

BGP – Lab

Traffic Engineering

Author: Thein Myint Khine

Version: 0.0

Last Update: September 11, 2024

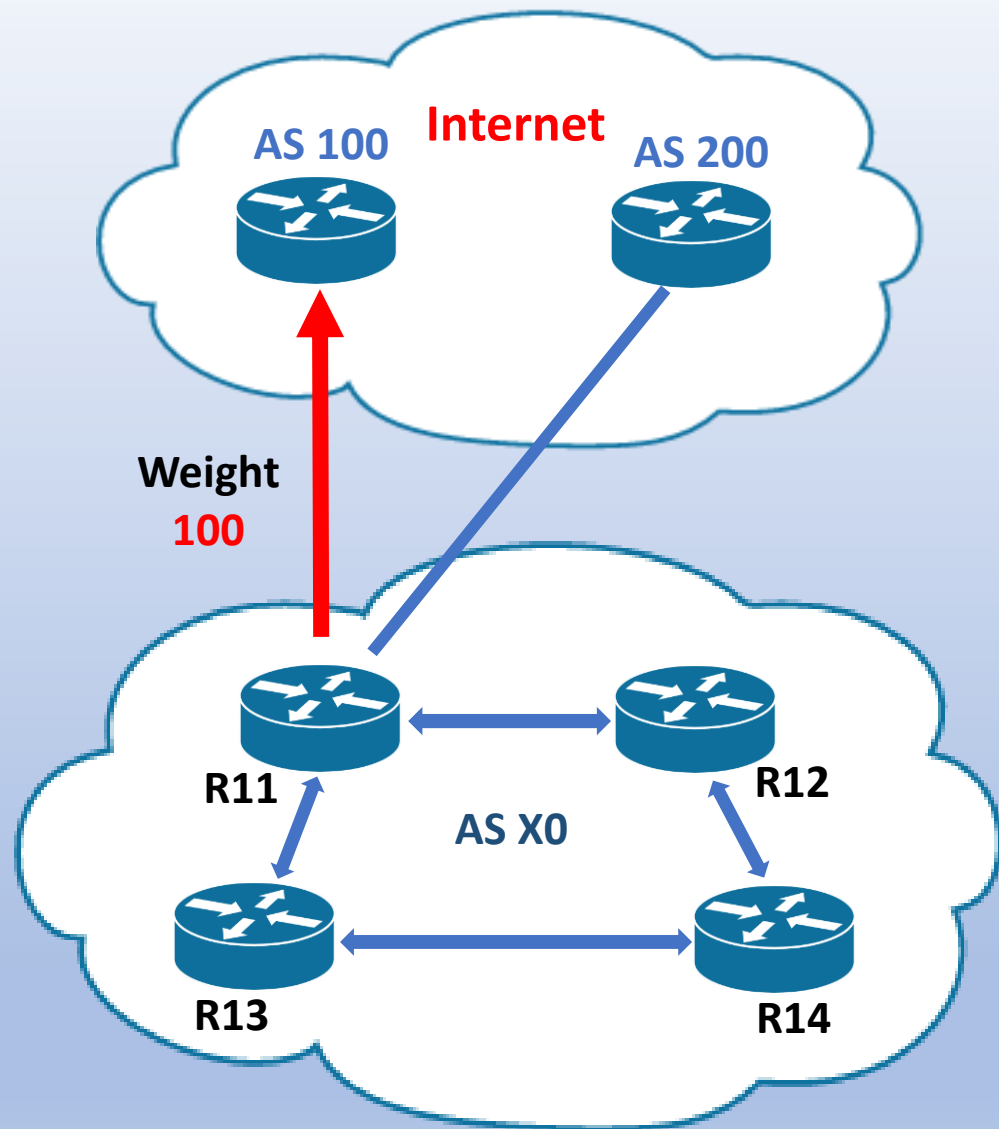


mmnog

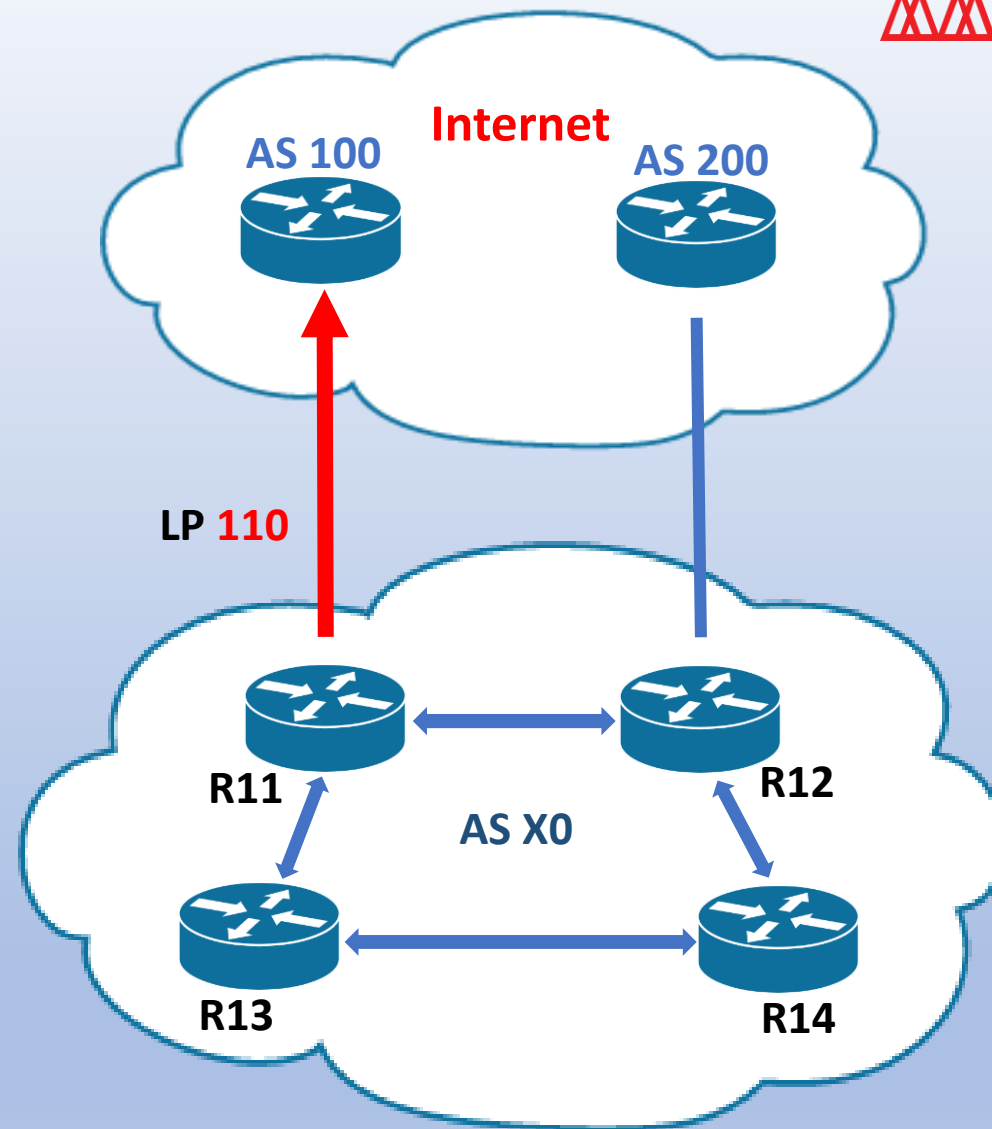
