

# Internet: rutas más fiables en la red de redes de comunicaciones

Por Iljitsch van Beijnum, Research Assistant IMDEA Networks, Instituto Madrileño de Estudios Avanzados en Networks

El uso simultáneo de varios caminos a través de una red de comunicaciones puede proporcionar muchas ventajas, como un mejor uso de la red y una mayor solidez. Una parte fundamental de una arquitectura de red multicamino es la capacidad de enrutar protocolos para instalar múltiples rutas a múltiples caminos en la tabla de enrutamiento. El BGP o Border Gateway Protocol es un protocolo mediante el cual se intercambia información de encaminamiento entre sistemas autónomos. Por ejemplo, los ISP registrados en Internet suelen componerse de varios sistemas autónomos y para este caso es necesario un protocolo como BGP. En este artículo proponemos cambios en el procesamiento del BGP local que permiten que un *router* BGP utilice múltiples caminos simultáneamente sin poner en peligro la libertad de bucle.



## ***Libertad de bucle en el BGP multicamino mediante la propagación del camino más largo***

### INTRODUCCIÓN

En el Internet actual es bastante frecuente que haya múltiples caminos en la red entre un origen y destino determinados. Los protocolos de enrutamiento se encargan de seleccionar el mejor camino para llegar a cada destino. Sin embargo, esto hace que a menudo queden otros caminos sin utilizar en la red. La utilización de varios caminos simultáneos para enviar paquetes a un destino único tiene muchas ventajas. En una red multicamino se reducen los casos en los que enlaces disponibles se quedan sin utilizar porque no se dirige tráfico hacia ellos. También se reduce la necesidad de optimizar el flujo de tráfico manualmente (*ingeniería de tráfico*). Se pueden conseguir más ventajas comunicando a los protocolos de transporte (como TCP) los múltiples caminos para que dirijan los flujos o subflujos a los diferentes caminos. En ese caso, los usuarios consiguen una mayor solidez ya que la respuesta a los fallos para un subconjunto de caminos disponibles a un destino se puede gestionar en escalas de tiempo de transporte, que tienden a ser mucho más cortas (segundos) que las escalas de tiempo de enrutamiento (medio minuto o más).

No obstante, para poder utilizar varios caminos es necesario que los protocolos de encaminamiento sean multicamino. Los protocolos de enrutamiento IP ya admiten el uso de múltiples caminos en su diseño o en su implementación. OSPF [1] admite multicaminos del mismo costo; EIGRP [2] es capaz de utilizar varios caminos con diferentes costos. Aunque la especificación no lo admite, muchas implementaciones del BGP [3], como las de Cisco y Juniper, son capaces de utilizar simultáneamente múltiples caminos del mismo costo. Pero como la definición actual del protocolo BGP no admite multicamino, existe el riesgo de enrutar bucles o de que los caminos de bucle libre no se consideren aptos para el uso del multicamino.

Proponemos realizar cambios en las normas de selección y diseminación de caminos del BGP que permitan el uso simultáneo de una amplia selección de caminos sin poner en peligro la libertad de bucle. Como un *router* que ejecuta el BGP recibe múltiples caminos para llegar al mismo destino desde diferentes *routers* vecinos, puede seleccionar, en base a

unos criterios de política, un subconjunto de los caminos recibidos para utilizarlos simultáneamente. A continuación diseminamos el camino que el BGP considera más largo en la dirección de *downstream* (de un ordenador remoto al nuestro). Aunque no parezca lógico diseminar un camino más largo, este procedimiento tiene la propiedad de permitir utilizar al *router* todos los caminos que tengan una longitud AS\_PATH más corta o igual sin poner en riesgo los bucles.

