KDDI Myanmar

OJT for Phase3
(Mingalabar ☺)

KDDI Myanmar KATSURA
katsura@kddi.com.sg

Training Agenda ;

KDDI Myanmar

scope of Phase 3 ;

- Phase3 summary
  -> Layout , network diagram , port info , address info
- Tech summary for Phase3
- Compare a configuration btw ASR and C76 for INGW
  (before and after)
- Internet overview
- BGP attribute (traffic control)
  -> how to control a traffic route
- Basic command for ios XR

additional requirement ;

(it is not Phase3 portion ,therefore it is just summary and if no time , we skip it)

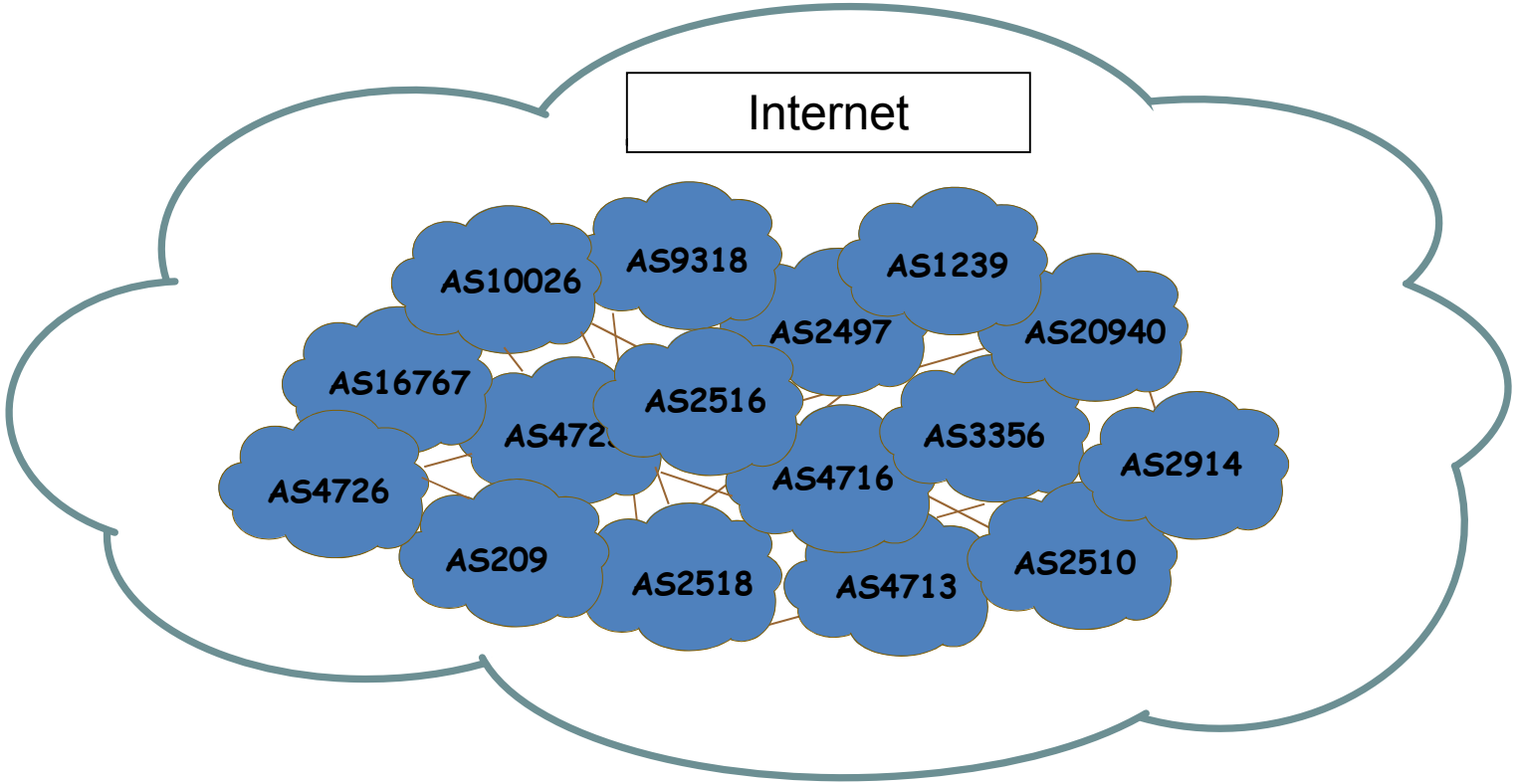
- IPv6 summary
- VPN service summary

Internet Overviwer

What's internet ?

KDDI Myanmar

Internet mean , to connect a network between ISP and ISP or IX
Each ISP network is called AS (Autonomous System)

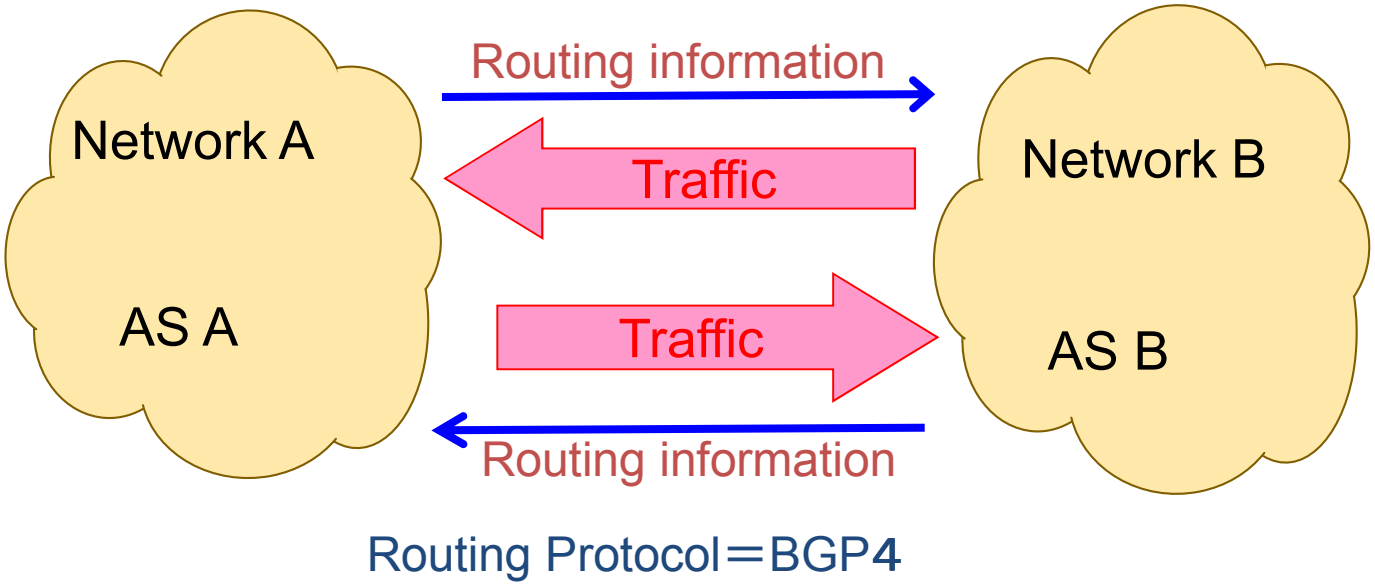


Connect between AS and AS

KDDI Myanmar

Exchange routing information by BGPv4

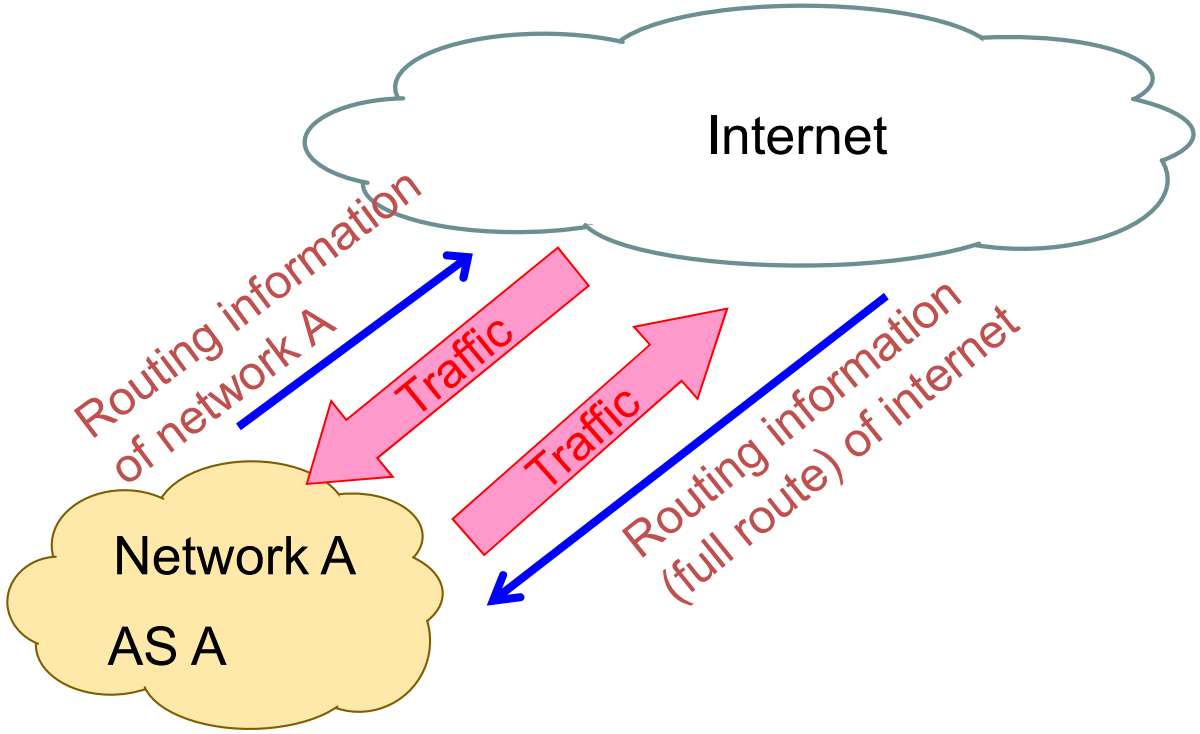
- Advertised routing information : sent a route info to partner ISP or IX
- Received routing information : receive a route info from partner ISP or IX



How to reach to internet ?

