

Hack Players

Cortes y lentitud en Internet a nivel global porque muchos routers no soportan más de 512k rutas BGP (o "la que se avecina")

PUBLICADO POR VICENTE MOTOS ON MIÉRCOLES, 13 DE AGOSTO DE 2014 ETIQUETAS: AMENAZAS , REDES

Ayer **usuarios de todo el mundo se quejaban de lentitud y algunos cortes en el acceso a Internet a través de múltiples proveedores:** Comcast, Level 3, AT&T, Cogent, Verizon, etc.

¿Se estaba produciendo un ataque DDoS a gran escala, un corte eléctrico en uno de los centros de datos de un gran proveedor de *cloud*, o quizás una avería en algún cable submarino de datos?... Ninguno de ellos: se trataba de un problema estructural en la forma en la que Internet está construida. Pero empecemos desde el principio...

BGP en Internet

Como todos sabéis y de forma muy genérica, Internet es una gran red global que interconecta cualquier host que se identifica por una dirección IP única. Esto se consigue reenviando los datos de un router a otro y, para conseguirlo, estos routers tienen que tener una tabla de rutas continuamente actualizada.

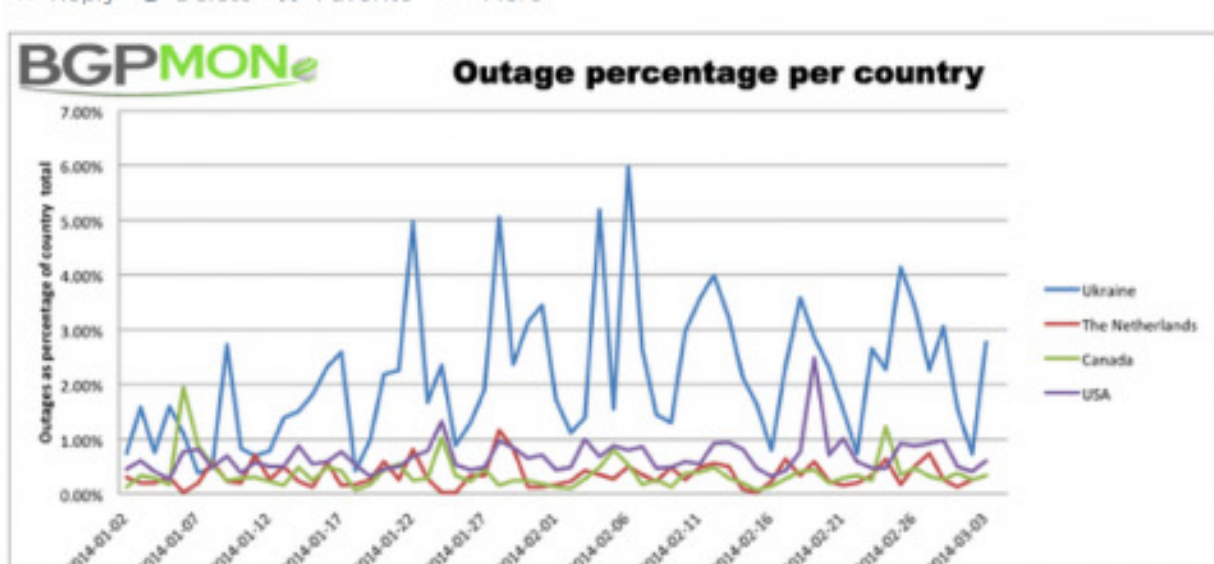
En esta tabla de rutas están las direcciones IP agrupadas en prefijos que pertenecen a distintos **Sistemas Autónomos (AS)**. Estos AS distribuyen sus rutas y se intercambian la información con otros AS mediante el protocolo **BGP (Border Gateway Protocol)**. Por ejemplo el prefijo de red 192.0.2.0/24 está dentro del AS 64496 y así lo anuncia al resto.

En el último año, de media en Internet se producían diariamente cortes o interrupciones de servicio que afectaban a 6.033 prefijos de 1.470 AS distintos. Es decir, estos cortes suelen ser normales y algunas redes y zonas geográficas son más estables que otras:



Curious how stable the Internet is in the Netherlands vs. Canada vs. USA vs. Ukraine? Attached graph shows outages

Reply Delete Favorite More



Sin embargo ayer y según BGPmon estas cifras subieron a 12.563 prefijos de 2.587 AS distintos, prácticamente el doble. ¿Y qué causó esta inestabilidad? Aquí está el meollo del asunto...

