Material de descarga - Módulo 2
Introducción a BGP

Introducción:

Bienvenidos al segundo módulo del curso “Fundamentos de BGP e introducción a RPKI” de LACNIC.
En esta oportunidad, se conocerán las características de BGP para luego ver los comandos básicos de configuración y análisis de algunos comandos más comunes.

Al final del módulo se podrá realizar una actividad práctica sobre los conceptos aprendidos, configurar una sesión BGP y analizar las tablas de BGP y de ruteo, entre otros comandos.

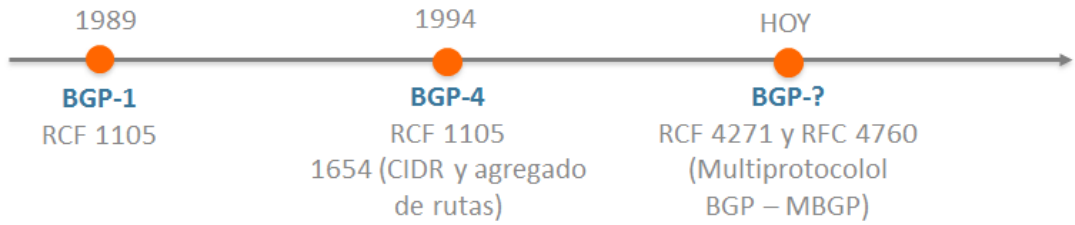
TEMA 1: Qué es BGP y cómo trabaja

BGP, por sus siglas en inglés de *Border Gateway Protocol*, es un protocolo de ruteo externo ó EGP por sus siglas en inglés de *Exterior Gateway Protocol*.

BGP se basa en el intercambio de información de ruteo al mismo tiempo que garantiza la elección de un camino libre de bucles o loops.

Esto lo logra básicamente porque se trata de un protocolo de ruteo del tipo “*path vector*”, o sea que tiene en cuenta el camino que hacen los paquetes para llegar a determinado destino, o lo que es lo mismo, considera un vector de sistemas autónomos.

Desde el año 1989 la versión que se utilizó fue BGP-1, la cual se definía en el RFC 1105. Fue desde el año 1994 que comenzó a utilizarse BGP-4, cuyo RFC que lo define es el 1654. Esta fue la primera versión de BGP que admitía el ruteo de prefijos CIDR (*Classless Interdomine Routes*) y el agregado de rutas. La RFC actual utilizada es la 4271 y la 4760 referida a BGP Multiprotocolo, o sea cuando BGP debe ser implementado bajo diferentes **address families**, como por ejemplo IPv4 e IPv6.



Existen dos tipos de BGP, el llamado **iBGP** y el llamado **eBGP**, según sea implementado dentro o fuera del Sistema Autónomo. De todas formas, no hay que adelantarse y habrá que esperar a ver el próximo tema para conocer mejor sobre estos dos tipos de BGP.

BGP trabaja estableciendo lo que se denomina “sesiones BGP”, para poder realizar el intercambio de información de ruteo. Las sesiones BGP se establecen entre dos dispositivos *routers*, los cuales pasan a denominarse “***neighbor***s” o “***peers***” BGP.

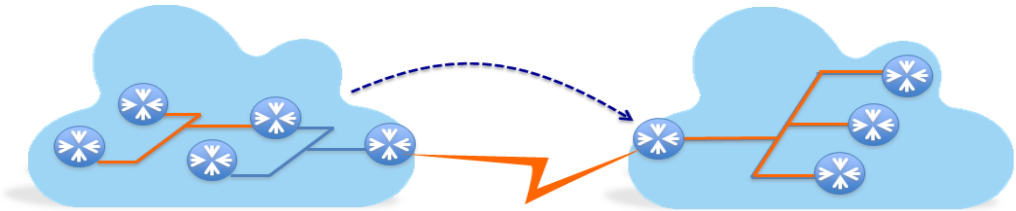
Para el establecimiento de la sesión, BGP utiliza el puerto TCP 179, en una sesión abierta que le permitirá el intercambio de ruteo entre los *neighbors*.