MYANMAR NETWORK OPERATORS GROUP

BGP Path Selection Algorithm for IOS Part One

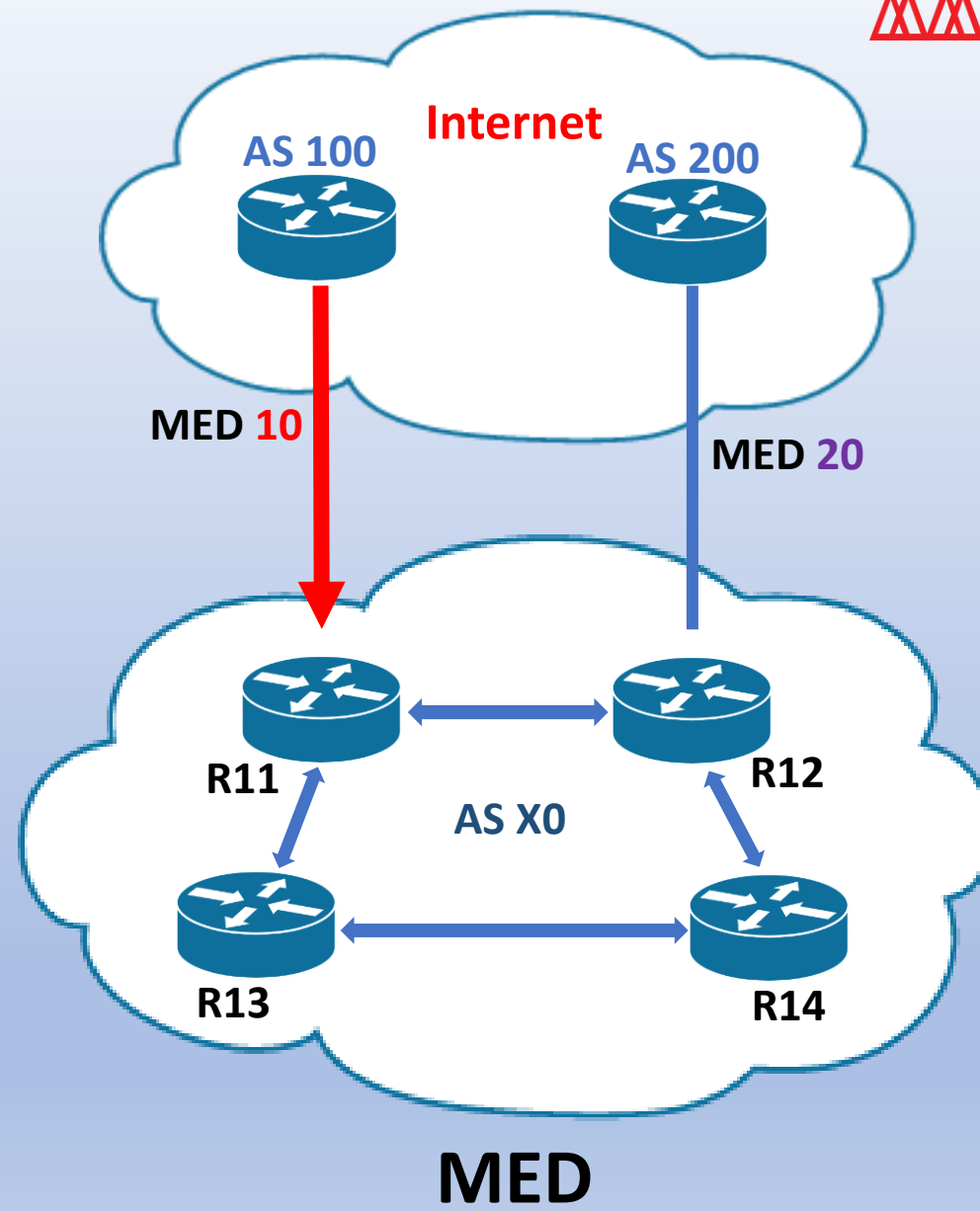
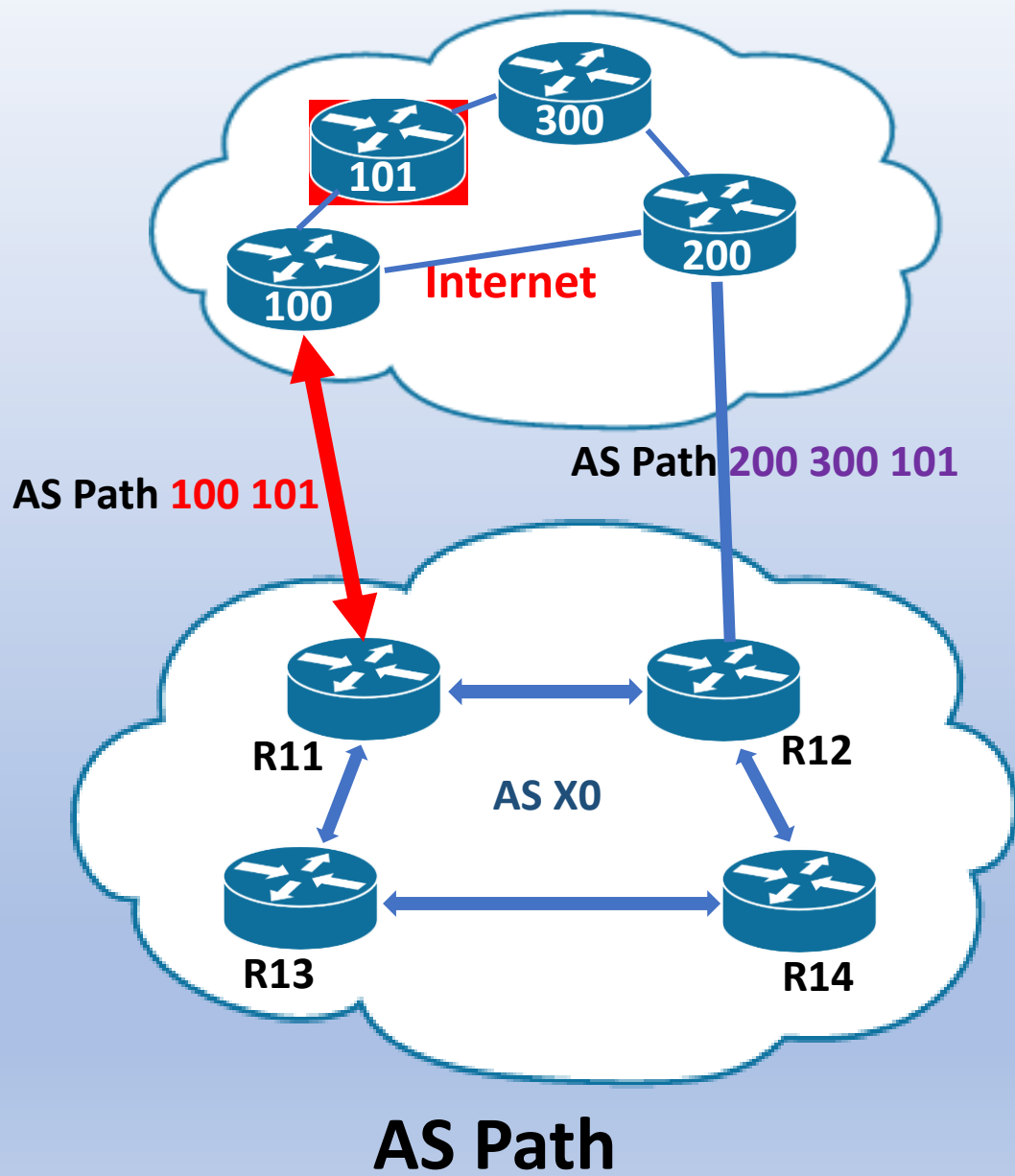
- Do not consider path if no route to next hop
- Do not consider iBGP path if not synchronized (Cisco IOS)
- Highest weight (local to router)
- Highest local preference (global within AS)
- Prefer locally originated route
- Shortest AS path
- Lowest origin code
- Lowest Multi-Exit Discriminator (MED)
- Prefer eBGP path over iBGP path
- Path with lowest IGP metric to next-hop
- Lowest router-id
- Shortest cluster-list
- Lowest neighbor address



