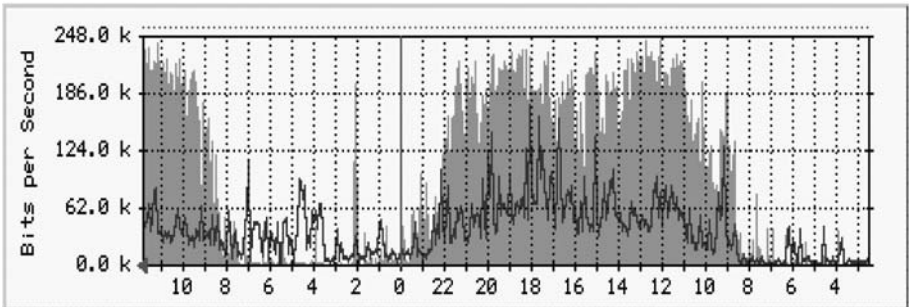


Figura 3: Evolución del ritmo transmisión promedio por un acceso a internet. Promedios tomados cada 5 minutos.

El hecho que durante un intervalo, el valor del promedio sea menor que la capacidad disponible se debe generalmente a que los intervalos de máxima demanda fueron de menor duración al intervalo del promedio utilizado. También es posible que no haya habido intervalos de máxima demanda sino solo tráfico inelástico, pero con un analizador como el descripto aquí no existirá esa ambigüedad.

Seguimiento individualizado de sesiones

El seguimiento individualizado de sesiones es necesario para saber por ejemplo en el caso de conexiones TCP si éstas están en estado NC, FC o DG, y así detectar intervalos

de máxima demanda. En próximos trabajos a realizar en la UP se agregará el seguimiento de flujos o sesiones de tráfico streaming como por ejemplo Real Time Protocol (RTP), etc.

El seguimiento individualizado permite a los administradores conocer otras cosas además de la capacidad disponible. Entre ellas: pérdida de datagramas, delay y throughput. La estimación de pérdidas y de delay utilizada aquí es solo aproximada (según veremos después) pero da información muy útil a los administradores. La estimación del throughput en cambio es muy buena ya que al seguir el estado se pueden distinguir las retransmisiones de datos y separar bytes de overhead de TCP/IP (incluyendo opciones) de bytes de datos.

Las limitaciones estarán dadas por capacidad de procesamiento de la estación de medición, por lo que se prestó especial atención a la eficiencia de los algoritmos en la implementación del analizador. Las experiencias en la UP muestran que una PC moderna puede manejar algunas decenas de Mbps en tiempo real, aunque para ello algunos análisis finos deban relegarse para su procesamiento off-line.

Detección de estado

Interesa seguir el **estado de la conexión TCP** a la vez que seguir su **estado de limitación**.

Al seguir el estado de la conexión se pueden detectar las pérdidas de datagramas, las retransmisiones, las variaciones del tamaño de los segmentos TCP, y los intervalos de tiempo entre segmentos. Esta información será utilizada para deducir el estado de la conexión como veremos más adelante.

En algunas conexiones se utilizan opciones como explicit congestion notification (ECN) o selective acknowledgments (SACK) que son muy útiles para hacer más exacto el análisis de las limitaciones de la conexión, pero como todavía representan una porción poco significativa del tráfico de Internet, no se utilizaron para la detección de limitaciones. En lugar de ello se usaron criterios más generales. La descripción de estas opciones en TCP puede verse en [3] y [4].

No es posible hacer una detección de limitaciones perfecta. Por ello se buscaron criterios que con una probabilidad muy alta acierten en la estimación de la limitación de cada conexión estudiada.

Cabe aclarar que el estado de limitación se estimará para el flujo en cada dirección de la conexión, ya que un extremo puede estar transmitiendo intermitentemente mientras que el otro no transmite ningún dato, o transmite continuamente.

NC: detección de limitación por capacidad de red

Cuando TCP tiene datos a su disposición para transmitir, prueba continuamente la capacidad de la red, aumentando el ritmo de transmisión de segmentos, más o menos rápidamente, hasta alcanzar y superar la capacidad disponible de la red. A partir de allí se producen pérdidas de segmentos que al ser detectadas hacen que TCP reduzca su ritmo de transmisión a un valor menor a la capacidad de la red. Luego comienza

nuevamente a aumentar el ritmo, repitiendo el ciclo. Esto mantiene el *ritmo promedio* de transmisión en un valor levemente menor a la capacidad disponible, a la vez que mantiene la *pérdida de segmentos TCP* en un valor bajo (para una descripción más extensa de los procesos de adaptación de TCP ver por ejemplo [1] y [2].)

