

Protocols d'Internet

Preguntes Topic 5. Inter-domain Routing (BGP)

Pregunta 1. Explica que es una política de encaminamiento y como se implementa.

Una política d'encaminament defineix el fluxe de rutes a través d'un sistema autònom, tan d'entrada com de sortida. Aquest fluxe és fruit de decisions preses pel AS que es basen en quines rutes seran anunciades als altres AS i quines rutes seran acceptades que passin a través d'ell. Són decisions que no responen únicament a decisions tècniques, sinó i sobretot, a decisions comercials.

En BGP, les polítiques d'encaminament s'implementen a través dels atributs.

Pregunta 2. Explica como escala la tabla de encaminamiento BGP en función de la cantidad de AS's a los que está conectado un AS.

Un AS pot estar connectar a N AS's, per tant, rebria un missatge UPDATE per cada possible camí per cada connexió BGP. Per evitar tenir una taula d'encaminament desproporcionada, cada AS decideix quina és la millor opció, en funció dels atributs. Un router només anunciarà la millor ruta en els missatges UPDATE BGP. Així, es té una visió parcial d'Internet, ja que només es veu allò que altres routers decideixen que s'ha de veure. D'aquesta manera, s'escala la taula d'encaminament.

Si es refereix a escalabilitat interna, BGP compta amb reflectors de ruta i confederacions.

Pregunta 3. ¿Pará que sirve definir una dirección de loopback en un router? ¿Qué tipo de dirección es?

En una connexió IBGP si s'utilitza la direcció de la interfície d'un router i aquest enllaç cau, ja no serà accessible. En canvi, si s'utilitza una direcció de loopback, que fa referència al router, i l'enllaç cau, es podrà trobar una ruta alternativa (en cas de que físicament n'hi hagi i que hagi estat prèviament anunciada amb protocol intern com OSPF).

Pregunta 4. ¿Cómo resuelve BGP el problema de los bucles?

Quan es retransmeten els missatges de update pels routers, si en un AS-PATH vector hi conté el del propi sistema autònom que el rep, ja no el reenvia i el descarta. Així evitem la creació de bucles en les connexions eBGP. Els routers iBGP, en canvi, no anuncien les rutes apreses per iBGP a altres routers iBGP per evitar bucles interns.

Pregunta 5. ¿Qué diferencia hay entre IBGP e EBGP?

iBGP fa referència a aquelles connexions establertes entre routers d'un mateix sistema autònom. eBGP a aquelles connexions establertes entre routers de diferents sistemes autònoms.

Pregunta 6. ¿Qué diferencia hay entre las redes que anuncia OSPF y las que anuncia BGP (e.g. con el comando network)?

OSPF només anuncia xarxes internes al sistema autònom dins el mateix sistema autònom (cada router farà network de totes les xarxes que té connectades i si alguna és un enllaç extern s'haurà d'usar passive interface). BGP és un protocol d'encaminament a nivell de AS i estableix connexions eBGP entre routers de diferents AS i iBGP entre routers interns. En aquest cas, però, la comanda network no s'utilitza per a sessions iBGP per routes apreses

internament.

Pregunta 7. Explica la diferencia entre un atributo BGP conocido ("well-known") y otro opcional. Idem si el atributo es mandatorio y discrecional. Menciona algún atributo que tenga la característica de ser conocido y discrecional, otro que sea conocido y mandatorio y otro que sea opcional y transitivo.

Els atributs well-known han de ser suportats per totes les implementacions de BGP, mentre que els opcionals no. Si un atribut és mandatori, cal que sigui sempre enviat a través d'un missatge UPDATE, mentre que els discrecionals no s'envien.

Conegut i discrecional: LOCAL-PREFERENCE i ATOMIC AGGREGATE

Conegut i mandatory: AS-PATH, NEXT-HOP i ORIGIN

Opcional i transitiu: AGREGATOR i COMMUNITY

Pregunta 8. ¿Qué significa que en una tabla BGP aparezca el atributo ORIGEN como incompleto? ¿Qué acción ha ejecutado el administrador del sistema para que aparezca como incompleto? ¿Qué efectos tiene dicha acción?

Significa que l'origen és desconegut. Els efectes de que aquest atribut aparegui com a incomplet és que en la presa de decisió serà la última opció (després de IGP i EGP). La causa per la qual l'atribut és incomplet pot ser que aquella ruta s'ha distribuït a BGP a través d'un protocol intern, com RIP, OSPF o IS-IS.