Weight



Local Preference



BGP-Attributes-Summary

	Traffic Direction	Prefer	Default	Effective Region
Local Preference	Outbound	Higher Value	100	Within AS
MED	Inbound	Lower Value	0	Other AS
Weight	Outbound	Higher Value	0	Local to Router
AS-Path	In/Out	Lower Hop Count		Multiple AS

BGP Filters

Three ways for BGP filtering

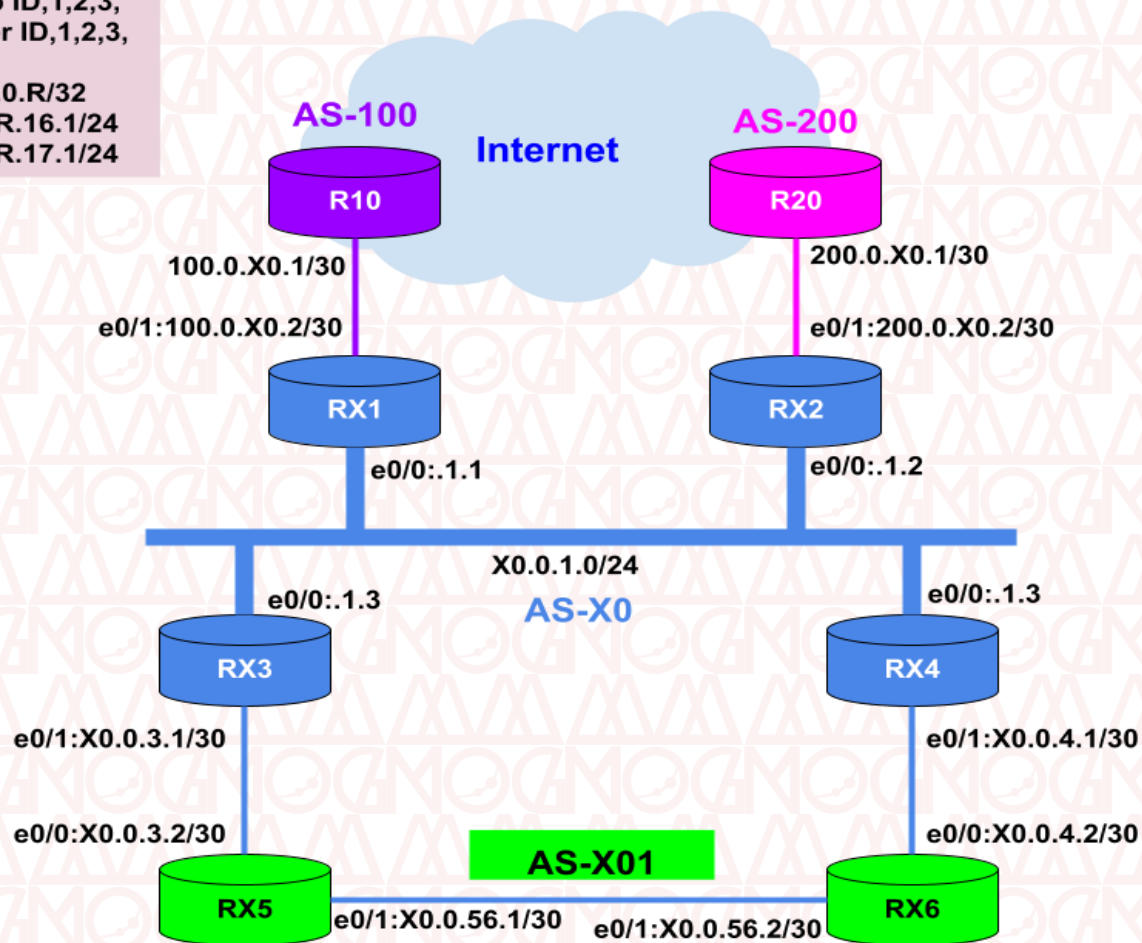
- Prefix-list - Prefixes
- Filter-list - AS-Path
- Route-map - Enhanced