El comportamiento de BGP es básicamente el de aprender y enseñar rutas, tal como se ha visto en la introducción de este curso. Un *neighbor* enseñará a otro aquellas rutas de redes IP que desee que sean alcanzadas, ó que estén visibles desde Internet. Al mismo tiempo, sólo podrá aprender las rutas que otro *neighbor* le quiera enseñar, o sea, no podrá obtener más información acerca de las rutas de otra organización, más allá de las que el *router* de borde que actúa de *neighbor* le ofrezca.

Es importante notar que la única forma que una red tiene de hacer conocer sus prefijos es si los anuncia, sino, estos no podrán ser alcanzados desde Internet. Todo esto se hace a través de BGP sin mostrar la complejidad interna de la red de la organización, sino simplemente anunciando qué redes pueden ser alcanzadas a través de qué *neighbors*. Aquí se ve claramente el concepto de Caja Negra que ya se ha mencionado.

Aprender y enseñar rutas incidirá claramente en el tráfico que fluya entre las organizaciones. Cuando un *router* está aprendiendo rutas, está afectando su tráfico saliente, pues es la única forma que tendrá de saber hacia dónde encaminar los paquetes que salen de la organización. Análogamente, el hecho de que un *router* anuncie qué rutas pueden ser alcanzadas a través de él, incide en el tráfico entrante a la organización. Si

una organización no publica ninguna ruta a otras organizaciones, nadie sabrá cómo llegar a ella, y el tráfico entrante será nulo. La siguiente figura grafica la interacción entre los routers de organizaciones diferentes, que establecen una sesión BGP:



TEMA 2 - BGP interno y externo

Como se ha mencionado antes en este mismo curso, BGP es un protocolo de ruteo externo, o EGP, por sus siglas en inglés de *Exterior Gateway Protocol*.

Se trata de un EGP porque es capaz de comunicarse con *routers* externos a nuestro Sistema Autónomo, logrando la comunicación **inter-ASs**.

Asimismo, existen dos tipos de BGP según se encuentre implementado fuera o dentro de nuestro Sistema Autónomo.

De esta manera, se puede decir que BGP se divide en dos subtipos: exterior BGP o **eBGP**, e interior BGP ó **iBGP**, ambos por sus siglas en inglés.

Así, **eBGP** se refiere al protocolo BGP utilizado entre Sistemas Autónomos. En cambio, **iBGP** se refiere a la utilización de BGP cuando es dentro del mismo sistema autónomo.

Sin embargo, que BGP pueda ser utilizado dentro del sistema autónomo, no lo convierte en un Protocolo de Ruteo Interno o **IGP**, por lo que habrá que tener mucho cuidado para no confundir a un iBGP con un IGP.

Esto se debe básicamente a cómo fue concebido BGP. BGP es un protocolo diseñado para llevar información de rutas que están fuera de nuestra organización, no para llevar información del ruteo interno, la cual debe ser llevada por un IGP como ISIS, OSPF, EIGRP, entre otros que ya se han nombrado.

Por ejemplo, es correcto planificar la red de un ISP considerando llevar la información de infraestructura a través de un IGP, así como también la información de las diferentes redes internas del ISP. No obstante, la información de ruteo dentro del Sistema Autónomo, pero relacionada con los

clientes del ISP, debería llevarse a través de un iBGP. Naturalmente, la información de ruteo relacionada con otros Sistemas autónomos, ingresaría y saldría del ISP a través de eBGP.

Esta separación y utilización apropiada de cada protocolo para el conjunto de prefijos que corresponde es una buena práctica de diseño y facilita la operación de las redes.

En cuanto a la conexión física, dentro de un AS, los *peers* BGP no necesitan estar directamente conectados. En cuanto a las sesiones de eBGP, la configuración natural es que los *peers* estén directamente conectados, aunque no es un requisito obligatorio.

TEMA 3 - Configuración y anáilisis de sesiones BGP

En el desarrollo de este tema se podrá ver cómo se realiza una configuración básica de sesiones BGP y se analizarán algunos comandos que resultarán muy útiles.