Hay algunos casos en donde esto puede estabilizarse en un valor bastante menor a la capacidad disponible debido a que TCP emite los segmentos de ráfagas que superan el ritmo promedio de emisión, provocando saturaciones pasajeras en nodos de la red que no se hubieran producido de haber mantenido un ritmo continuo [1]. Pero suponemos que este problema no es tan frecuente en los casos de interés, donde se analiza el tráfico de muchas conexiones simultáneas. Y por lo tanto no tendremos en cuenta.

Mientras TCP actúa de esta manera diremos que está en estado de limitación por capacidad de la red (NC), ya que su ritmo promedio de transmisión no aumenta más debido a una limitación de la red.

La detección de este estado por parte del analizador estará basada en la *observación de pérdidas de segmentos*.

Existen pérdidas de segmentos debido a saturaciones y debido a errores en las transmisiones. Las primeras siempre indican un problema de congestión, e inducirán al analizador a la conclusión correcta de estado de limitación. Las segundas inducirán al analizador erróneamente a concluir que hay un problema de capacidad cuando no lo hay. Sin embargo, el TCP corriendo en el host transmisor también concluirá que hay un problema de congestión y disminuirá su ritmo de transmisión consecuentemente.

Entonces si bien no será correcto decir que si hay pérdidas, siempre estamos ante una limitación por congestión, podemos decir que estaremos ante una limitación por red (además las pérdidas por errores son mucho menos comunes que las pérdidas por saturaciones.)

DG: detección de limitación por ritmo de generación de datos

Cuando TCP vacía su buffer de transmisión, TCP cesa de transmitir segmentos (excepto para reconocer datos llegados desde el otro extremo, en cuyo caso emite ACK's - segmentos sin datos.) Por lo tanto cuando el ritmo de transmisión alcanzado en estado NC supera el ritmo de generación de datos por parte de la aplicación, TCP no superará la capacidad de la red y no producirá pérdidas por congestión (sí pueden producirse pérdidas por errores pero las suponemos bajas.)

Existe otro tipo más de limitación por generación de datos, que se da por ejemplo cuando el emisor se conecta a la red por medio de un módem. En este caso no se producirán pérdidas por congestión ya que el host emitirá segmentos a un ritmo mucho menor al que la red soporta (siempre que la red no esté saturada) por una limitación impuesta por su propio dispositivo de conexión a la red.

Por lo tanto incluimos dentro de los casos de limitación por generación de datos (DG) a los casos de limitación por la aplicación, por los dispositivos propios del host, por el sistema operativo, y por el canal de acceso a la red.

La detección de este estado estará basada en varios hechos. En primer lugar el host en estado DG no transmite a pesar que el otro extremo le da permiso, es decir que no lo

bloquea publicándole una ventana de transmisión nula. En segundo lugar es poco probable que vacíe su buffer de transmisión justo con un segmento que contenga una cantidad de datos igual a la máxima (MTU), por lo que se verán *variaciones en el tamaño* de los segmentos. Por último no produce pérdida de segmentos por congestión (excepto pérdidas por errores o por saturaciones causadas por otras conexiones, que supondremos bajas.)

FC: detección de limitación por control de flujo

Cuando el host emisor y la red son ambos mas rápidos que el host receptor, el ritmo de la transferencia de información esta limitado por el receptor. Este le informara al emisor que se detenga frecuentemente para poder leer y procesar la información, usando los mecanismos para el control de flujo que ofrece TCP.

Este estado se llamara **FC rígido** o **hard FC** y se puede detectar fácilmente porque el emisor publica una *ventana de recepción nula* cuando tiene lleno su buffer de recepción como para aceptar mas segmentos de datos.

Existe otro caso de limitación por control de flujo que no esta relacionado con la velocidad de procesamiento en el host sino que tiene que ver con una reserva de poca memoria (menos de la necesaria) para el buffer de recepción. Cuando el tiempo necesario para llenar la ventana de recepción a un ritmo de transmisión igual a la capacidad disponible de la red es menor que el round trip time (el doble del tiempo de retardo entre extremos: $2d$), *TCP detendrá su transmisión periódicamente* disminuyendo el ritmo de transmisión a un valor menor al que alcanzaría con sus mecanismos de control de congestión (valor que alcanzaría en estado NC.) Esto es debido a que el emisor asumirá que ha llenado el buffer de recepción del receptor con los segmentos ya emitidos pendientes de reconocimiento, cuyos ACK's no le llegan a tiempo por el problema descrito a principio del párrafo.