Pregunta 9. ¿Qué relación hay entre los atributos ATOMIC AGGREGATE y AGGREGATOR?

L'Agreggator és opcional i transitiu. S'envia un UPDATE indicant una subxarxa que pot ser agregada. El router que agrega pot indicar a l'AS-PATH vector la partició de la subxarxa que ha agregat (opció AS-SET). No té influència en la selecció de camí.

L'Atomic Aggregate és well-known i discrecional. Alerta als routers BGP d'una ruta que part de la informació ha estat perduda a causa d'un procés d'agregació i que, per tant, podria no ser el millor camí cap a la destinació. Si quan s'ha agregat no s'ha activat AS-SET pot passar que el vector AS-PATH perdi informació de les agregacions prèvies. Atomic Aggregate és mandatori que es tregui actiu.

Pregunta 10. Qué diferencia hay entre una política BGP inbound y una outbound. Qué atributo BGP te permite generar una política outbound?

La política BGP inbound serveix per triar quin és el link d'entrada mentre que la outbound per triar la de sortida. L'atribut Local Preference permet generar una política outbound.

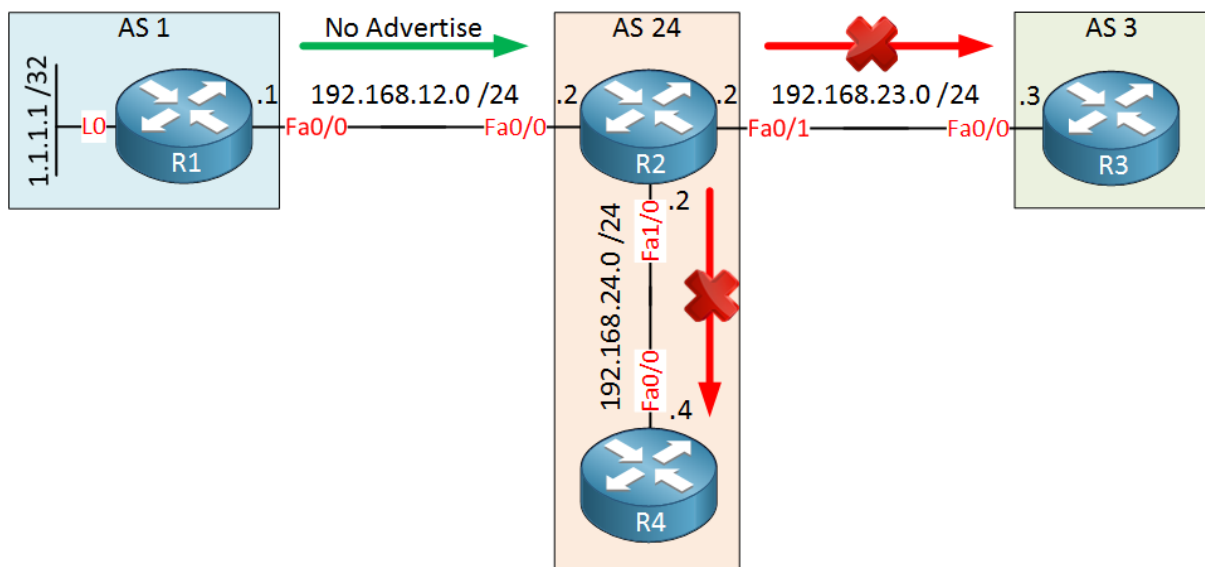
Pregunta 11. ¿Qué es una política de "AS-path-prepend"? Explica mediante un ejemplo sencillo como un ISP puede usar esta política. ¿Qué atributo BGP permite definir a un ISP una política de tráfico de tipo "outbound"? Explica mediante un ejemplo sencillo como un ISP puede usar esta política.

Una política AS-path-prepend implica inflar el nombre de salts d'un camí per evitar que sigui triat com a millor camí d'altres as AS. Un ISP X que té una sessió BGP amb un altre ISP Y vol evitar en la mesura del possible que aquest últim passi a través d'ell. Per aconseguir-ho X envia l'atribut AS-PATH a Y amb més AS dels que realment hi ha. D'aquesta manera, si hi ha una ruta alternativa, Y marcarà com a preferible una altra opció.

L'atribut LOCAL-PREF permet generar una política outbound al ISP. L'atribut LOCAL-PREF és el primer atribut que es mira l'algoritme que tria la millor ruta, si hi ha dues rutes i es puja el LOCAL-PREF a una, l'algoritme triarà la ruta amb el LOCAL-PREF més alt.

