DE-CIX Academy: Best Path Selection

Links and Explanations

Where networks meet

Notice of Liability

Despite careful checking of content, we accept no liability for the content of external links.

Content on linked sites is exclusively the responsibility of the respective website operator.

Links visited during the webinar

- → Definition of terms (all from RFC4271):
  - *Next Hop* is defined in Section 5.1.3
  - *AS Path* is defined in Section 5.1.2
  - *Local Preference:* Section 5.1.5
  - Origin: Section 5.1.1
  - *Multi Exit Discriminator (MED):* Section 5.1.4
- → Best Path Selection process: Section 9.1
- → BGP Route Selection Algorithm by vendor:
  - Cisco
  - Juniper
  - Mikrotik
  - Nokia
  - BIRD
  - Quagga

DE-CIX Academy: Best Path Selection

Links and Explanations

Where networks meet

BGP Routing Algorithm

Bolded items were covered in this webinar.

1		NextHop reachable?	Continue if "yes"
2		Local Preference	higher wins
3		AS Path	shorter wins
4		**Origin Type**	**IGP over EGP over Incomplete**
5		**MED**	**lower wins**
6		**eBGP, iBGP**	**eBGP wins**
7		**Exit**	**nearest wins**
8		**Age of route**	**older wins**
9		**Router ID**	**lower wins**
10		**Neighbor IP**	**lower wins**

Local Preference is...

- → a 32bit integer value (0-4294967295)
- → Propagated via iBGP inside an Autonomous System
- → Usually set using rules when receiving prefixes
  - According to your routing policy
- → Typical values
  - 10000 (high value) for customer prefixes
  - 1000 (medium value) for prefixes received via peering
  - 100 (low value) for prefixes received via upstream
- → Rules to adjust local preference can be as complex as your router software allows it to be.

AS Path is...

- → an ordered list of AS numbers...
- → ...with the originator AS at the rightmost side
- → automatically built when prefixes are sent via eBGP
- → length of the path is used for selection (shorter wins)

DE-CIX Academy: Best Path Selection

Links and Explanations

Where networks meet

Origin Type is...

- → a historic, but mandatory attribute
- → set by originator AS and forwarded unchanged
- → can have the values (in order of preference):
  - IGP - prefix was originated via a network statement
  - EGP - prefix was learned from Exterior Gateway Protocol (RFC904, historic)
  - incomplete - prefix was learned by another protocol

Multi Exit Discriminator (MED) is...

- → a 32Bit value, lower wins
- → optional, if it is not there it's either treated as zero (best) or as 2^32-1 (worst)
- → non-transitive (set by an eBGP speaker and only sent to the next-hop AS)
- → usually set using rules when sending prefixes (according to the sender's routing policy)
- → only compared between eBGP speakers if next-hop AS is the same

Router ID is...

- → also called **BGP Identifier**
- → a 4 byte, unsigned integer (mostly it's the IPv4 loopback address of a router)
- → unique within one AS
- → set at startup and stays unchanged
- → the same for all BGP sessions

Neighbor IP is...

- → the last tie-breaker in the BGP Best Path Selection
- → the IP address of the eBGP speaker a prefix was learned from