KDDI Myanmar

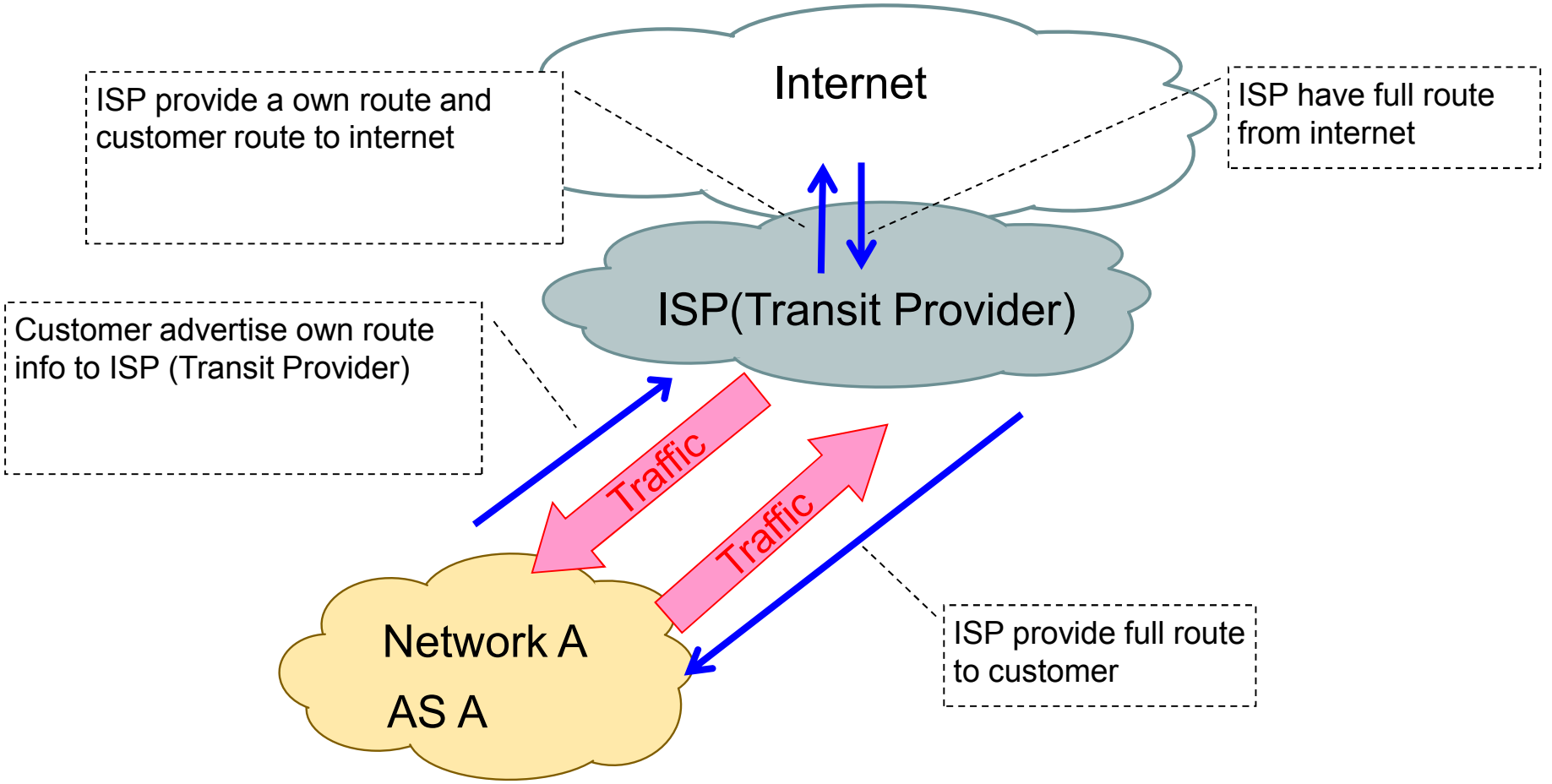
for reach to internet ;
- AS needs to advertise its own routing information to Internet (ISP or IX)
- AS needs to receive full-route (or default gate) information from Internet (ISP or IX)



To provide a internet connection

KDDI Myanmar

IP Transit Provider can provide a full route information to customer ;



KDDI Myanmar

Single Home

- Single connection to internet
- Simple provisioning and operation(no need to consider a traffic route)
- Internet route is depend on provider

Multi Home

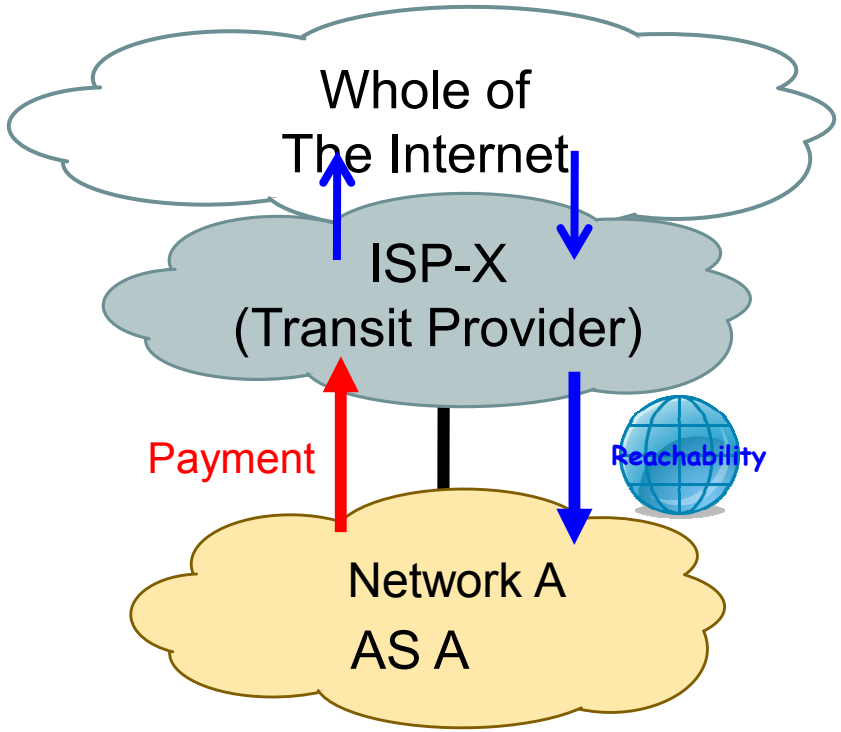
- Multi connection to internet
- Have a redundancy to internet
- Can control a traffic route by BGP attribute
- Mast consider best route because multi home have many route to reach same destination

KDDI Myanmar

IP Transit ;

Provider provide and sell a internet (full route) to customer

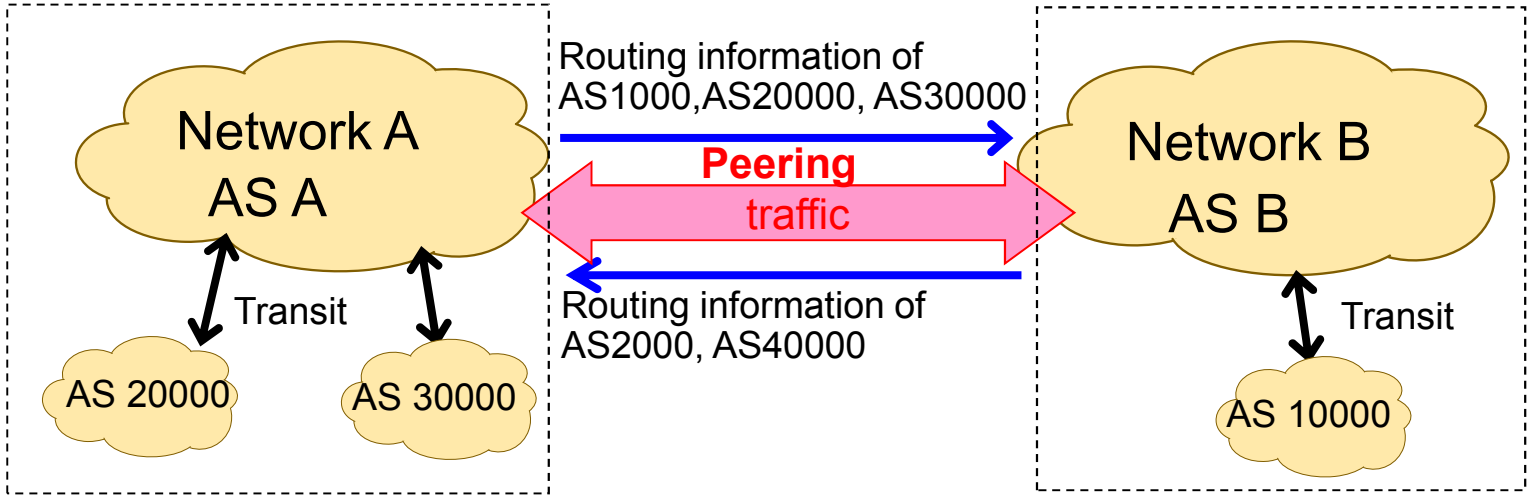
Customer buy a internet (full route) from Provider



KDDI Myanmar

Peering ;

- **Exchange own route info between ISP and ISP by free**
- **Basically ISP do not use this way with small scale ISP (should be customer in that case)**
- **exchange own AS and customer AS only between each other**



What is advantage/disadvantage of Peering ?