Community – tagging mechanism to mark BGP prefixes

LAB-BGP-Attributes

X = Group ID,1,2,3,
R = Router ID,1,2,3,

Lo0: X0.0.0.R/32
Lo16: X0.R.16.1/24
Lo17: X0.R.17.1/24



Configure to meet the following requirements.

1. Check configuration and ping test link IP addresses. Check also OSPF routes, BGP sessions and BGP routes.
2. Blue ISP will use AS100 link as primary for incoming traffic for Client prefixes of Rx1 and Rx3 while the link is still a backup link for Client prefixes of Rx2 & Rx4. (AS prepend)

! @Rx2, Prefix list defines for prefixes of R1 and R3.

```
ip prefix-list PRF-13 seq 5 permit 10.1.16.0/23 le 24
ip prefix-list PRF-13 seq 10 permit 10.3.16.0/23 le 24
```

! @Rx2, create route-map for bgp

```
route-map RM-UPLINK-OUT permit 10
  match ip address prefix-list PRF-13
  set as-path prepend 10 10
route-map RM-UPLINK-OUT permit 65535
```

! bind route-map at BGP session

```
router bgp 10
  neighbor 200.0.10.1 route-map RM-UPLINK-OUT out
```


! for immediate effective, soft clear bgp session

```
clear ip bgp 200.0.10.1 soft out
```

! check routing table from R10 & R20 for effectiveness

```
R10#sh ip bgp regexp _10$
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.1.16.0/24	100.0.10.2	0		0	10 i
*>	10.1.17.0/24	100.0.10.2	0		0	10 i
*	10.2.16.0/24	100.0.10.2			0	10 10 10 i
*>		12.0.1.2			0	200 10 i
*	10.2.17.0/24	100.0.10.2			0	10 10 10 i
*>		12.0.1.2			0	200 10 i
*>	10.3.16.0/24	100.0.10.2			0	10 i
*>	10.3.17.0/24	100.0.10.2			0	10 i
*>	10.4.16.0/24	12.0.1.2			0	200 10 i
*		100.0.10.2			0	10 10 10 i
*>	10.4.17.0/24	12.0.1.2			0	200 10 i
*		100.0.10.2			0	10 10 10 i

```
R20#sh ip bgp regexp _10$
  Network          Next Hop          Metric LocPrf Weight Path
 *> 10.1.16.0/24    12.0.1.1          0 100 10 i
 *                200.0.10.2        0 10 10 10 i
 *> 10.1.17.0/24    12.0.1.1          0 100 10 i
 *                200.0.10.2        0 10 10 10 i
 *> 10.2.16.0/24    200.0.10.2        0 10 i
 *> 10.2.17.0/24    200.0.10.2        0 10 i
 *> 10.3.16.0/24    12.0.1.1          0 100 10 i
 *                200.0.10.2        0 10 10 10 i
 *> 10.3.17.0/24    12.0.1.1          0 100 10 i
 *                200.0.10.2        0 10 10 10 i
 *> 10.4.16.0/24    200.0.10.2        0 10 i
 *> 10.4.17.0/24    200.0.10.2        0 10 i
```

- Blue ISP will use AS200 link as primary for incoming traffic for Client prefixes of Rx2 & Rx4 while the link is still a backup link for Client prefixes of Rx1 & Rx3. (AS Prepend) <Same like 2>

4. From Blue ISP, outgoing traffic destined to prefixes of AS100 should go via Rx1; outgoing traffic destined to prefixes of AS200 should go via Rx2. (AS-Path & Local Preference)