DE CIX

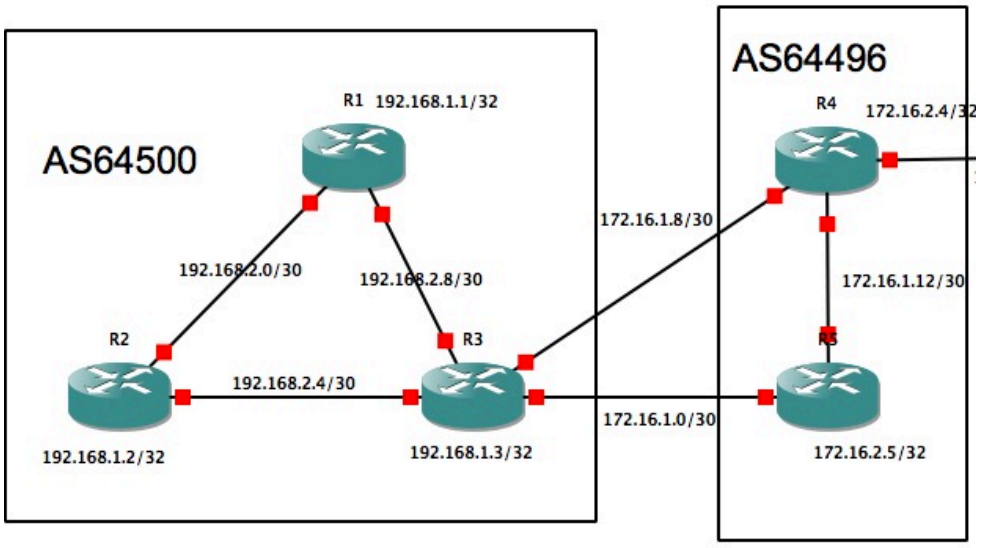
DE-CIX Academy: Best Path Selection

Links and Examples

Where networks meet

Router Configs (Cisco IOS)

Example Network



Experiment 1: Set outgoing MED on R4 and R5

On R4:

```
conf t
route-map customer-out permit 100
  set metric 0
end
clear ip bgp 64500 soft out
```

On R5:

```
conf t
route-map customer-out permit 100
  set metric 1000
end
clear ip bgp 64500 soft out
```

Experiment 2: Age of route

- Set metric on R4 and R5 to the same value
- on R3 shut down interface Gig0/0 or Gig 3/0 and see how best prefix changes

DE-CIX Academy: Best Path Selection

Links and Examples

Where networks meet

AS64500 Router R1

```
hostname R1
!
interface Loopback0
 ip address 192.168.1.1 255.255.255.255
!
interface GigabitEthernet0/0
 ip address 192.168.2.2 255.255.255.252
 description to R2
!
interface GigabitEthernet2/0
 ip address 192.168.2.9 255.255.255.252
 description to R3
!
router ospf 64500
 redistribute connected subnets route-map internal-only
 network 192.168.2.0 0.0.0.3 area 0
 network 192.168.2.8 0.0.0.3 area 0
!
router bgp 64500
 bgp log-neighbor-changes
 neighbor internal peer-group
 neighbor internal remote-as 64500
 neighbor internal update-source Loopback0
 neighbor internal next-hop-self
 neighbor internal send-community both
 neighbor 192.168.1.2 peer-group internal
 neighbor 192.168.1.3 peer-group internal
!
ip prefix-list internal seq 5 permit 192.168.0.0/16 le 32
route-map internal-only permit 10
 match ip address prefix-list internal
!
end
```

DE-CIX Academy: Best Path Selection

Links and Examples

Where networks meet

AS64500 Router R2

```
hostname R2
!
interface Loopback0
 ip address 192.168.1.2 255.255.255.255
!
interface GigabitEthernet0/0
 ip address 192.168.2.1 255.255.255.252
 description to R1
!
interface GigabitEthernet1/0
 ip address 192.168.2.5 255.255.255.252
 description to R3
!
router ospf 64500
 redistribute connected subnets route-map internal-only
 network 192.168.2.0 0.0.0.3 area 0
 network 192.168.2.4 0.0.0.3 area 0
!
router bgp 64500
 bgp log-neighbor-changes
 neighbor internal peer-group
 neighbor internal remote-as 64500
 neighbor internal update-source Loopback0
 neighbor internal next-hop-self
 neighbor internal send-community both
 neighbor 192.168.1.1 peer-group internal
 neighbor 192.168.1.3 peer-group internal
!
ip prefix-list internal seq 5 permit 192.168.0.0/16 le 32
route-map internal-only permit 10
 match ip address prefix-list internal
!
end
```