KDDI Myanmar

Pros ;

- **Can exchange a traffic route without payment**
- **Can control a traffic route individually**
- **Can have a direct connection with partner (will be best path)**

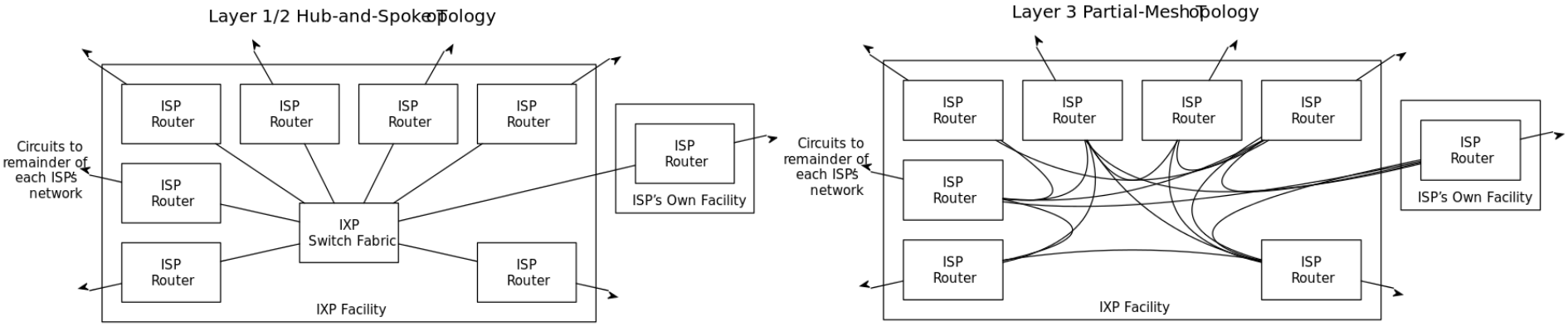
Cons ;

- **Can not get full route by peering (must use IP transit for get a full route)**
- **Need to consider physical circuit cost (Cost-effectiveness is lower than IP transit)**

What is IX ?

IX (Internet exchange) ;

- Interconnection
  - one physical connection to IX L2SW
  - IX do not provide a carrier negotiation
- IX value is depend on No# of ISP (many connection with ISP is high value)
- IX provider co-locate a interconnection SW in DC
- Basically IX if for peering only (should not use for IP Transit)



How can be Tier-1 ?

KDDI Myanmar

Tier-1 (Carrier) can control a whole internet traffic route because they have a global backbone and do not need to buy a IP Transit from ISP .
Most of them is US companies yet .

Major Tier 1 ;
Level3
Cogent
Sprint
Century Link
Global Crossing
AT&T

BGP session
(these information is for cisco router and some information will be different on other device like juniper , fotigate , etc …)

Agenda for BGP session ;

KDDI Myanmar

- Introduction
- eBGP Peering
- iBGP Peering
- Attributes and Best Path Selection Algorithm
- Route Origination
- AS-PATH
- NEXTHOP
- Communities
- Controlling Outbound Traffic
- BGP Multipath
- Controlling Inbound Traffic
- Route Reflectors
- Initial Convergence
- BGP Routing Convergence
- High Availability
- Practice

KDDI Myanmar

Introduce a BGP summary on this part …

Autonomous System

KDDI Myanmar

Within the internet , an autonomous system (AS) is a collection of connected IP routing prefixes under the control of one or more network operators on behalf of a single administrative entity or domain that presents a common, clearly defined routing policy to the Internet .

AS Numbers ;

2-byte bytes AS number

- 1 to 65535
- 64512 to 65535 are private
- RFC 1771 (original definition)
- RFC 1930 (newer definition)

4-byte AS number

- RFC 4893
- Unique AS for every IPv4 address
- 2^{32} or 4,294,967,296 AS numbers (ranging from 0 to 4294967295)

IGP and EGP

KDDI Myanmar

IGP – Interior Gateway Protocol

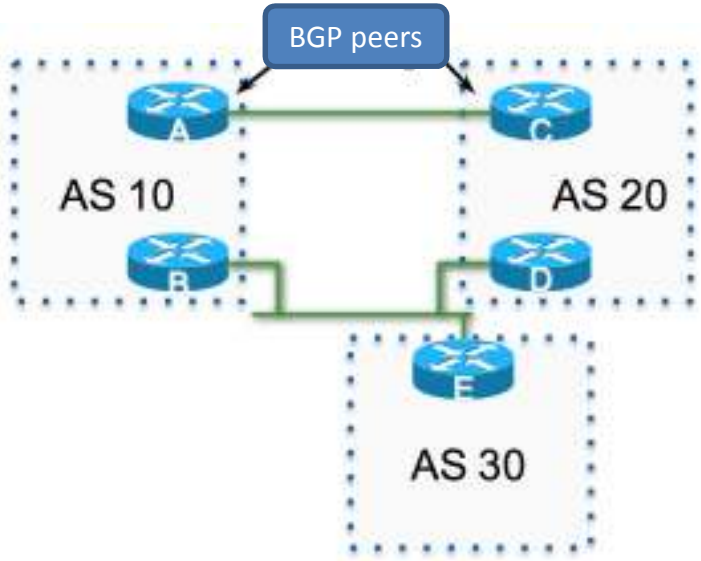
– Exchange routes within an Autonomous Systems
– Limited Scalability
– Sub-second convergence
– OSPF, ISIS, EIGRP, etc.

EGP – Exterior Gateway Protocol

– Exchange routes between Autonomous Systems
– Once was an EGP called “EGP”
– BGP is standard EGP
– Slower convergence in exchange for scalability

KDDI Myanmar

- An AS originate's routes into BGP
- BGP peers with other BGP speakers
  - Peer is also called "neighbor"
  - Uses TCP port 179
- BGP peers exchange routes
- Picks the best path
  - Installs in the forwarding table
  - Advertises to BGP peers via UPDATEs
- UPDATEs have Attributes
  - **Routing policies tweak attributes to influence best path selection**



KDDI Myanmar

-BGP is classified as a path vector routing protocol **(see RFC 1322)**

A path vector protocol defines a route as a pairing between a destination and the attributes of the path to that destination.

12.6.126.0/24	**207.126.96.43**	**1021**	**0 6461 7018 6337 11268 i**

AS Path

KDDI Myanmar

- Once BGP sends a route to a peer, it assumes the peer will keep it

- There is no periodic refresh

- New UPDATEs are sent when

  – Best path change
  – Peer bounces
  – Route-Refresh

eBGP peers

KDDI Myanmar

What is eBGP peers and what is different with iBGP ?

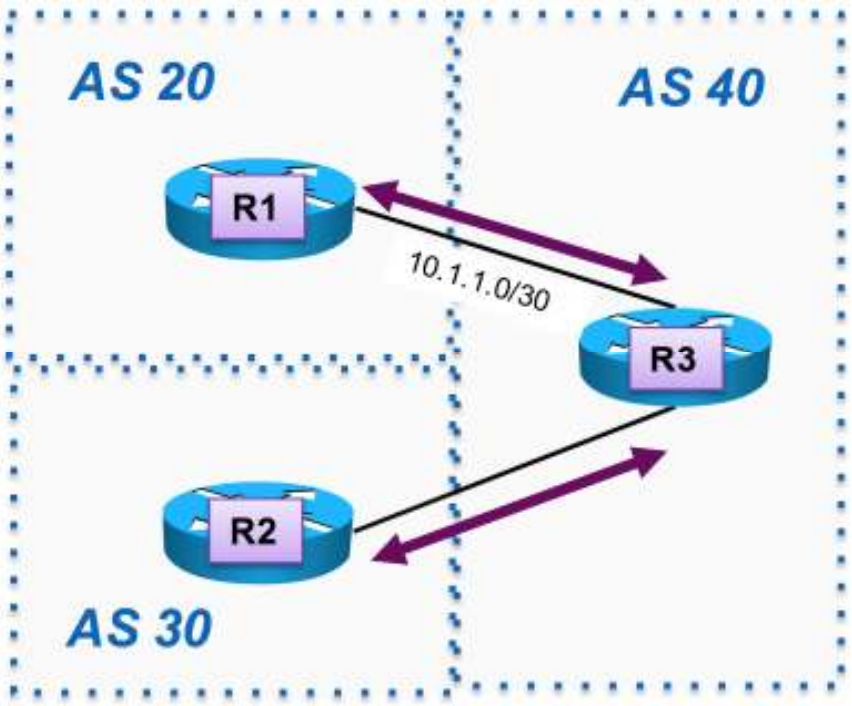
eBGP – External BGP

KDDI Myanmar

- Neighbor in different AS
- Usually directly connected
- NEXTHOP set to self
- R1 -> R3 and R2- > R3
- R1 -> R3 config

R1
router bgp 20
neighbor 10.1.1.2 remote-as 40

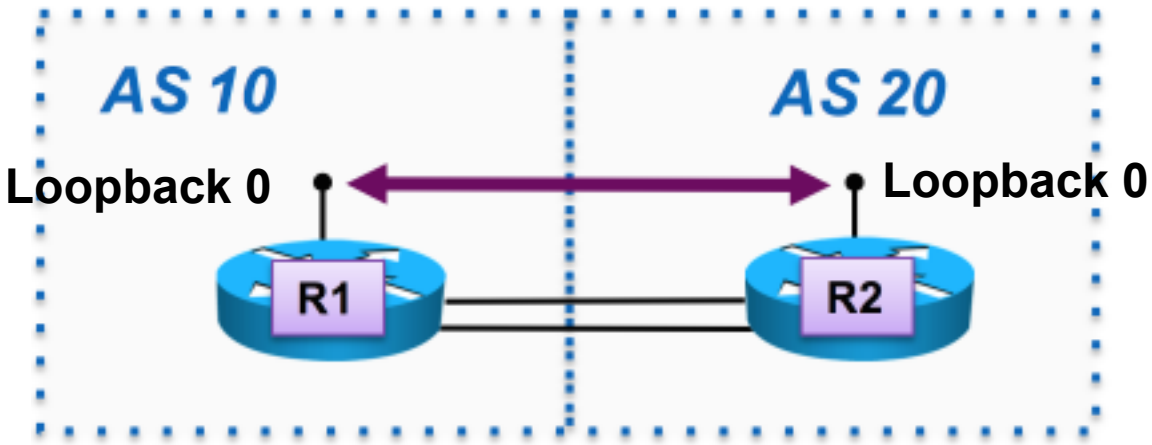
R3
router bgp 40
neighbor 10.1.1.1 remote-as 20



KDDI Myanmar

- Peer between loopbacks
- Often used to load-balance traffic over multiple links

```
router bgp 10
 neighbor 10.1.20.1 remote-as 20
 neighbor 10.1.20.1 update-source loop0
 neighbor 10.1.20.1 ebgp-multihop 2
ip route 10.1.20.1 255.255.255.255 s0/0
ip route 10.1.20.1 255.255.255.255 s1/0
```



What is iBGP peers and what is deferent with eBGP ?