No obstante, este cambio implica que no seguirá habiendo una relación de uno a uno entre los caminos que los paquetes siguen a través de la red y el camino que se anuncia en el BGP. Según los observadores de la vanguardia de la red, la consecuente confusión de la topología de la red se puede considerar dañina por los que quieren estudiar las redes o aplicar políticas basadas en la presencia de unos dominios intermedios determinados, o útil por los que están dispuestos a esconder los trabajos internos de su red. Nos limitamos a la situación en la que un *router* BGP individual equilibra el tráfico de forma local a múltiples caminos, sin cambiar la semántica del protocolo BGP, de modo que los cambios se puedan desplegar gradualmente en *routers* individuales que después puedan aprovechar las ventajas del multicamino sin exigir cambios en otros *routers* BGP.

## EL BORDER GATEWAY PROTOCOL

En los últimos 15 años, BGP-4 ha sido el protocolo de enrutamiento entre dominios que se ha usado en Internet [4]. BGP es un protocolo de enrutamiento exterior que se ejecuta entre dominios de enrutamiento. En el BGP estos dominios se llaman sistemas autónomos y se diferencian por un número SA. En la mayoría de los casos, un SA equivale a una red de un proveedor de servicios de Internet (ISP), aunque es posible que los usuarios finales que se conectan a varios ISP también constituyan un SA propio.

El BGP se clasifica como un protocolo de enrutamiento de *vector de caminos*, estrechamente relacionado con los protocolos de vector de distancia, que incluyen RIP y EIGRP. Al añadir un camino al BGP se mejora la detección de bucles de enrutamiento: cuando un *router* ve el número propio de un sistema autónomo (SA) de la red en un camino, interpreta que el camino traza un bucle y lo rechaza. Asimismo, a diferencia de los protocolos de enrutamiento interior como RIP, EIGRP y OSPF, el BGP permite que los operadores de red expresen su política. Esto es fundamental para un protocolo de enrutamiento entre dominios: puede que el camino más corto entre dos ISP sea a través de un cliente mutuo. No obstante, esta forma de utilizar un camino hasta un cliente viola la relación comercial entre el cliente y sus ISP, de modo que el protocolo de enrutamiento entre dominios debe configurarse para que ignore este camino.

Conceptualmente, el protocolo BGP selecciona el mejor camino calculando un grado de preferencia para todos los caminos válidos a un prefijo de destino determinado (grupo de direcciones IP) que reciben los altavoces del BGP en dominios de enrutamiento vecinos. Después selecciona el camino que tenga el mayor grado de preferencia (expresado como el atributo LOCAL\_PREF). La especificación del BGP no autoriza a una función o algoritmo a calcular el grado de preferencia; en la práctica, la preferencia se deriva de unas normas de política configuradas por el administrador. Durante el proceso de selección del camino, el BGP selecciona el camino o caminos con el LOCAL\_PREF más alto para incluirlos en la tabla de enrutamiento y utilizarlos para enviar paquetes.

## VALLEY-FREENESS

En [5], Gao introduce el concepto de *valley-freeness* (libertad de valle). En el modelo *valley-free*, hay tres tipos de relaciones entre sistemas autónomos: una relación proveedor-cliente,

una relación de pares y una relación de hermanos. En la relación proveedor-cliente el proveedor proporciona conectividad a todo Internet al cliente, y a los clientes de su cliente, si los hay. Se trata de una relación común, ya sea entre ISP comerciales y sus clientes, o entre proveedores de servicio académicos o gubernamentales y sus usuarios.

En una relación de pares, los dos SA intercambian información de enrutamiento y, por tanto paquetes, con un SA (o un cliente de ese SA) como origen y el otro SA (o un cliente de ese SA) como destino. Esto suele suceder entre proveedores de servicios de Internet (ISP) de similar envergadura sin que haya traspaso de dinero. En una relación de hermanos, cada SA le proporciona al otro conectividad segura. Esta relación puede darse si dos SA tienen el mismo propietario o mantienen una relación estrecha de otro tipo.