```
! check existing routes
```

```
! define as-path @Rx1
```

```
R11(config)#ip as-path access-list 1 permit ^100$
```

```
! create route-map for BGP
```

```
R11(config)#route-map RM-AS100-IN permit 10
```

```
match as-path 1
```

```
set local-preference 110
```

```
route-map RM-AS100-IN permit 65535
```

```
! bind to bgp sessions
```

```
router bgp 10
```

```
neighbor 100.0.10.1 route-map RM-AS100-IN in
```

```
! for immediate effective, soft clear bgp session
```

```
clear ip bgp 100.0.10.1 soft in
```

```
! check result
```

```
R11#sh ip bgp regexp _100$
  Network          Next Hop           Metric LocPrf  Weight Path
 *> 100.1.16.0/24   100.0.10.1         0      110      0 100 i
 *> 100.1.17.0/24   100.0.10.1         0      110      0 100 i
! do the same for R12
```

5. Blue ISP will not accept Full BGP routes but default via AS-100 and partial via AS-200.

! Create prefix list for default gateway @Rx1 and bind to bgp session

```
R11(config)#ip prefix-list PRF-DGW permit 0.0.0.0/0
router bgp 10
  nei 100.0.10.1 prefix-list PRF-DGW in
```

!check routes

```
R11#sh ip bgp sum
Neighbor    V  AS  MsgRcvd  MsgSent    TblVer    InQ  OutQ  Up/Down  Sta/PfxRcd
100.0.10.1  4 100      31       24         30     0    0 00:12:18      5
```

!

Q 5 : BGP table optimize



```
! re-check routes
```

```
R11#sh ip bgp sum
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	Sta/PfxRcd
100.0.10.1	4	100	31	24	30	0	0	00:12:18	1

```
R11#sh ip bgp regexp ^100$
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 0.0.0.0	100.0.10.1		110	0	100 i

```
! check route tables or Rx2
```

```
R12#sh ip bgp regexp _200_
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 100.1.16.0/24	200.0.10.1			0 200	100 i
*> 100.1.17.0/24	200.0.10.1			0 200	100 i
*> 200.1.16.0	200.0.10.1	0		0 200	i
*> 200.1.17.0	200.0.10.1	0		0 200	i

! Partial routes

! Create as-path access list and bind to route-map

```
ip as-path access-list 1 permit ^200$
route-map RM-AS200-IN permit 10
  match ip address prefix-list PRF-DGW
  set as-path prepend 10 10
route-map RM-AS200-IN permit 15
  match as-path 1
```

! Check result

R12#sh ip bgp

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>i	0.0.0.0	10.0.0.1	0	110	0	100 i
*		200.0.10.1			0	10 10 200 i
*>	100.1.16.0/24	200.0.10.1			0	10 10 200 100 i
*>	100.1.17.0/24	200.0.10.1			0	10 10 200 100 i
*>	200.1.16.0	200.0.10.1	0		0	10 10 200 i
*>	200.1.17.0	200.0.10.1	0		0	10 10 200 i

6. Green ISP will accept default-GW only from Blue ISP.

! check existing routes

```
R15#sh ip bgp sum
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	St/PfxRc
10.0.3.1	4	10	1086	1076	60	0	0	16:09:11	13
11.0.56.2	4	101	1084	1084	60	0	0	16:09:04	15

! create prefix-list for DGW and bind to bgp session

```
R15(config)#
```

```
ip prefix-list PRF-DGW permit 0.0.0.0/0
```

```
router bgp 101
```

```
neighbor 10.0.3.1 prefix-list PRF-DGW in
```

! Check result

```
R15#sh ip bgp sum
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	St/PfxRc
10.0.3.1	4	10	1086	1076	60	0	0	16:09:11	1
11.0.56.2	4	101	1084	1084	60	0	0	16:09:04	15

! do the same config for Rx6.