Pregunta 12. Explica la diferencia entre una comunidad "NO-EXPORT" y una comunidad "NO-ADVERTISE". Pon un ejemplo de uso de cada una de ellas.

La comunitat "NO-EXPORT" obliga a que totes les rutes que el rebin no el retransmetin fora del sistema autònom. La comunitat "NO-ADVERTISE" obliga a que totes les rutes que el rebin no el retransmetin a altres veïnes BGP dintre del sistema autònom.



En aquest exemple, R1 anuncia la xarxa a R2 que se l'emmagatzema però no la retransmet perquè és NO-ADVERTISE.

Si en canvi, la comunitat entre R1 i R2 fos NO-EXPORT, R2 emmagatzemaria la xarxa anunciada per R1 i només la retransmetria a R4 perquè està dintre el seu mateix sistema autònom però no a R3 perquè l'AS és diferent del de R2.

Pregunta 13. ¿Qué diferencia hay entre asignar un "route-map" con el comando neighbor en modo "in" o en modo "out"? ¿Qué efectos tienen ambas acciones sobre la tabla BGP?

En mode in és per missatges BGP que entren al router. El seu efecte és que les condicions-accions del route-map s'inclouran en la taula BGP.

El mode out és per missatges que surten del router. El seu efecte és que les condicions-accions del route-map s'inclouran en el missatge BGP que surt. Es modificaran les taules BGP destinació.

Pregunta 14. Justifica porqué los routers BGP tienen que estar i-BGP totalmente mallados y explica la diferencia entre el funcionamiento i-BGP y e-BGP respecto al anuncio de rutas.

Per evitar bucles, els routers iBGP entre ells no s'exporten les rutes internes. Per això, cal fer una malla completa. Per contra, en les connexions eBGP sí que s'hauran d'anunciar totes les rutes que es vulguin fer públiques.

Pregunta 15. Explica que es multi-homing y explica como se puede implementar una línea

de back-up con un ISP.

El multi-homing és quan un client té més d'un connexió amb un o diversos ISP. Amb l'atribut MED es pot forçar el link d'entrada i amb LOCAL-PREFERENCE el de sortida. Així s'utilitzarà només un enllaç i si aquest cau, utilitzarà el de backup.

Pregunta 16. Explica qué significa que el encaminamiento externo e interno estén sincronizados.

Quan un router d'un AS aprèn una ruta externa i l'exporta internament i tots els routers del AS coneixen aquesta ruta significarà que els dos encaminaments estan sincronitzats. Mentre hi hagi routers interns que no coneixen la ruta, no ho estarà.

Pregunta 17. Asume que tienes un ISP con 100 routers BGP. Indica cuantas sesiones I-BGP necesita para funcionar correctamente. Indica que tecnicas hay para reducir el número de sesiones I-BGP y explica brevemente el funcionamiento de una de ellas. Pon un ejemplo de las técnicas que has explicado enseñando la reducción de sesiones BGP a los 100 routers.

Hi haurà $N*(N-1)/2$ connexions. $(100*99)/2 = 4950$ sessions. Per reduir el nombre de sessions iBGP hi ha els reflectors de ruta i les confederacions.

Les confederacions són subsistemes autònoms dins un mateix sistema autònom i utilitza números de AS privats. Aquests subsistemes es tracten entre ells com si fossin sistemes diferents, establint connexions eBGP entre ells i fent xarxa mallada completa a dins.

Els reflectors de ruta consisteix en la creació de clusters dins un AS. Existeix la figura del reflector de ruta, qui manté sessions iBGP amb els seus clients. La malla completa només es fa entre els reflectors de ruta, mentre que els clients no és necessari.

Per exemple, si als 100 routers posem 5 reflectors de ruta, hi haurà 20 clients per reflector, per tant:

$$20+20+20+20+20+(20*(20-1)/2) = 290 \text{ sessions iBGP}$$

Si fos el cas de posar 4 subsistemes autònoms, tindríem 25 routers per confederació:

$$4 * (25*(25-1)/2) = 1200 \text{ sessions iBGP.}$$

Pregunta 18. Asume que tienes un ISP con 50 routers BGP. Para que funcionen correctamente necesitas una red totalmente mallada i-BGP implicando $50*49/2=1225$ sesiones i-BGP. Explica una técnica que sea capaz de reducir el número de sesiones i-BGP y pon un ejemplo con los 50 routers BGP en que se muestre dicha reducción.