Se comenzarán a ver los comandos que se involucran en la creación de una sesión BGP. Antes de comenzar, se deberá tener en cuenta que los ejemplos aquí mostrados están basados en *routers Cisco*, y que, para agilizar el avance de los temas, sólo se harán en IPv4. No obstante, en el Anexo de esta documentación se podrán encontrar los mismos ejemplos de configuración, pero implementados con IPv6.

Otro punto importante a tener en cuenta es que, al principio de este módulo se ha aprendido que una sesión BGP se establece entre dos *neighbors* o *peers*. Para que la sesión BGP pueda establecerse cada uno de ellos deberá ser alcanzable por IP desde el otro. Esto es muy importante ya que no tener conectividad IP entre los *peers* es una de las principales razones para que se frustre el establecimiento de una sesión BGP.

El primer paso para establecer una sesión BGP es crear el proceso BGP dentro del *router*.

En el caso de los *routers* Cisco sólo se permite un proceso BGP corriendo por *router*, lo que significa que un *router* sólo podrá pertenecer a un único Sistema Autónomo.

El siguiente paso será determinar contra qué *neighbor* se establecerá la sesión, teniendo en cuenta que éste puede estar dentro o fuera del sistema autónomo. En general, los *neighbors* externos comparten una subred, o sea

son adyacentes. Los internos en cambio, pueden estar en cualquier parte del Sistema Autónomo.

Con estos dos pasos, una sesión BGP podrá quedar establecida. Se verá cómo se llevan a cabo.

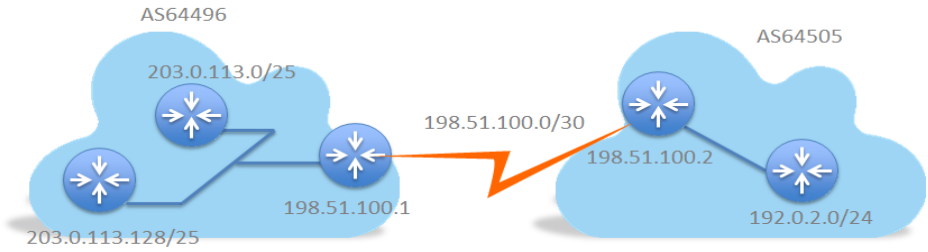
Como se mencionó anteriormente, el primer paso será crear el proceso BGP. Para ello, en modo de configuración global se escribirán las palabras: "router BGP" seguidas por el número de Sistema Autónomo local, o sea, el número de nuestro sistema autónomo. Esta acción habilitará un nuevo modo de configuración del *router*.

Luego, se tendrá que identificar el *neighbor* contra el que se quiere levantar la sesión BGP.

Para ello, será necesario introducir el comando **neighbor**, seguido de la dirección IP del potencial *neighbor* e indicando el sistema autónomo al que pertenece esa IP destino. Este valor será nuestro sistema autónomo, si estamos configurando una sesión iBGP o bien un sistema autónomo distinto, en cuyo caso se tratará de una sesión eBGP.

Notar que la creación del proceso sólo será necesario hacerlo una vez y al comenzar a configurar BGP, pues luego sólo será cuestión de levantar sesiones BGP contra distintos *neighbors*.

Veamos el siguiente ejemplo. Se supone que se intenta levantar una sesión BGP entre los *routers* que se muestran en la figura:



Configuración Router A

```
router bgp 64496
 neighbor 198.51.100.2 remote-as 64505
```

Configuración Router B

```
router bgp 64505
 neighbor 198.51.100.1 remote-as 64496
```

El *router* A que pertenece al ASN 64496, tendrá que configurarse tipeando: **router bgp 64496**. Luego, se indicará que se quiere conectar por BGP con el *neighbor* que está en el ASN 64505, así que se deberá introducir

el comando "**neighbor 198.51.100.2 remote-as 64505**".

Por otro lado, el administrador del *router* B deberá configurar: "**router bgp 64505**", y luego especificar el *neighbor* que corresponde al ASN remoto, colocando: "**neighbor 198.51.100.1 remote-as 64496**".