KDDI Myanmar

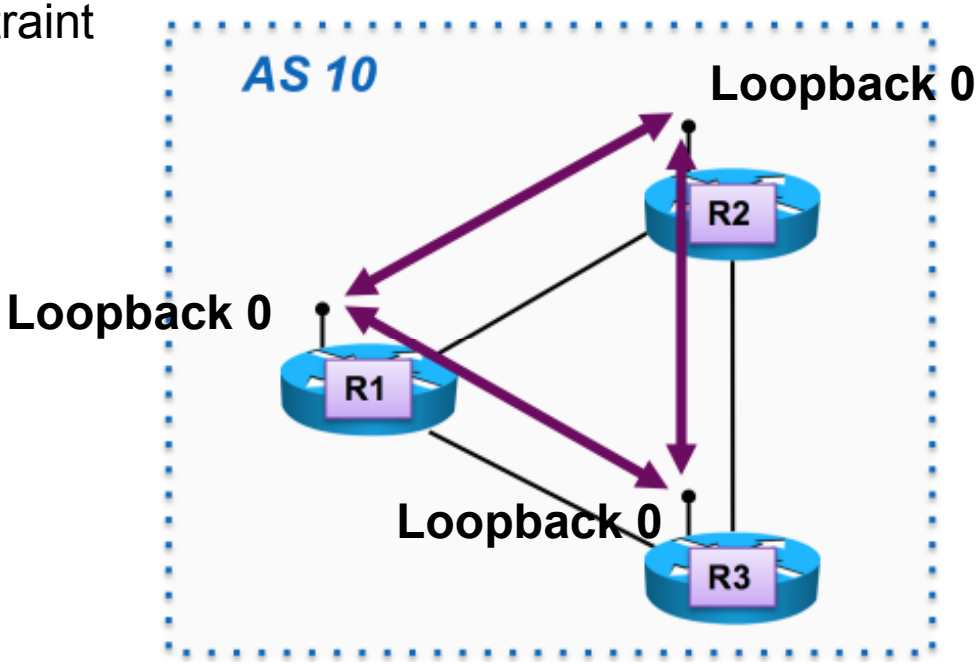
- Neighbor in same AS
- NEXTHOP is unchanged
- Peer to loopbacks (/32)
- Cannot advertise route received from one iBGP peer to another iBGP peer
  - Full iBGP mesh is required
  - n*(n-1)/2 peering mesh – scaling problem!
  - Route-Reflectors relax this constraint

R1

```
router bgp 10
 neighbor 10.1.1.2 remote-as 10
 neighbor 10.1.1.2 update-source loop0
```

R2

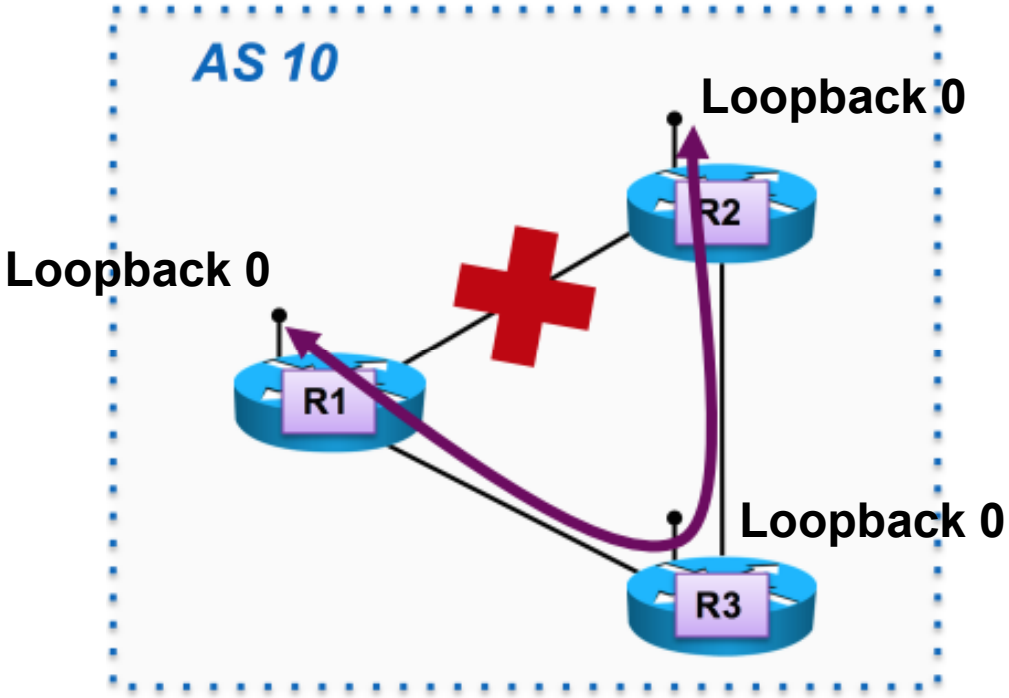
```
router bgp 10
 neighbor 10.1.1.1 remote-as 10
 neighbor 10.1.1.1 update-source loop0
```



iBGP – Loopback Peer (what is Pros of loopback peer ?)

KDDI Myanmar

- Loopback peer promotes stability
- There are two paths between R1 and R2
- If the link between them fails
  – Peering to the interface IP would bring down the BGP session
  – Peering to a loopback allows the session to stay up



Attributes and Best Path Selection Algorithm

KDDI Myanmar

How to control a traffic route ?

Attributes

KDDI Myanmar

IGP :

– Primary attribute is a cost/metric
– The path with the lowest metric is the best...nice and easy

EGP :

– Routing Policy between AS is usually more complex than this
– Shortest path is not necessarily the best one
– Has many attributes to describe reachability to a destination
– The “Best Path Algorithm” compares attributes between different paths to select a best path
– Route-policies are used tweak attributes to influence routing

BGP Route and BGP Path

KDDI Myanmar

- “show ip bgp summary” provides the total number of routes and paths
- Paths and routes both consume memory
- The more paths you have per route, the more memory consumed

```
r1#show ip bgp summary
BGP router identifier 10.1.1.1, local AS number 65100
BGP table version is 479140003, main routing table version 479140003
478079 network entries using 118563592 bytes of memory
4217545 path entries using 472365040 bytes of memory
330373/82377 BGP path/bestpath attribute entries using 76646536 bytes of memory
…
BGP using 674846186 total bytes of memory
BGP activity 3211514/2716352 prefixes, 433901680/429525155 paths, scan interval 60 secs
```

Your router is ok ?

BGP Path Selection Algorithm

KDDI Myanmar

	Attribute	Logic
1	**Weight**	Higher is better. Local to the router…not really an attribute.
2	**Local Preference**	Local to an AS…higher is better
3	**Locally Originated**	Corner case…"network 10.0.0.0" vs. "aggregate 10.0.0.0" vs. "redistribute" on the same router
4	**AS-PATH**	Shorter AS-PATH is better
5	**ORIGIN**	IGP < EGP < Incomplete
6	**MED**	Is often a reflection of IGP metrics so lower is better
7	**eBGP vs. iBGP**	Prefer eBGP path over iBGP path
8	**IGP cost to NEXTHOP**	Lower is better
9	**Lowest Router ID**	Lower is better
10	**Shortest CLUSTER_LIST**	Lower is better
11	**Lowest neighbor IP address**	Lower is better

Take a break : Which route is best ?