Tècniques: Reflectors de ruta o confederacions (explicades pregunta anterior).

Podem fer servir la tècnica abans explicada de les confederacions. Fem 10 confederacions, cadascuna amb 5 routers. Cada confederació tindria $5*4/2 = 10$ connexions I-BGP i en total 100 connexions I-BGP i les connexions mínimes E-BGP serien 9 connexions. En total, tindriem 109 connexions.

Pregunta 19. Explica el funcionamiento de los reflectores de rutas en BGP.

Els reflectors de ruta consisteix en la creació de clusters dins un AS. Existeix la figura del reflector de ruta, qui manté sessions iBGP amb els seus clients. La malla completa només es fa entre els reflectors de ruta, mentre que els clients no és necessari. S'aconsegueix

modificant la norma split horizon.

Pregunta 20. Explica el funcionamiento de las confederaciones en BGP.

Les confederacions són subsistemes autònoms dins un mateix sistema autònom i utilitza números de AS privats. Aquests subsistemes es tracten entre ells com si fossin sistemes diferents, establint connexions eBGP entre ells i fent xarxa mallada completa a dins.

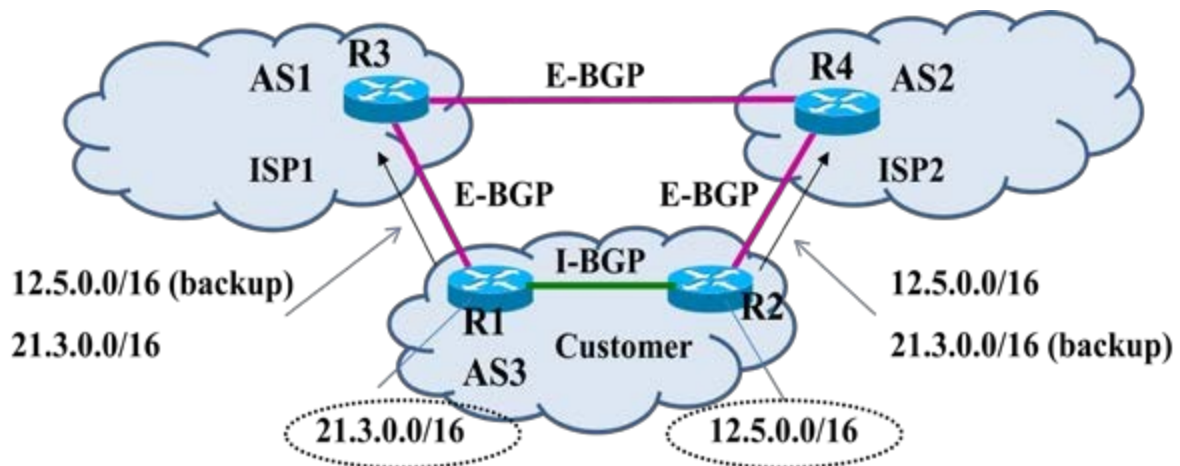
Pregunta 21. Explica los conceptos de escalabilidad, sincronización y convergencia en BGP y como se solucionan cada uno de ellos.

Escalabilitat: Capacitat d'adaptar-se al creixement de la xarxa. En aquest sentit, BGP implementa confederacions i reflectors de ruta.

Sincronització: Es diu que una xarxa BGP està sincronitzada quan l'intern i l'extern tenen mapejades les mateixes rutes. És a dir, que allò que s'ha après per eBGP també s'hagi après als routers iBGP.

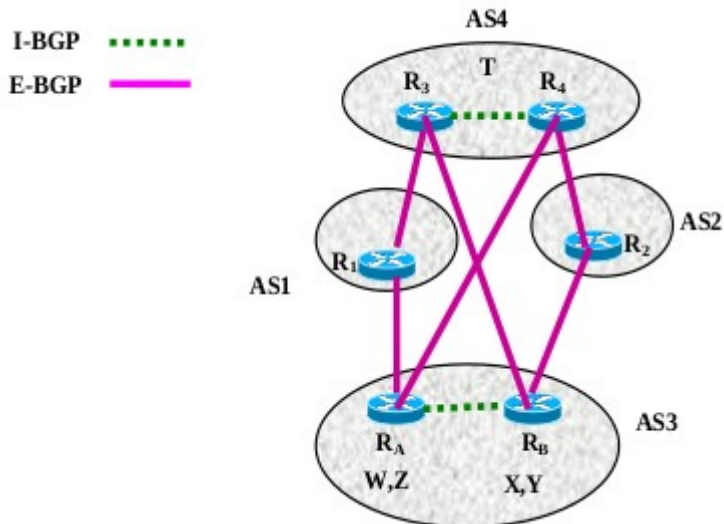
Convergència: O temps de convergència és el temps que tarda la xarxa a convergir, és a dir, a que tots els nodes tinguin la mateixa informació. Es soluciona retardant el temps en el qual L2 alerta a L3 que ha caigut un enllaç (per solucionar flapping). Reduir també la freqüència dels missatges UPDATE. També s'utilitza el dampening.