Al finalizar la acción explicada en el ejemplo anterior, una sesión eBGP deberá quedar establecida entre los *neighbors* 198.51.100.1 y 198.51.100.2. Para verificarlo, se puede salir del modo de configuración BGP y modo de configuración global, e intentar el comando "show ip bgp summary", tal como se puede ver en el ejemplo:

```
router> show ip bgp summary
BGP router identifier 198.51.100.1, local AS number 64496
BGP table version is 48347, main routing table version 48347
2558 network entries and 3869 paths using 389968 bytes of memory
527 BGP path attribute entries using 28600 bytes of memory
250 BGP AS-PATH entries using 6304 bytes of memory
1 BGP community entries using 24 bytes of memory
191 BGP route-map cache entries using 3056 bytes of memory
1028 BGP filter-list cache entries using 12336 bytes of memory
BGP activity 14929/71963 prefixes, 42175/38306 paths

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
198.51.100.2    4 64505   38357   31873    48347    0    0 6d08h     1342
203.0.113.2     4 64496   31728   34362    48347    0    0 6d09h     Active
```

Lo que muestra la salida del comando es un resumen de las sesiones BGP que el *router* tiene configurado. En este caso, se ve que el *router* A tiene 2 sesiones BGP configuradas, una con el *neighbor* 198.51.100.2 y otra con el 203.0.113.2. Una columna importante para visualizar es la que muestra el ASN al que pertenece el *neighbor* remoto, la tercera columna (recuadrada en azul en la figura). Claramente se puede deducir que la primera sesión es eBGP y la segunda sesión es un BGP interno, pues está levantado contra un *neighbor* del mismo sistema autónomo.

Otra columna muy importante para tener en cuenta sobre la salida de este comando es la última de todas, la denominada State/PrefijosRecibidos (recuadrada en color rojo). En el primer caso, claramente se están recibiendo 1342 prefijos o rutas a través de esa sesión BGP, para que el *router* A las aprenda, o lo que es lo mismo: las incorpore en su tabla de BGP. En el segundo caso, no muestra la cantidad de prefijos recibidos sino la palabra "Active". ¿Qué significa que diga "Active" en lugar de los prefijos recibidos?

Por el contrario de lo que pudiera parecer, "Active" en una sesión BGP significa que la sesión está "intentando" establecerse, pero no está establecida. En este caso, debe aparecer un número entero, aunque sea 0,

que indica la cantidad de prefijos recibidos. Se verá en la próxima diapositiva cuáles son los posibles valores del campo **State**.

Por último, siguiendo con otro de los datos arrojados por este comando que resultan de mucha utilidad, es interesante ver que el campo "UP/Down" está indicando cuánto tiempo hace que una sesión BGP está levantada (recuadro violeta). Este campo, naturalmente estará en concordancia con lo descrito anteriormente para el campo "State/PrefijosRecibidos".

Por supuesto que hay más información a la salida de este comando, que tiene que ver con la versión de BGP utilizada, las versiones de la tabla de BGP, cantidad de mensajes recibidos y enviados, entre otros. Pero los que se han descrito son los más utilizados para la resolución de problemas y la operación diaria de la red.:

Se ha visto hasta ahora cómo configurar una sesión BGP entre dos *neighbors* y cómo verificar su estado.

En este punto resulta interesante conocer otro comando que ayudará mucho en la administración de BGP, y es el comando "show ip bgp". Se verá cómo es la salida de este comando:

router> show ip bgp

BGP table version is 134358, local router ID is 198.51.100.1

Status codes: s suppressed, d damped, h history, * valid, > best, i – internal, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	192.0.2.0/26	198.51.100.1		0		64496 65511 i
*>	192.0.2.128/26	198.51.100.1		0		64496 64506 i
*>i		198.51.100.22	0	0		64506 e
*	203.0.113.128/25	198.51.100.1		0		64496 64501 i
*>		198.51.100.114	0	0		64501 i
*>	203.0.113.0/25	198.51.100.1		0		64496 ?

Al aplicarlo, se obtendrá un listado del contenido de la tabla de BGP, con las redes ordenadas en forma numérica (recuadro azul).

Las primeras tres columnas listan el *estado* de cada ruta (recuadr rojo). Un asterisco en la primera columna, indica que la ruta tiene un *next-hop*

válido, el cual se analizará más adelante. Existen más opciones para este campo, las cual son descritas en el encabezado de la salida del comando.