Las políticas y los caminos que éstas permiten por la red son *valley-free* si un camino sólo tiene una conexión de pares y si, después de la conexión de pares, las otras conexiones del camino son enlaces proveedor-cliente en la dirección proveedor a cliente o enlaces de hermanos. Después de un enlace proveedor-cliente, no se permiten enlaces cliente-proveedor. En la figura 1, el camino de A a B a través de los ISP 2, 1 y 4 es *valley-free*, pero no sucede lo mismo con el camino a través de los ISP 2, 3 y 4.

Por lo general, se entiende que sólo pueden considerarse correctos los caminos *valley-free*. Puede haber caminos que no sean *valley-free*, pero dejan uno o más SA en el camino sin compensación por transportar tráfico. Por ejemplo, en el camino que no es *valley-free* A-2-3-4-B de la figura 1, el ISP 3 transporta tráfico de A a B, pero al ISP 3 no se le compensa por ello, porque ni el SA A ni el SA B es un cliente directo o indirecto del ISP 3.

## MODIFICACIONES MULTICAMINO DEL BGP

Para cada destino posible un *router* puede aprender un camino a dicho destino desde cada *router* BGP vecino. No obstante, cuando hay varios caminos con un mismo valor LOCAL\_PREF, el BGP utiliza específicamente un grupo de siete normas de desempate para definir un único camino hacia cada destino. A fin de utilizar varios caminos para un destino, seguimos las normas de selección de camino del BGP, concretamente la norma de seleccionar sólo caminos que compartan el mayor LOCAL\_PREF, hasta llegar a las normas de desempate, que eliminamos. Aun así, hay otras normas que determinan qué caminos se pueden utilizar para enviar paquetes y más normas que determinan cuáles de estos caminos se diseminan para los vecinos. Para ser breves, sólo describimos las normas más importantes.

## LIBERTAD DE BUCLE

En una situación normal, el atributo AS\_PATH del BGP garantiza la libertad de bucle: un *router* rechaza todos los prefijos con su propio SA en el AS\_PATH. Dado que nuestros cambios permiten que el BGP utilice varios caminos de forma simultánea, pero sólo se propaga un camino a SA vecinos, ya no basta con comprobar el AS\_PATH para la incidencia del número SA local para evitar bucles. En su lugar, dependemos de las condiciones de Vutukury-García-Luna-Aceves Loop-free Invariant (LFI) [6].

De forma intuitiva, estas condiciones son muy sencillas: como un *router* sólo puede utilizar caminos que tengan un costo inferior al del camino que propaga a sus vecinos (o, puede que sólo propague un camino con un costo mayor al de los caminos que utiliza), no puede haber bucles. Un bucle se produce cuando un *router* utiliza un camino que ya ha propagado anteriormente, en cuyo caso el camino que utiliza debe tener tanto un costo mayor como otro menor que el camino que disemina, situaciones que evidentemente no pueden darse al mismo tiempo.

## A. Propagar caminos de bucle libre en eBGP

Todos los caminos que quedan en el grupo multicamino después de realizar los pasos previos y aplicar la política se instalan en la tabla de enrutamiento y se utilizan para enviar paquetes. La determinación de los ratios de división de tráfico entre los caminos disponibles es un tema para futuras investigaciones.

En este punto, el camino con el AS\_PATH más largo del grupo es el que se selecciona para la propagación a los vecinos del BGP. Si hay varios caminos con la longitud AS\_PATH máxima, el camino que se propagó previamente a vecinos del BGP, si los hay, es el que se selecciona para la propagación. Esto produce un efecto amortiguador de las oscilaciones en los caminos más cortos.

## CONVERGENCIA

De forma intuitiva, es fácil ver que las topologías del BGP con políticas contradictorias presentan problemas de convergencia [7]. Por ejemplo, si A prefiere enviar tráfico a través de B, pero B prefiere enviarlo a través de A, la detección de bucles del BGP se asegurará de que los dos envíos no se produzcan al mismo tiempo, pero no se podrá alcanzar un estado estable y convergido: el estado final depende del orden de los acontecimientos.