DE-CIX Academy: Best Path Selection

Links and Examples

Where networks meet

AS64500 Router R3

```
hostname R3
!
interface Loopback0
 ip address 192.168.1.3 255.255.255.255
!
interface GigabitEthernet0/0
 description to AS64496 R5
 ip address 172.16.1.2 255.255.255.252
!
interface GigabitEthernet1/0
 description to R2
 ip address 192.168.2.6 255.255.255.252
!
interface GigabitEthernet2/0
 description to R1
 ip address 192.168.2.10 255.255.255.252
!
interface GigabitEthernet3/0
 description to AS64496 R4
 ip address 172.16.1.10 255.255.255.252
!
router ospf 64500
 redistribute connected subnets route-map internal-only
 network 192.168.2.4 0.0.0.3 area 0
 network 192.168.2.8 0.0.0.3 area 0
!
router bgp 64500
 bgp log-neighbor-changes
 neighbor internal peer-group
 neighbor internal remote-as 64500
 neighbor internal update-source Loopback0
 neighbor internal next-hop-self
 neighbor internal send-community both
 neighbor upstream peer-group
 neighbor upstream send-community both
 neighbor upstream soft-reconfiguration inbound
 neighbor upstream route-map upstream-in in
 neighbor upstream route-map upstream-out out
 neighbor 172.16.1.1 remote-as 64496
 neighbor 172.16.1.1 peer-group upstream
 neighbor 172.16.1.9 remote-as 64496
 neighbor 172.16.1.9 peer-group upstream
 neighbor 192.168.1.1 peer-group internal
 neighbor 192.168.1.2 peer-group internal
!
ip prefix-list internal permit 192.168.0.0/16 le 32
!
route-map upstream-out permit 100
!
route-map upstream-in permit 100
 set local-preference 100
!
route-map internal-only permit 10
 match ip address prefix-list internal
!
end
```

DE CIX

DE-CIX Academy: Best Path Selection

Links and Examples

Where networks meet

AS64496 Router R4

```
hostname R4
!
interface Loopback0
 ip address 172.16.2.4 255.255.255.255
!
interface GigabitEthernet1/0
 description to AS64500 R3
 ip address 172.16.1.9 255.255.255.252
!
interface GigabitEthernet2/0
 description to R5
 ip address 172.16.1.13 255.255.255.252
!
router ospf 64496
 redistribute connected subnets
 network 172.16.1.12 0.0.0.3 area 0
!
router bgp 64496
 network 172.16.0.0
 neighbor internal peer-group
 neighbor internal remote-as 64496
 neighbor internal update-source Loopback0
 neighbor internal next-hop-self
 neighbor internal send-community both
 neighbor customer peer-group
 neighbor customer send-community both
 neighbor customer soft-reconfiguration inbound
 neighbor customer route-map customer-in in
 neighbor customer route-map customer-out out
 neighbor 172.16.1.10 remote-as 64500
 neighbor 172.16.1.10 peer-group customer
 neighbor 172.16.2.5 peer-group internal
!
ip route 172.16.0.0 255.255.0.0 Null0
!
route-map customer-in permit 100
 set local-preference 10000
!
route-map customer-out permit 100
!
end
```

DE-CIX Academy: Best Path Selection

Links and Examples

Where networks meet

AS64496 Router R5

```
hostname R5
!
interface Loopback0
 ip address 172.16.2.5 255.255.255.255
!
interface GigabitEthernet0/0
 description to AS64500 R3
 ip address 172.16.1.1 255.255.255.252
!
interface GigabitEthernet2/0
 description to R5
 ip address 172.16.1.14 255.255.255.252
!
router ospf 64496
 redistribute connected subnets
 network 172.16.1.12 0.0.0.3 area 0
!
router bgp 64496
network 172.16.0.0
 neighbor internal peer-group
 neighbor internal remote-as 64496
 neighbor internal update-source Loopback0
 neighbor internal next-hop-self
 neighbor internal send-community both
 neighbor customer peer-group
 neighbor customer send-community both
 neighbor customer soft-reconfiguration inbound
 neighbor customer route-map customer-in in
 neighbor customer route-map customer-out out
 neighbor 172.16.1.2 remote-as 64500
 neighbor 172.16.1.2 peer-group customer
 neighbor 172.16.2.4 peer-group internal
!
ip route 172.16.0.0 255.255.0.0 Null0
route-map customer-in permit 100
 set local-preference 10000
!
route-map customer-out permit 100
!
end
```