KDDI Myanmar

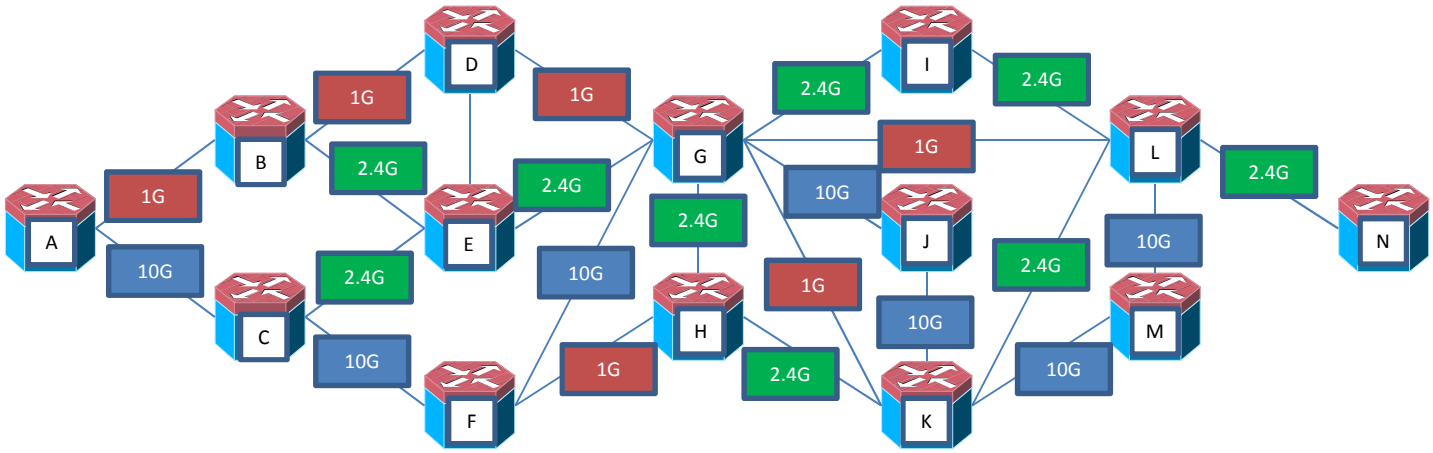
- BGP can have multiple paths per router
- Here show command result , which is best path route ?
- Is it external route or internal route ?
- Why do you think so ?

```
r1#show ip bgp 40.1.1.0
BGP routing table entry for 40.1.1.0/24, version 14
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  20 40
    20.1.1.1 from 20.1.1.1 (20.1.1.1)
      Origin IGP, localpref 100, valid, external
  30 40
    10.1.1.2 (metric 11) from 10.1.1.3 (10.1.1.3)
      Origin IGP, metric 0, localpref 200, valid, internal, best
      Originator: 10.1.1.2, Cluster list: 10.1.1.3
```

Take a break : Which route is best ?

KDDI Myanmar

- Each peers are different ISP and be connected each other by eBGP
- Each peers bandwidth is below (on this scenario , we do not consider latency however actual case , we must consider it because latency also effect to traffic speed)
- For get a best path to N from A , how should each ISP do ?



Route Origination

KDDI Myanmar

As first step .. (welcome to BGP world)

Route Origination – Network Statements

KDDI Myanmar

- Easiest/Cleanest method

- Network 10.1.1.0 mask 255.255.255.0
  – Requires 10.1.1.0/24 to be in the RIB (routing information base)
  – Floating static route to Null0 is common
  – Originates 10.1.1.0/24

- Easy to determine/control what you are originating

```
R1#
router bgp 10
  network 10.1.1.0 mask 255.255.255.0
!
ip route 10.1.1.0 255.255.255.0 Null0
```

Route Origination – Network Statements

KDDI Myanmar

- The Origin is IGP
- Weight is 32768
- “0.0.0.0 from 0.0.0.0”
  – The first 0.0.0.0 is the NEXTHOP
  – The second 0.0.0.0 is the peer we learned it from

R1#show ip bgp 10.1.1.0 255.255.255.0
BGP routing table entry for 10.1.1.0/24, version 2 Paths: (1 available, best #1, table default)
Advertised to update-groups: 9 12
Local
0.0.0.0 from 0.0.0.0 (10.1.1.2)
Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local, best

Route Origination – Redistribution

KDDI Myanmar

- Routes can be redistributed into BGP

- Pros
  - Easy to configure and setup

- Cons
  - IGP instability is passed along to BGP
  - Isn't always obvious what routes you are originating

(*) use route-maps to control what you're redistributing

R1#
router bgp 10
redistribute ospf 10 route-map FOO

Address-Family-Identifier Syntax

KDDI Myanmar

- In the beginning BGP only supported IPv4 Unicast
- IPv4 is the AFI, Unicast is the SAFI
- Today , there are many supported AFI/SAFIs
  ->IPv6 Unicast, VPNv4, IPv4 Multicast, etc
- AFI/SAFI specific configuration happens in a sub-context
  ->Network statements, route-maps on neighbors, etc
- Non AFI/SAFI configuration still happens directly under ‘router bgp’
  ->Remote-as, update-source, keepalive timers, etc

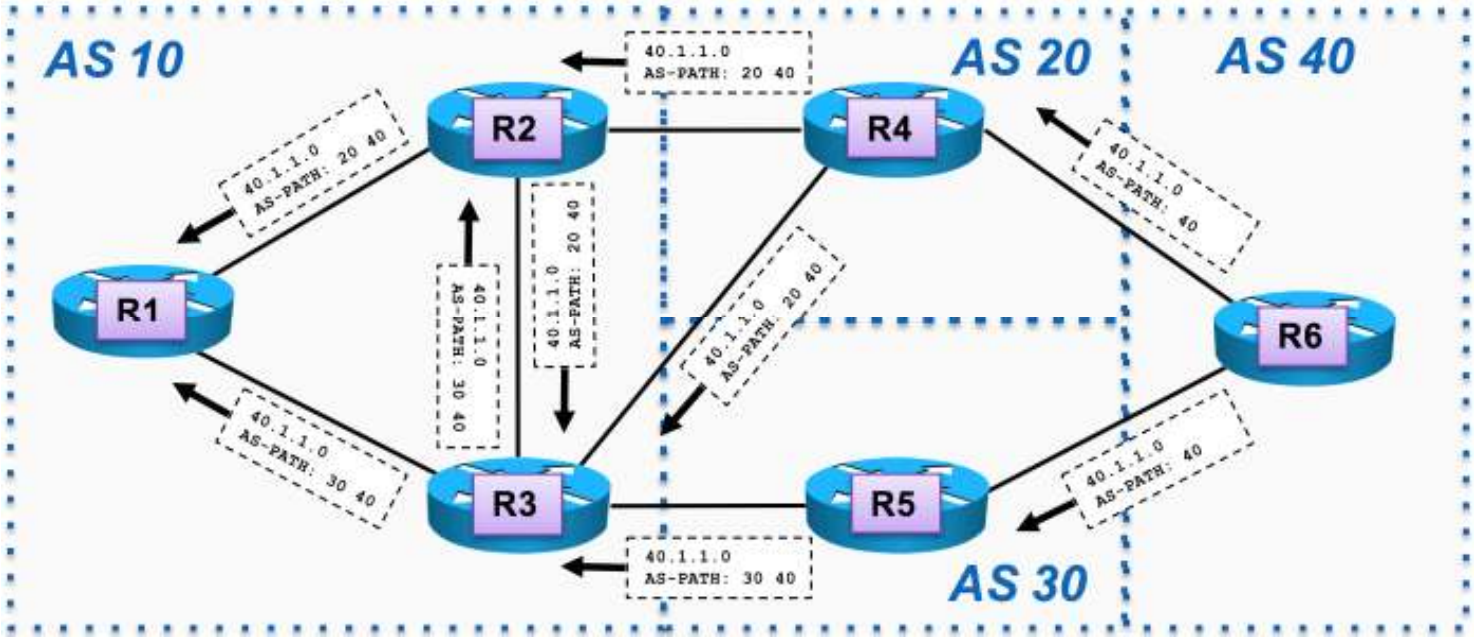
Original syntax	AFI/SAFI syntax
!	!
router bgp 100	router bgp 100
neighbor 1.1.1.1 remote-as 50	neighbor 1.1.1.1 remote-as 50
neighbor 1.1.1.1 route-map FOO out	!
network 10.1.1.0 mask 255.255.255.0	address-family ipv4 unicast
!	network 10.1.1.0 mask 255.255.255.0
	neighbor 1.1.1.1 route-map FOO out
	neighbor 1.1.1.1 activate
	!
	address-family vpnv4 unicast
	neighbor 1.1.1.1 send-community ext
	neighbor 1.1.1.1 activate
	!
	Example with multiple AF per neighbour

AS path

AS path is key person on internet (internet master)