Por otro lado, cuando se respetan las limitaciones de la libertad de valle, se garantiza la convergencia a un estado estable porque no hay ciclos en las políticas configuradas. Esto significa que cuando un SA selecciona un camino, las decisiones posteriores de los SA *upstream* (de nuestro ordenador a uno remoto) no harán que el primer SA seleccione un camino diferente.

Al igual que el BGP estándar, nuestras normas de selección de camino exigen que sólo se incluyan caminos con el mayor LOCAL\_PREF al grupo de rutas candidatas. Como cada valor LOCAL\_PREF traza el mapa a una única clase *valley-free* (de hermanos, proveedor de servicios, de pares o cliente), nuestro uso de varios caminos no rompe la propiedad *valley-free* si estaba presente en el caso de camino único.

De modo que, en el caso de las topologías *valley-free*, la convergencia eventual está garantizada y es bastante parecida a la de una topología donde sólo existen los caminos más largos que se han utilizado. No obstante, es posible que haya más estados intermedios y las actualizaciones de dichos estado pueden activar el intervalo de anuncio de la ruta mínima, forzando los tiempos de convergencia hacia el máximo impuesto por este intervalo. Las topologías que no son *valley-free* no pueden converger nunca. La presencia de caminos más largos inyectados por *routers* multicaminos puede agravar esta situación ya que dichos *routers* intentan encontrar los caminos de bucle libre más largos que permite la política. La cuantificación de estos efectos forma parte de nuestro calendario de futuras investigaciones.

## EVALUACIÓN

Para evaluar el impacto de nuestros cambios del BGP, y en concreto para comprender la dinámica del sistema resultante, hemos creado un simulador que implementa nuestras normas del BGP modificado. El simulador es un *script* que produce el resultado de la toma de decisiones del BGP (multicamino) según una topología de entrada determinada. El *script* también puede simular la toma de decisiones en routers BGP tradicionales. Curiosamente, para el ejemplo de topología que hemos elegido, el número de iteraciones necesarias para que el BGP multicamino converja es en realidad ligeramente menor que el número de iteraciones que necesita el BGP tradicional. Esto se debe, probablemente, a que cada

*router* obtiene con ansia todos los caminos que puede, limitando las oportunidades de otros sistemas autónomos. Aunque es necesario entender mucho mejor la dinámica, los resultados obtenidos son prometedores, ya que los experimentos realizados muestran que el BGP multicamino propuesto converge en un número similar (si no menor) de iteraciones que el BGP actual y que consigue evitar caminos largos, todo esto utilizando topologías que constituyen un reto.

#### AGRADECIMIENTOS

Este artículo está basado en una tesis de máster del autor y en un artículo aún no publicado y redactado conjuntamente con Jon Crowcroft, Marcelo Bagnulo y Francisco Valera.

#### REFERENCIAS

- [1] J. Moy. OSPF Version 2. RFC 2328, abril de 1998.
- [2] R. Albrightson, J.J. García-Luna-Aceves, y J. Boyle. EIGRP—A Fast Routing Protocol Based On Distance Vectors. Proc. Networld/Interop 94, mayo de 1994.
- [3] Y. Rekhter, T. Li, S. Hares. A Border Gateway Protocol 4 (BGP-4). RFC 4271, enero de 2006.
- [4] I. van Beijnum. BGP: Building Reliable Networks with the Border Gateway Protocol. O'Reilly & Associates, Inc., Sebastopol, CA, 2002. ISBN: 0-596-00254-8.
- [5] L. Gao. On inferring autonomous system relationships in the Internet. IEEE/ACM Trans. on Networking, 2000.
- [6] S. Vutukury, J.J. García-Luna-Aceves. A Simple Approximation to Minimum-Delay Routing. Proc. of ACM SIGCOMM, 1999.
- [7] K. Varadhan, R. Govindan, D. Estrin. Persistent route oscillations in inter-domain routing. ISI technical report 96-631, USC/Information Sciences Institute, 1996.