En cierta manera, todo pende de un hilo

En marzo de 2014 un **informe del CIDR** (Classless Inter-Domain Routing) incluía estadísticas en las que se mostraba que, **por primera vez en la historia de Internet, la tabla de rutas global había sobrepasado las 500.000 rutas**. Como se puede ver en la siguiente tabla, actualmente el número de rutas IPv4 en la tabla BGP que maneja cada AS ronda esa cifra:

Provider	Location	Number of IPv4 routes on full BGP feed
Level3	Amsterdam	497,869
GTT	Amsterdam	499,272
NTT	Amsterdam	499,610
Telstra	Sydney	491,522
NTT	Sydney	500,265
Singtel	Singapore	501,061
NTT	Singapore	499,830
Level3	Chicago	497,470
GTT	Chicago	499,326
NTT	Chicago	499,523

El problema es que **muchos proveedores manejan estas tablas con cores de red Cisco antiguos que por defecto soportan sólo 512K rutas en memoria**. La mayoría de las plataformas soportan tablas de rutas mayores, pero requieren ajustar la configuración.

Entremos en detalle

Las series Catalyst 6500 y 7600 almacenan la información de routing en una **memoria especial de alta velocidad llamada TCAM**. Hay distintos tipos de TCAM:

- Forwarding Information Base (FIB), or routing TCAM
- Access Control List (ACL) TCAM
- Netflow TCAM

Cuando se añade una ruta en la tabla Cisco Express Forwarding (CEF) de la memoria RAM, una segunda copia se almacena también en la memoria TCAM, en ya sea el supervisor o el módulo DFC (Distributed Forwarding Card) de la tarjeta. **Los módulos 3BXL por defecto reservan un espacio máximo de sólo 512k rutas:**

```
7600#show mls cef max
FIB TCAM maximum routes :
=====
Current :-
-----
IPv4 + MPLS           - 512k (default)
IPv6 + IP Multicast - 256k (default)
```

También he podido comprobar en uno de nuestros COREs, un 6509, que esta cifra es aún menor:

```
MAIN_CORE#show mls cef max
FIB TCAM maximum routes :
=====
Current :-
-----
IPv4 + MPLS           - 192k (default)
IPv6 + IP Multicast - 32k (default)
```

¿Y qué pasó exactamente ayer?

Pues que debido a un error Verizon, anunciaron en dos de sus AS (701 y 705) cientos de prefijos /24 provocando que **la tabla de rutas global alcanzara temporalmente las 515.000 entradas y provocando el fallo de numerosos equipos antiguos de muchos ISPs...**

El workaround

Como habéis visto anteriormente, la FIB TCAM es un bloque de memoria que comparten las rutas IPv4/MPLS y IPv6/Multicast:



Evidentemente, no se puede incrementar físicamente la suma total del espacio de esta memoria al menos que se sustituya el hardware (módulo supervisor y DFC). Pero si podemos realizar un **sencillo workaround en la FIB TCAM concediendo la memoria de IPv6/Multicast al espacio IPv4/MPLS mediante el comando 'mls cef maximum-routes ip'**.

```
Por ej.:
7600 (config)#mls cef maximum-routes ip 1000
Maximum routes set to 1024000. Configuration will be effective on reboot.
```

De esta manera y tras reiniciar el equipo aumentaremos la suma del espacio de la TCAM para IPv4 restándosele al resto.

¿Y ahora qué?

Ya sabemos que el límite de las 512k rutas globales está muy cerca y que muchos equipos fallarán de nuevo si se supera (que seguro volverá a ser pronto). Los ISPs están ya en preaviso, aunque por otro lado saben que el *workaround* les obliga como mínimo a reiniciar unos equipos muchas veces complejos y asumir cierto riesgo. Por otro lado saben que no tienen opción: o modifican la configuración o sustituyen el hardware. Mientras el tiempo corre inexorablemente y urge una solución...