Este estado se llamara **FC blando** o **soft FC**, y no es fácil de detectar ya que no es fácil conocer la capacidad disponible para una conexión individual (si se puede estimar para una cantidad grande de conexiones), ni es fácil conocer el delay entre extremos. Si bien se puede medir el delay entre el analizador y el receptor, no es fácil medir el delay entre el emisor y el analizador, a menos claro esta que haya suficiente transferencia de información en ambas direcciones, o que se alcancen también estados de hard FC que al destrabarse permitan medir el tiempo de respuesta del emisor ante la apertura de la ventana.

Por ello la detección del estado de soft FC estará basada en métodos menos directos que los anteriores, en base a la observación de una dinámica periódica particular causada por el proceso descrito anteriormente. Para ello es necesario distinguir esta dinámica temporal del jitter introducido por los routers y las redes en el camino entre extremos. El hecho que los routers hagan caching de las decisiones de ruteo recientes, estabiliza el camino para cada conexión en particular, lo que facilita un poco esta estimación.

Tercera parte

Construcción de un analizador

Se ha construido un analizador en el Departamento de Comunicaciones de la Universidad de Palermo que implementa las ideas tratadas anteriormente. El analizador se llama Limiting State Analyzer (LSA.)

LSA clasifica el tráfico como perteneciente a las diferentes comunidades de interés, configurables como grupos de varios bloques de direcciones de IP, y obtiene métricas independientes para cada una de ellas. Estas métricas se obtienen de analizar los protocolos IP, ICMP, UDP, y TCP. Puede seguir el estado de todas las conexiones TCP con su estado y sus limitaciones individuales en cada dirección. Detecta todos los estados de limitación descriptos excepto soft FC que es el menos común de los estados ya que se da solamente en casos que hay simultáneamente un enorme delay y una enorme capacidad disponible (bit rate promedio) para cada conexión individual.

LSA realiza todos los análisis en tiempo real para los niveles de tráfico máximos en una LAN de 10Mbps. Puede funcionar sobre plataformas Linux y FreeBSD.

Actualmente se esta trabajando en la detección de intervalos de máxima demanda de corta duración, y en la representación de sus salidas.

Los detalles constructivos de LSA se están por publicar próximamente.

Conclusiones

A través de estos métodos de análisis de tráfico y de demanda se logra obtener información de suma utilidad para los administradores de redes, como por ejemplo capacidad disponible, pérdida de datagramas y retransmisiones, throughput, proporción de tráfico elástico e inelástico, y otros datos. Si bien existen otras alternativas de medición de estos datos, necesitan generalmente de acceso administrativo a los elementos de red que se encuentran en redes de proveedores u otros dominios, que conducen y afectan el tráfico de interés.

Se puede auditar contratos de nivel de servicio (SLA's) con proveedores sin necesidad de su cooperación, y sin introducir tráfico con mediciones activas que afecten el servicio a la comunidad de interés.

Además un analizador como el LSA es una muy buena herramienta para la investigación de tráfico, comportamiento, eficiencia de redes IP y protocolos. Permite por ejemplo comparar throughput entre conexiones individuales con diferentes implementaciones de TCP que convivan en una misma agregación de tráfico. O comparar throughput de agregaciones de tráfico sobre accesos a internet con diferentes combinaciones de delay y bit rate, etc.

Actualmente se está analizando extender el análisis y seguimiento de estado de sesiones de tráfico streaming (RTP), y de tráfico tipo consulta respuesta (DNS), para ver pérdida de datagramas en tráfico no elástico, etc. Asimismo se extenderán los

criterios de especificación del tráfico de interés en LSA para que discrimine entre tráfico web, de correo, etc.

Referencias

- [1] *Geoff Huston, "TCP Performance", Internal Protocol Journal, Vol 3, Number 2, June 2000, Cisco.*
- [2] *Douglas Comer, "Internetworking with TCP/IP Volume I; principles, protocols and architectures", 2nd edition, 1991, Prentice Hall.*
- [3] *M. Mathis, J. Mahdavi, S. Floyd, A. Romanow, "TCP Selective Acknowledgment Options", RFC 2018, October 1996, IETF.*
- [4] *K. Ramakrishnan, S. Floyd, "A Proposal to add Explicit Congestion Notification (ECN) to IP", RFC2481, January 1999, IETF.*