Luego, si la segunda columna tiene un signo mayor, significa que esa ruta fue seleccionada como el mejor camino hacia una red determinada.

Si la tercera columna está en blanco, indica que el router aprendió la ruta de un vecino externo. Una ruta aprendida de un vecino iBGP debería tener una letra **i**.

La cuarta columna (nuevamente el recuadro azul) lista las rutas que contiene esta tabla de BGP. Cuando no se especifica máscara de subred, se utilizará la máscara ***classfull*** (o sea, /8 para las clases A, /16 para las clases B y /24 para las clases C).

Cuando esta columna tiene campos en blanco, significa que se refiere al mismo prefijo que está justo en la fila anterior. Esto es así cada vez que un *router* recibe, a través de BGP, varios caminos para un mismo prefijo.

El resto de las columnas se mostrarán completas para cada instancia del prefijo.

Las siguientes columnas muestran atributos que se verán más adelante.

Ahora bien, se ha estado hablando en este curso de que BGP trabaja anunciando y aprendiendo rutas. Se ha dicho también que cuando un *router* anuncia una ruta por BGP, equivale a enseñar a otro *router* sobre una entrada que el *neighbor* tiene en su propia tabla de BGP.

Existen diferentes formas de incorporar una ruta a la tabla de BGP, y se verá cuáles son:

Las rutas que se incorporan a una tabla de BGP pueden ser:

- Aprendidas de otras sesiones BGP, ya se verá cómo y cuándo se anuncian en cada caso, pero lo importante en este punto es la forma en la que las rutas llegan a formar parte de la tabla de BGP del *router*.

- Configuradas dentro del proceso BGP el comando **redistribute <routing-protocol>**, el cual significa que todas las rutas del protocolo indicado en el campo **<routing-protocol>** serán transferidas dentro de BGP.

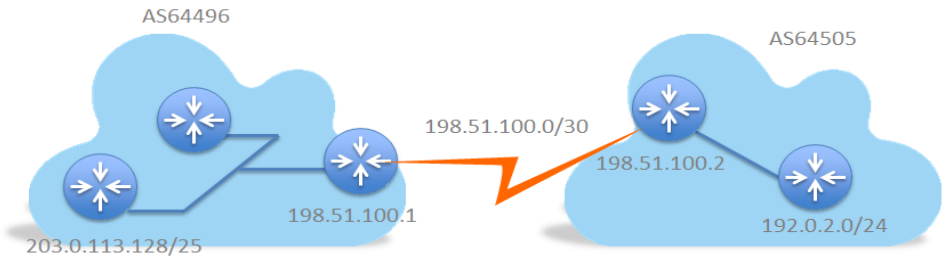
No es muy recomendable utilizar este comando, sobre todo porque suele perderse el control de los prefijos a anunciar y la administración se vuelve compleja.

- Finalmente, otra forma de incorporar prefijos en la tabla de BGP para

que puedan ser anunciados a un *neighbor*, es a través del comando **network**, especificando el prefijo a anunciar y su máscara.

Es muy importante tener en cuenta que para que el prefijo indicado en el comando **network** pueda ser anunciado, una ruta coincidente debe estar en la tabla de ruteo. Se verá en un ejemplo cómo se implementa el comando **network**.

Volviendo al esquema anterior:



Configuración Router A

```
router bgp 64496
    network 203.0.113.128 mask 255.255.255.128
```

Configuración Router B

```
router bgp 64505
    network 192.0.2.0
```

Se puede observar que el *router* A podría querer anunciar al *router* B una ruta para llegar a la red 203.0.113.128/25. Para ello, en el *router* A se deberá configurar: router bgp 64496, y luego, network 203.0.113.128 mask 255.255.255.128
De esta forma, el *router* B, estará recibiendo información desde el *router* A de cómo alcanzar a la red 203.0.113.128/25.

Si la máscara no es especificada, los prefijos se tomarán como clases A, B o C.

Si el *router* B deseara anunciar al *router* A sus redes, por ejemplo, la 192.0.2.0/24, la configuración sería, en forma análoga, la que se muestra en figura.