Pregunta 22. Explica cómo puede el AS3 forzar que la línea R3-R1 es backup para la red 12.5.0.0/16 y principal para la 21.3.0.0/16 y la línea R4-R2 es backup para la red 21.3.0.0/16 y principal para la 12.5.0.0/16.



Caldrà que R1 pugi el Local-Preference per a la xarxa 21.3.0.0/16 i que R2 la pugi per a 12.5.0.0/16.

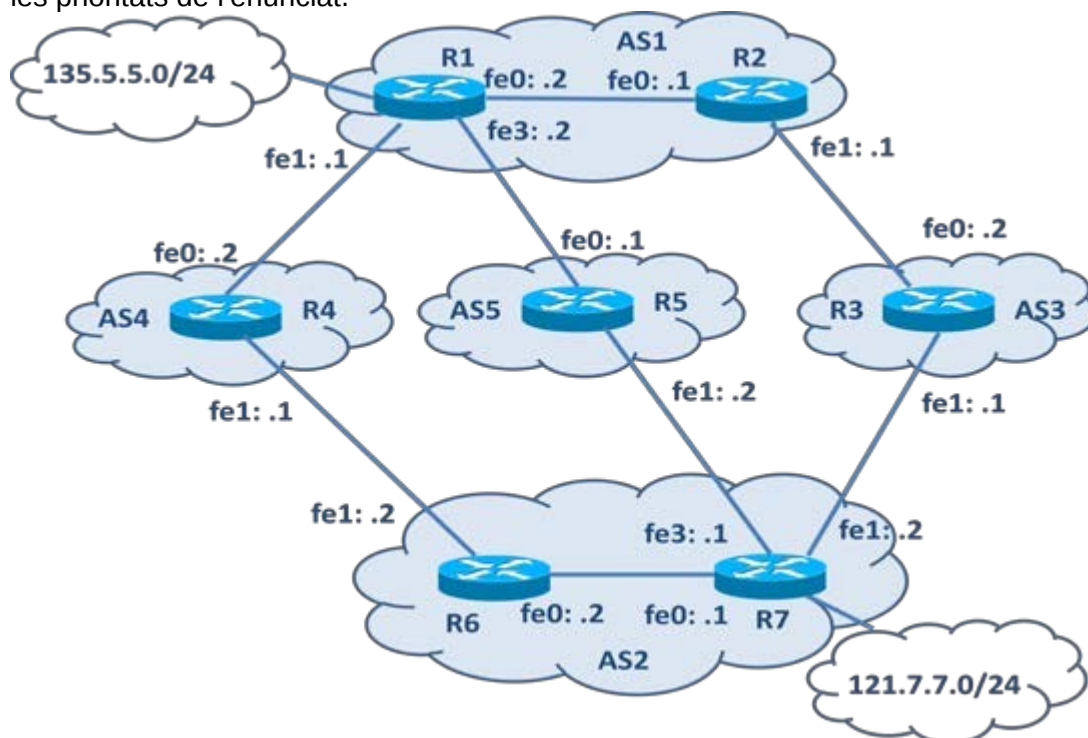
Pregunta 23. Asume que tienes la siguiente red. Explica como AS3 puede definir una comunidad para que el tráfico que va desde AS4 hacia las redes X,Y en AS3, vaya preferentemente vía R3-RB en vez de usar otras rutas y que el tráfico que va desde AS4 hacia las redes W,Z en AS3, vaya preferentemente vía R4-RA.



Cal que AS3 crei una comunitat i que AS4 la tracti. En el tractament de la comunitat en AS4, si fa match amb W,Z caldrà pujar Local-Preference a R4; si fa match amb X,Y caldrà pujar Local-Preference a R3.