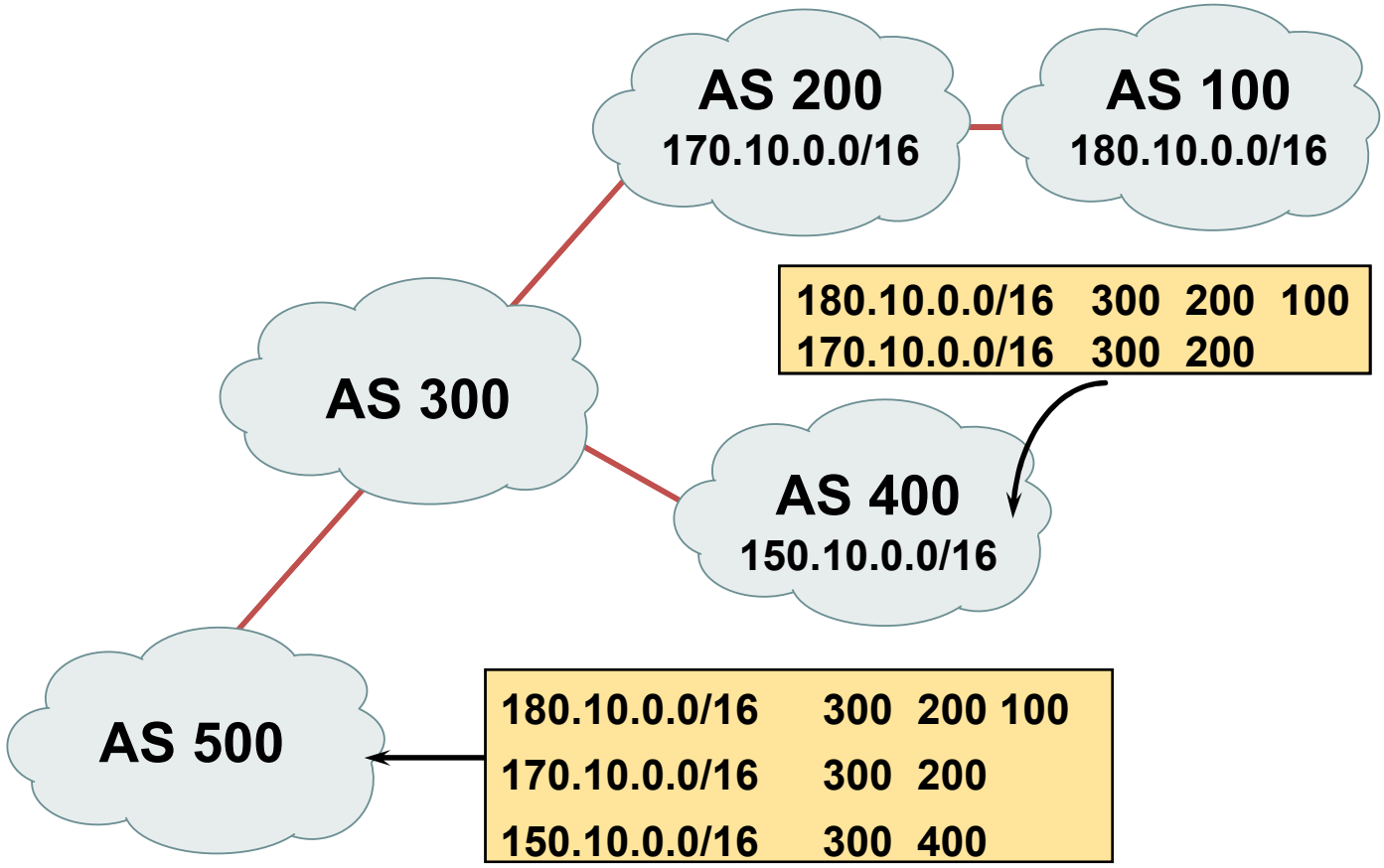
KDDI Myanmar

- The AS-PATH tells the story of what ASes a route has been through
- AS-Path is used for loop detection on the border of the AS
  ->BGP drops an external update if it sees its own AS in the path
- The most recent AS is on the left, the originating AS is on the far right
- BGP prepends his own AS# to the AS-PATH when advertising to an eBGP peer
- Shortest AS-PATH is often the tie-breaker for best path selection



KDDI Myanmar

Best route is decided by no# of ASes



KDDI Myanmar

How to check a AS – path ?

R1#show ip bgp 40.1.1.0
BGP routing table entry for 40.1.1.0/24, version 54 Paths: (1 available, best #1, table default)
Advertised to update-groups: 5
30 40, (Received from a RR-client)
10.1.1.3 (metric 11) from 10.1.1.3 (10.1.1.3)
Origin IGP, metric 0, localpref 200, valid, internal, best R1#

AS path is key of them on internet

KDDI Myanmar

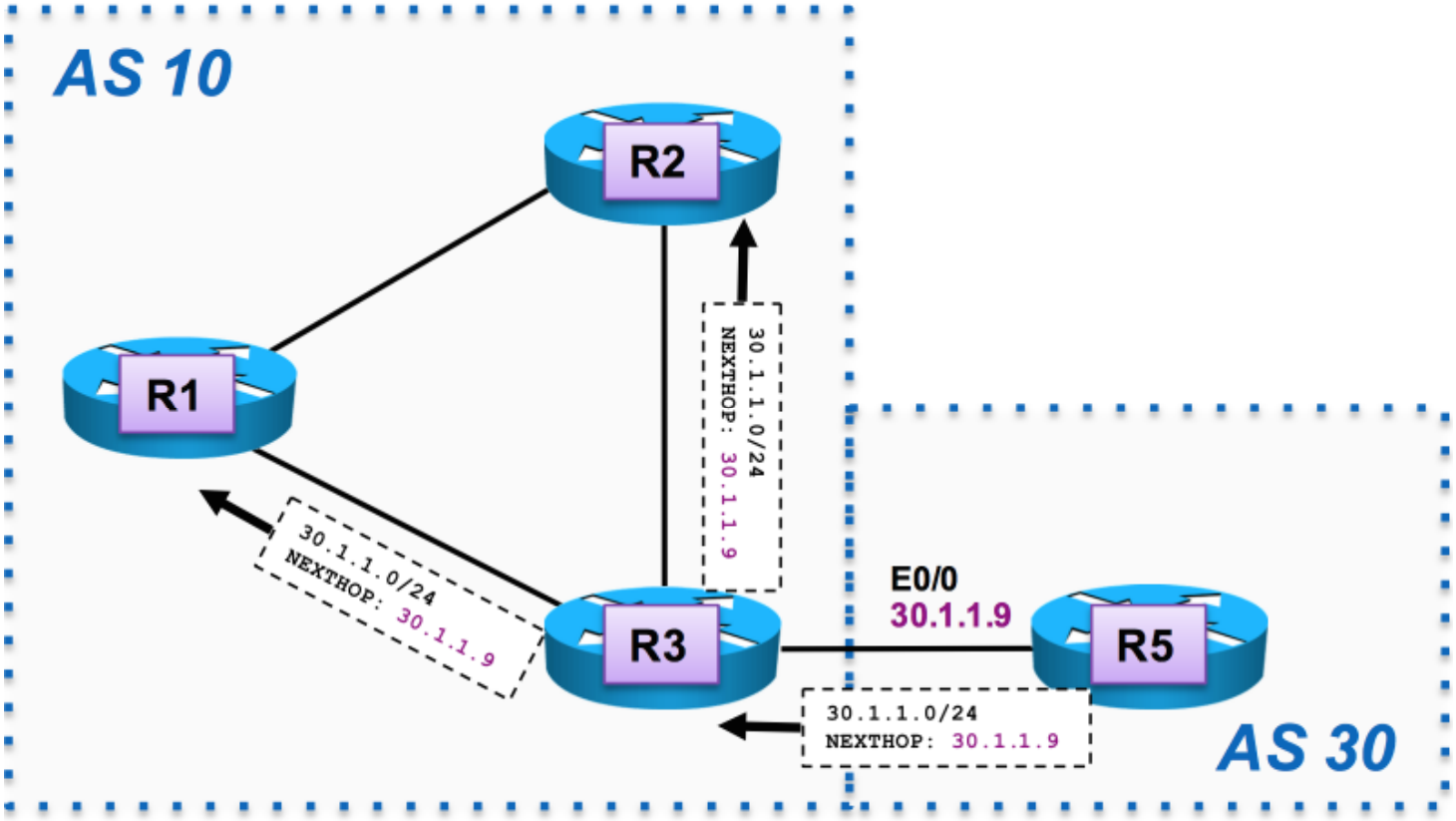
- NEXTHOP is the address that we must route towards in order to reach the BGP prefix
  - Paths where the next-hop is unreachable are not considered for best-path calculation

- eBGP does “next-hop-self” automatically
  - Multiple eBGP peers on the same subnet is an exception

- iBGP does not modify the NEXTHOP by default
  – NEXTHOP will remain as the IP of the eBGP peer
  – Forces BGP speakers in an AS to have routes for the eBGP facing links
  – Would need to carry many /30 eBGP facing links in our IGP
  – Best practice is to use “next-hop-self” to avoid this