Algunas consideraciones respecto al comando **network** son las siguientes:

En primer lugar, no se debe olvidar el requisito que ya se ha mencionado acerca de que el prefijo a anunciar deberá estar en la tabla de

ruteo para que pueda ser efectivo el comando **network**. Esto podrá verificarse inspeccionando la ruta a anunciar dentro de la tabla de ruteo, o sea, a través del comando **show ip route** y deberíamos obtener una salida similar a la que se muestra:

```
R1#show ip route
.....
192.0.2.0/30 is subnetted, 4 subnets
C 192.0.2.0 is directly connected, FastEthernet0/1
O 192.0.2.4 [110/2] via 192.0.2.2, 00:10:04, FastEthernet0/1
O 192.0.2.8 [110/2] via 192.0.2.13, 00:10:04, FastEthernet0/0
C 192.0.2.12 is directly connected, FastEthernet0/0
```

Si la ruta no está en la tabla de ruteo, lo indicado en el comando **network** para esa ruta, será ignorado.

Otra consideración importante a tener en cuenta es que cuando un prefijo se inserta en la tabla de BGP a través del comando **network**, esto afecta a todas las sesiones BGP que se encuentran configuradas, lo cual debe ser tomado muy en cuenta pues se podría estar anunciando prefijos a *neighbors* a los que no se les desea anunciar. Se verá más adelante cómo manejar esta situación.

Cuando se configura una sesión BGP contra un *neighbor*, es muy recomendable confirmar efectivamente qué es lo que le estamos anunciando a ese *neighbor* y lo que aprendemos de él, de manera tal de no encontrarse con sorpresas de estar anunciando o recibiendo prefijos no deseados.

Para ello, existen dos comandos muy útiles que deberán ejecutarse para cada *router* que se desee verificar. Se verá cómo es esto.

Para saber qué rutas se le anuncia a determinado vecino, el comando a ejecutar será: **show ip bgp neighbor <ip_del_neighbor> advertised-routes**. Aquí se ve un ejemplo:

```
Show ip bgp neighbor 192.0.2.51 advertised-routes
BGP table version is 48402, local router ID is 192.0.2.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? – incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 192.0.2.64/26    0.0.0.0                  0         32768 i
*> 192.0.2.128/27   198.51.100.3             0             64496 64506 i
*> 192.0.2.160/27   198.51.100.22       0    0             64506 e
*> 192.0.2.192/27   198.51.100.35           100            65550 i
*> 192.0.2.224/27   198.51.100.2             0             65549 i
```

En el ejemplo se puede ver que la salida de este comando lista todos los prefijos que se anuncian a un determinado *neighbor*, el que especificamos en el encabezado del comando, junto con todos los atributos que acompañan al anuncio de cada prefijo. Así, el ejemplo muestra que el *router* 192.0.2.65, anuncia al *neighbor* 192.0.2.51, exactamente estos 6 prefijos, acompañados con los atributos de cada uno.

En forma análoga, para saber cuáles son las rutas que se reciben de un *neighbor*, se puede utilizar el comando **show ip bgp neighbors <ip_del_neighbor> routes**. En el ejemplo puede verse que el *router* 192.0.2.65, aprende del *neighbor* 192.0.2.51 seis prefijos para incorporar en su tabla de BGP:

```
Show ip bgp neighbor 192.0.2.51 routes
BGP table version is 48402, local router ID is 192.0.2.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? – incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i203.0.113.0/26	198.51.100.1	0	600	0	65549 i
*>i203.0.113.64/26	198.51.100.65	0	600	0	65547 65548 i
*>i203.0.113.128/27	198.51.100.129	0	600	0	65547 i
*>i203.0.113.160/27	198.51.100.161	0	600	0	65548 i
*>i203.0.113.192/27	198.51.100.193	0	600	0	65496 e
*>i203.0.113.224/27	198.51.100.225	0	600	0	65500 ?

Se ha visto acerca de cómo configurar una sesión BGP y comandos básicos, lo cual permitirá realizar las actividades de este módulo, antes de pasar al módulo 3 que trata sobre **Atributos de BGP**.