Pregunta 24. Asume que tienes la arquitectura de la figura. Explica como AS2 puede definir una comunidad para que el tráfico que va desde la red 135.5.5.0/24 en el router R1 vaya a la red 121.7.7.0/24 conectada al router R7 preferentemente vía R1-R5 como primera opción, vía R1-R2 como segunda opción y finalmente vía R1-R4 como tercera opción.

Cal que AS2 faci una comunitat per a la xarxa 127.7.7.0/24. Quan faci match a R1 amb aquesta xarxa com a destí caldrà que estableixi tres Local-Preference diferents en funció de les prioritats de l'enunciat.



Pregunta 25. Asume que tienes un ISP con 100 routers BGP. Para que funcionen

correctament necessites una red totalment mallada i-BGP. Obten el número total de sessions iBGP necessaries per que funcione correctament el AS. Definim ara una configuració amb 5 confederacions: en les 3 primeres confederacions se configuren reflectors de routers (4 Reflectors amb 4 clients cada un), mentre que en les 2 últimes confederacions no hi ha reflectors. Obten el nombre total de sessions i-BGP de cada confederació i el total de la nova configuració en el AS.

Per a 100 routers, tenim $100 \cdot 99 / 2 = 4950$ connexions i-BGP.

Per cada confederació amb reflectors:

$N=20$

$N_R=4$

$N_{Ri}=4$

Sessions iBGP = $4+4+4+4+(4 \cdot (4-1)/2) = 22$ connexions iBGP per confederació.

Per cada confederació sense reflectors:

Sessions iBGP = $20 \cdot (20-1)/2 = 190$

Session iBGP totals = $22 \cdot 3 + 190 \cdot 2 = 446$ sessions.

Pregunta 26. Calcula el throughput mitjà i el "oversubscription ratio" d'un commutador amb 8 enllaços de 10 Gb/s al nivell d'agregació i 192 ports de 1Gb/s de capacitat al nivell d'accés. Si els 192 servidors del nivell d'accés ocupen un 55% de l'enllaç, ¿Està bé dissenyada la red (justifica la teua resposta)? Si la resposta és no, indica com hauria de ser el commutador per suportar els 192 servidors del nivell d'accés.

???

Pregunta 27. Explica què és el "flapping". ¿Per a què serveixen les tècniques de "slow-down" en BGP? Explica el "exponential back-off".

Flapping: Un enllaç canvia constantment d'un estat a un altre, provocant missatges d'update i provocant una convergència baixa, bucles i fallades de xarxa i meltdown.

Slow-down: Són tècniques que consisteixen en reduir la freqüència a la qual els missatges d'update són enviats a altres routers BGP.

Exponential back-off: És un algorisme que utilitza el feedback per decrementar la ràtio d'algun procés. En aquest cas, per endarrerir el processament de missatges update.

Pregunta 28. Explica què és el dampening en BGP i per a què serveix. Explica el mecanisme de dampening i la relació entre els temporitzadors half-time, max-suppress-limit i el valor màxim del suppress-limit (dampening threshold).

El dampening consisteix en que cada cop que succeeix un esdeveniment, s'incrementa un comptador de penalització. Si després d'un temps sense que cap esdeveniment succeeixi, el comptador es redueix. Si el comptador assoleix un "suppress" límit, l'esdeveniment entra en un DAMPENED estat, on l'enllaç i la ruta passen a un estat baix. Només es tornen a posar en funcionament, quan el comptador assoleix el "reuse" límit per baix. Serveix per endarrerir l'enviament de missatges d'UPDATE.

El half-time és el temps en el qual si no hi ha esdeveniments, el comptador es decrementa, el max-suppress-limit és el temps màxim en que un router pot estar down i el dampening threshold és el límit en el qual el router passa a l'estat de DAMPENED. Aquestes tres variables es

relacionen perquè s'han d'anar regulant per a què el router no estigui massa temps fora de servei.

Pregunta 29. ¿Qué es y que implicaciones tiene el max-penalty en dampening? Si tienes un penalti = 1000, un reuse-limit = 2000, un half-life = 15 minutos, y un max-supress-limit = 60 minutos, ¿Cuál es el valor máximo del supress-limit (dampening theshold) que puedes configurar?

El max-penalty s'utilitza per assegurar-se que no s'han introduït valors de dampening que no funcionaran.

$\text{max-penalty} = \text{reuse-limit} * 2^{(\text{max-suppress-time}/\text{half-life})}$

$\text{max-penalty} = 2000 * 2^{(60/15)} = 32000$