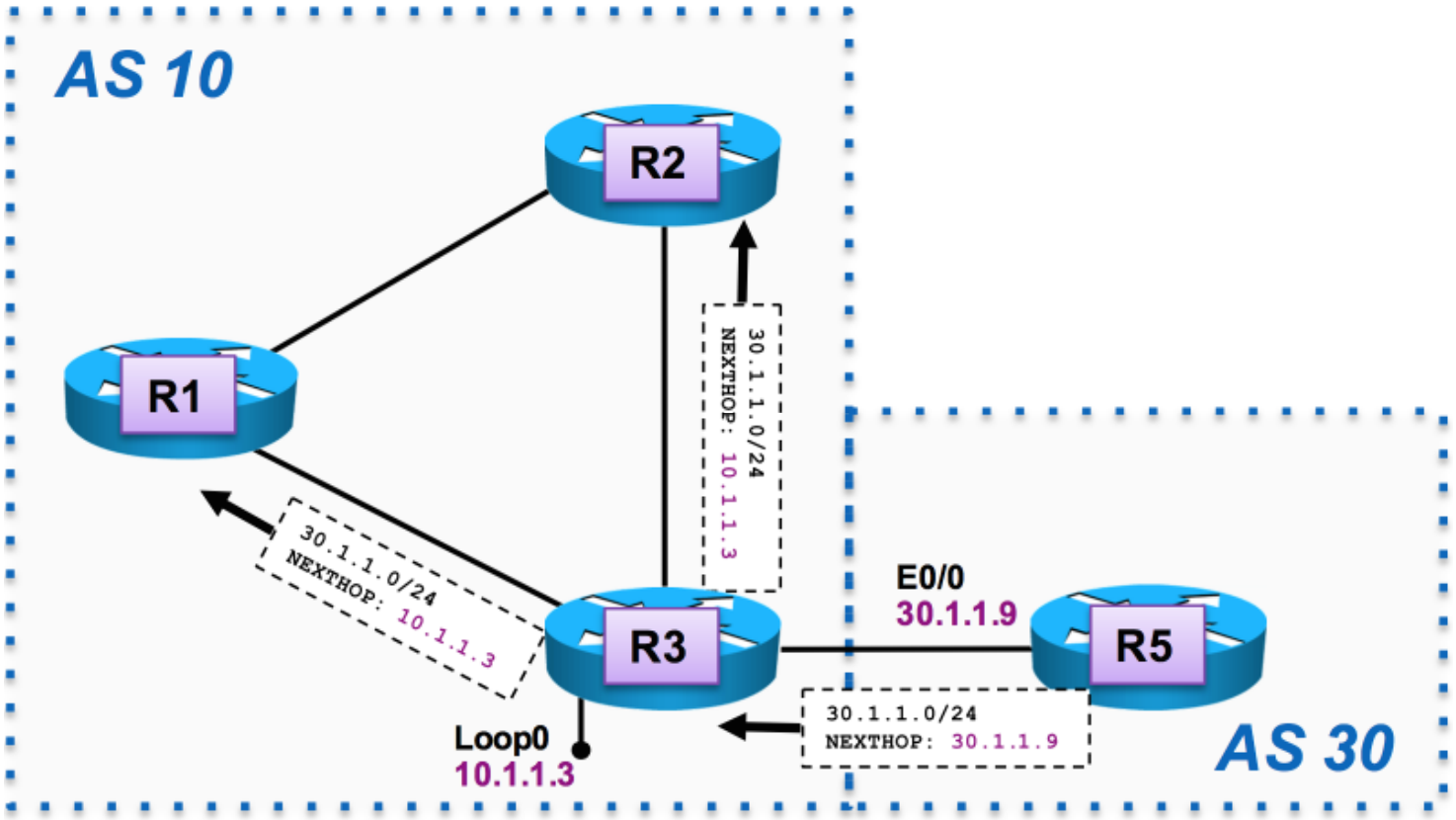
KDDI Myanmar

- NEXTHOP does not change
- AS 10's IGP must have route to 30.1.1.9



iBGP with next-hop-self

- R3 changes NEXTHOP to his “update-source” interface
- iBGP should always use loopback peering
- AS 10’s IGP has a route to R3’s loopback **10.1.1.3**



Communities

KDDI Myanmar

Communities for grouping and traffic control

Communities

KDDI Myanmar

- A COMMUNITY is an attribute that stores a number
  - 4-byte number that is usually displayed in X:Y notation
  - "ip bgp-community new-format" triggers X:Y notation
- Set communities via a route-map
- Communities are not advertised by default
- A community by itself does nothing
  – Tagging a prefix with 100:1 or 100:2 will not change routing in any way

How to send a community ;

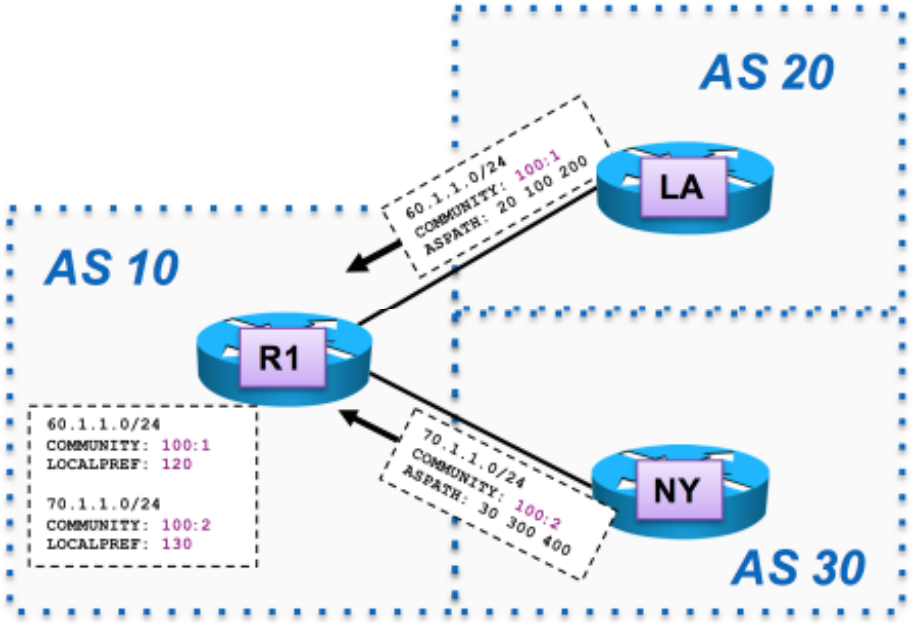
```
router bgp 20
neighbor 10.1.1.2 remote-as 10
neighbor 10.1.1.2 send-community
neighbor 10.1.1.2 route-map Com-send out !
ip bgp-community new-format
!
route-map Com-send permit 10
set community 10:1
```

Communities

KDDI Myanmar

- Applying Policy towards communities does impact routing
- Use route-maps and community-list to
  - Match against a certain community
  - Modify a BGP attribute as a result
    - LOCALPREF, ASPATH prepending, etc
- You can impact 1000s of prefixes by applying policy based on a single community

```
R1#
router bgp 10
 neighbor 20.1.1.1 description LA_PEER
 neighbor 20.1.1.1 route-map NY_OR_LA in
 neighbor 30.1.1.1 description NY_PEER
 neighbor 30.1.1.1 route-map NY_OR_LA in
!
ip community-list standard VIA_LA permit 100:1
ip community-list standard VIA_NY permit 100:2
!
route-map NY_OR_LA permit 10
 match community VIA_LA
 set local-preference 120
!
route-map NY_OR_LA permit 20
 match community VIA_NY
 set local-preference 130
!
route-map NY_OR_LA permit 30
```



Well Known Communities

KDDI Myanmar

- *A community by itself does nothing"*
- There are exceptions to every rule
- Well Known Communities do have an automatic impact

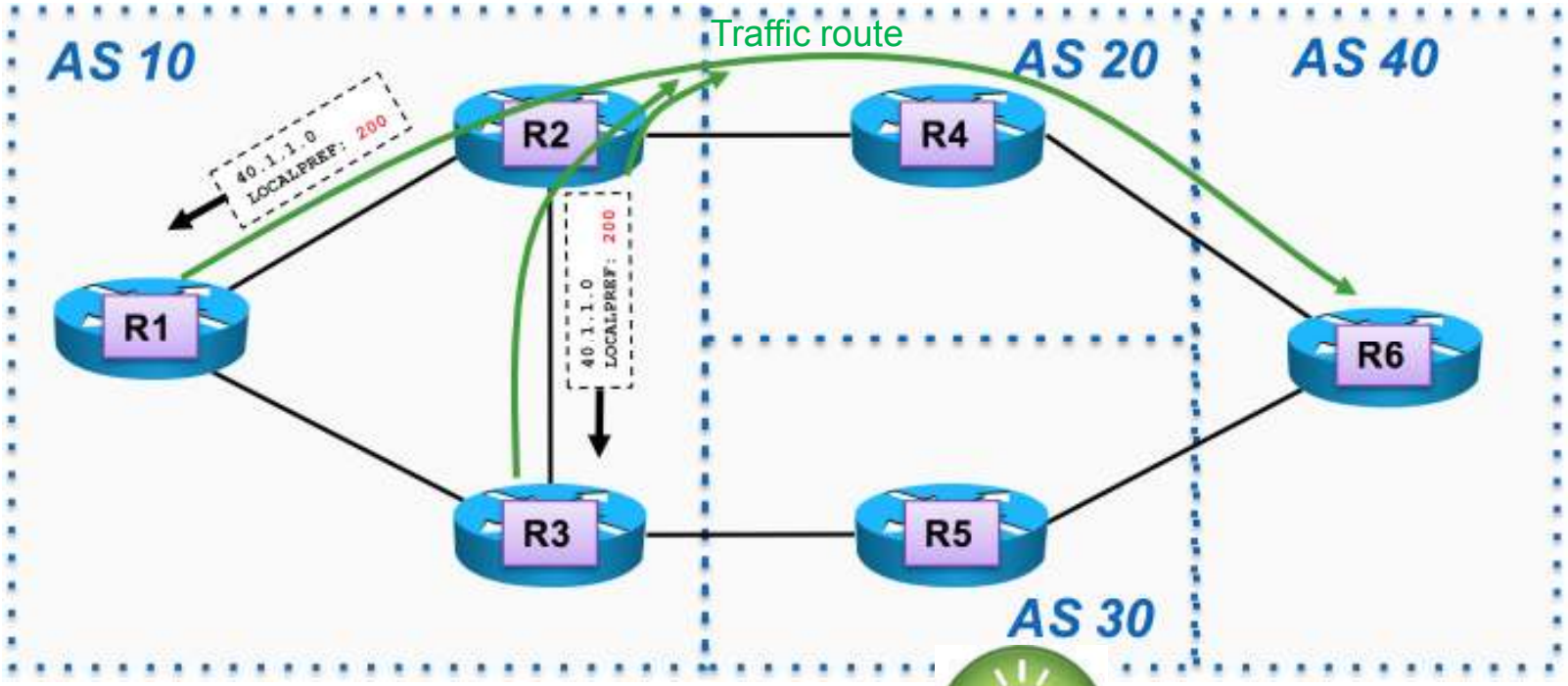
Community	Impact
local-AS	*Do not send to EBGP peers*
no-advertise	*Do not advertise to any peer*
no-export	*Do not export outside AS/confed*