7. Green ISP will use link to Rx3 as primary link while link to Rx4 shall be backup link. At Rx5 & Rx6, adjust necessary bgp parameters to meet the above requirement. (MED & LP)

! Check routes of Green from Blue

```
R13#sh ip bgp regexp _101$
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
* i	11.5.16.0/24	10.0.0.4	0	100	0	101 i
*>		10.0.3.2	0		0	101 i
* i	11.5.17.0/24	10.0.0.4	0	100	0	101 i
*>		10.0.3.2	0		0	101 i
*>	11.6.16.0/24	10.0.3.2			0	101 i
* i		10.0.0.4	0	100	0	101 i
*>	11.6.17.0/24	10.0.3.2			0	101 i
* i		10.0.0.4	0	100	0	101 i

Q 7 : multi-link to single upstream



! Create Route-map for higher metric value at Rx6 and bind to bgp

```
route-map RM-AS10-OUT permit 10
  set metric 20
router bgp 101
  neighbor 10.0.4.1 send-community
  neighbor 10.0.4.1 route-map RM-AS10-OUT out
```

! also recommend changing the lower metric value at Rx5 to Rx3.

! check result after bgp update enforce

```
R14#sh ip bgp regexp _101$
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i 11.5.16.0/24	10.0.0.3	0	100	0	101 i
*	10.0.4.2	20		0	101 i
*>i 11.5.17.0/24	10.0.0.3	0	100	0	101 i
*	10.0.4.2	20		0	101 i
*>i 11.6.16.0/24	10.0.0.3	0	100	0	101 i
*	10.0.4.2	20		0	101 i
*>i 11.6.17.0/24	10.0.0.3	0	100	0	101 i
*	10.0.4.2	20		0	101 i

Q 7 : multi-link to single upstream



! Outbound traffic adjustment

! set higher local preference @Rx5 for outbound traffic adjustment

```
R15(config)#
route-map RM-AS10-IN permit 10
  set local-preference 110
router bgp 101
  neighbor 10.0.3.1 route-map RM-AS10-IN in
```

! BGP force-update and check result

```
R15#sh ip bgp
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
* i	0.0.0.0	10.0.4.1	0	100	0	10 100 i
*>		10.0.3.1		110	0	10 100 i

```
R16#sh ip bgp
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>i	0.0.0.0	10.0.3.1	0	110	0	10 100 i
*		10.0.4.1			0	10 100 i

8. Advertise prefix x1.5.16.0/24 of Rx5 and Rx6 only to ASx0. Do not advertise any other AS beyond it. (no-export, no-advertise)

```
! create prefix, update route-map
R15(config)#
ip prefix-list PRF-5.16 permit 11.5.16.0/24
route-map RM-AS10-OUT permit 10
  match ip address prefix-list PRF-5.16
  set metric 20
  set community no-export
route-map RM-AS10-OUT permit 100
  set metric 20
```

! do the same config at Rx6 but keep lower metric value

```
! update bgp session and check result
R13#sh ip bgp 11.5.16.0/24
BGP routing table entry for 11.5.16.0/24, version 77
Paths: (1 available, best #1, table default, not advertised to any peer)
  Not advertised to any peer
  Refresh Epoch 4
  101
    10.0.3.2 from 10.0.3.2 (11.5.17.1)
      Origin IGP, metric 10, localpref 100, valid, external, best
      Community: no-export
      rx pathid: 0, tx pathid: 0x0
R10#sh ip bgp 11.5.16.0/24
% Network not in table
R20#sh ip bgp 11.5.16.0/24
% Network not in table
```

Thank you

Q&A

Thein Myint Khine
theinmyintkhine@mm-ix.net

www.mmnog.net.mm
event@mm-ix.net