Not only AS path for traffic control .. But also …

Local Preference

KDDI Myanmar

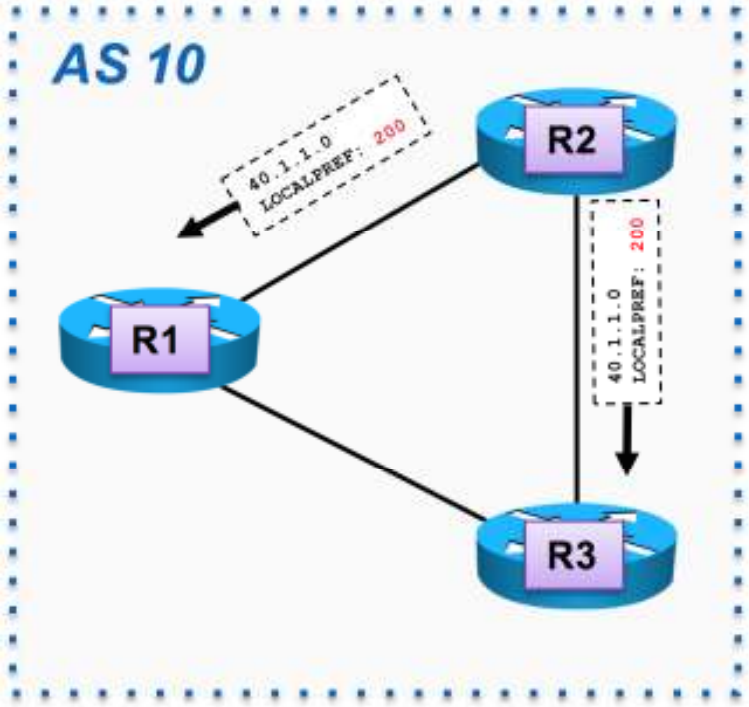
- An attribute used to influence outbound traffic
- Higher LOCAL_PREF is preferred
- Is compared very early in the Best Path Algorithm
- Is local to an AS
  – Local preference is never transmitted to an eBGP peer
  – A default LP of 100 is applied to routes from eBGP peers



(*) need a input policy for influence to inbound

Local Preference

- You can also control a inbound traffic route (iBGP)




```
R2#
!
router bgp 10
 neighbor 10.1.1.1 remote-as 10
 neighbor 10.1.1.1 route-map SET_LOCAL_PREF out
 neighbor 10.1.1.3 remote-as 10
 neighbor 10.1.1.3 route-map SET_LOCAL_PREF out
!
route-map SET_LOCAL_PREF permit 10
 set local-preference 200
!
```

Or: set localpref inbound on eBGP session.. There are always multiple ways to skin a cat :-}

- Local preference is a very “heavy” attribute to influence routing, as it is evaluated very early in best path algorithm
- Especially with Internet routing, AS path length is very important (how “far” is the destination away from me)
- Hence, evaluate other attributes like MED for best path manipulation
- No one size fits all, there are lots of ways to implement BGP routing policies...

BGP Policy

KDDI Myanmar

- Policy based on various attributes:
  - ASPATH
  - Community
  - Destination prefix
  - Many, many others…

- Reject/accept selected routes

- Set attributes to influence path selection

- Tools(IOS):
  - Distribute-list or prefix-list
  - Filter-list (as-path access-list)
  - Community-list
  - Route-maps (the Swiss army knife)

Policy Control - Prefix List

KDDI Myanmar

- Per-peer prefix filter, inbound or outbound
- Allows coverage for ranges of prefix lengths (ge, le)
- Based upon network numbers in NLRI (using familiar IPv4 address/mask format)
- Example configuration:

```
router bgp 200
 neighbor 220.200.1.1 remote-as 210
 neighbor 220.200.1.1 prefix-list PEER-IN in
 neighbor 220.200.1.1 prefix-list PEER-OUT out
!
ip prefix-list PEER-IN deny 218.10.0.0/16
ip prefix-list PEER-IN permit 0.0.0.0/0 le 32
ip prefix-list PEER-OUT permit 215.7.0.0/16
ip prefix-list PEER-OUT deny 0.0.0.0/0 le 32
```

le : between subnet mask and /32
i.g.) 10.10.0.0/16 le 32
between /16 to /32 is matched on this policy

ge : between no of value and /32
i.g.) 10.10.0.0/16 ge 24
between /24 to /32 is matched on this policy

Policy Control - Filter List

KDDI Myanmar

- Filter routes based on AS path
- Inbound or Outbound
- Example Configuration:

```
!
router bgp 100
 neighbor 220.200.1.1 filter-list 5 out
 neighbor 220.200.1.1 filter-list 6 in
!
ip as-path access-list 5 permit ^200$
ip as-path access-list 6 permit ^150$
```

Policy Control - Regular Expressions

KDDI Myanmar

Regular Expressions	Remark
.	Match with anything
^	Match with first word
$	Match with last word
_	Match with Space or last word or first word or , or { or }
*	Match with atom sequence (>= 0)
+	Match with atom sequence (>= 1)
?	Match atom or null
-	Between
[ ]	Single character pattern

Regular Expressions	Remark
^100_	Start from 100
^100$	Start from 100 and end at 100
_100$	Originl of 100
_100_	Include 100
^$	Own AS
.*	match anything
^[1-2]00_	Start from 100 or 200
^100_[0-9]*$	Start from 100 ,then , end at 1-9

For route redundancy …

BGP Multipath

KDDI Myanmar

- receives a route from multipath
- Best-path algorithm selects one and installs it in routing table
- Assuming all attributes are equal, uses the one from the lower neighbour IP address
- By default, all of the traffic goes via one link only
- We could do some manual load-sharing via localpref/MED, but that's cumbersome
- Enable eBGP multipath on router to install multipath paths

```
router bgp 10
        maximum-paths 2
```

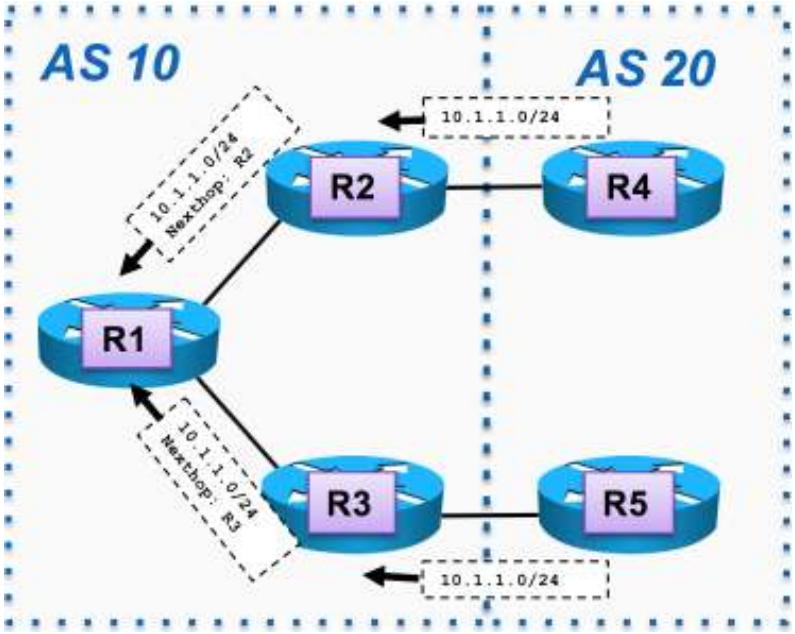
- When R1's IGP cost to R2 and are R3 is equal, and all other path attributes are the same, iBGP multipath can be used

```
router bgp 10
        maximum-paths ibgp 2
```

KDDI Myanmar

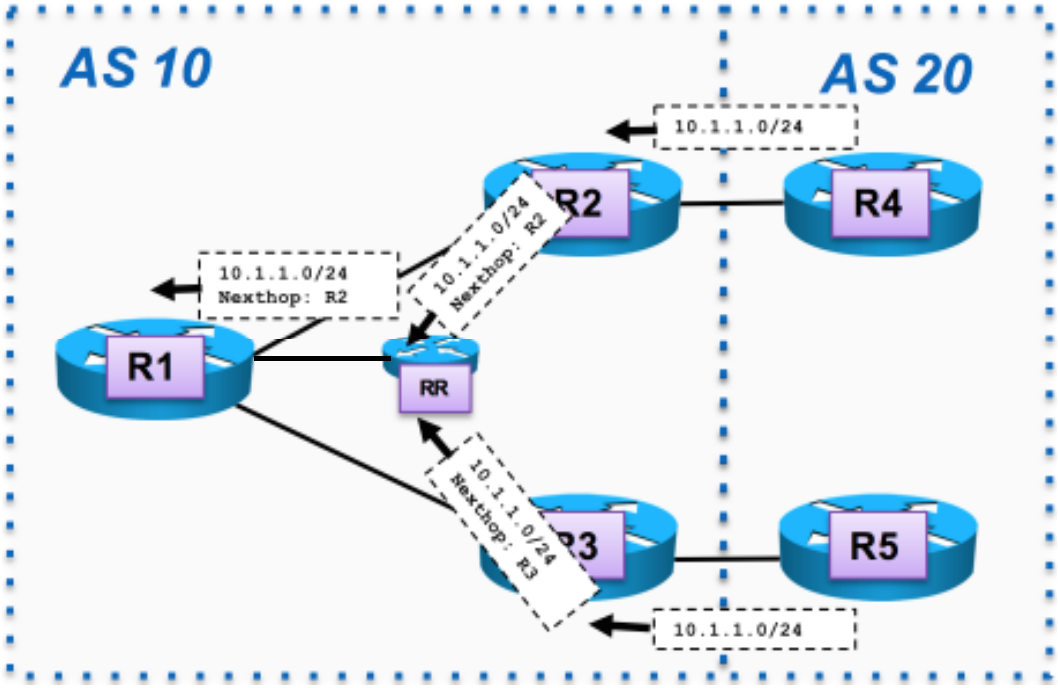
- eBGP Multipath will not help by default
- R1 will choose one best path
- When R1's IGP cost to R2 and are R3 is equal, and all other path attributes are the same, iBGP multipath can be used

router bgp 10
maximum-paths ibgp 2



KDDI Myanmar

- BGP multipath allows to install multiple paths to be installed in the routing table
- BGP will still select one best path and advertises it to its peers
- With RR between R1 and R2/R3, R1 will only receive one path, no load-sharing possible
- Multiple solutions:
  - Enhance BGP protocol behaviour
    using “ADD-PATH” (however no need it as normal operation)
  - Use multiple iBGP sessions on RR (“shadow-RR”)



More deep BGP technology …
Be Mr route controller …

Controlling Inbound Traffic

KDDI Myanmar

The first rule of controlling inbound traffic...

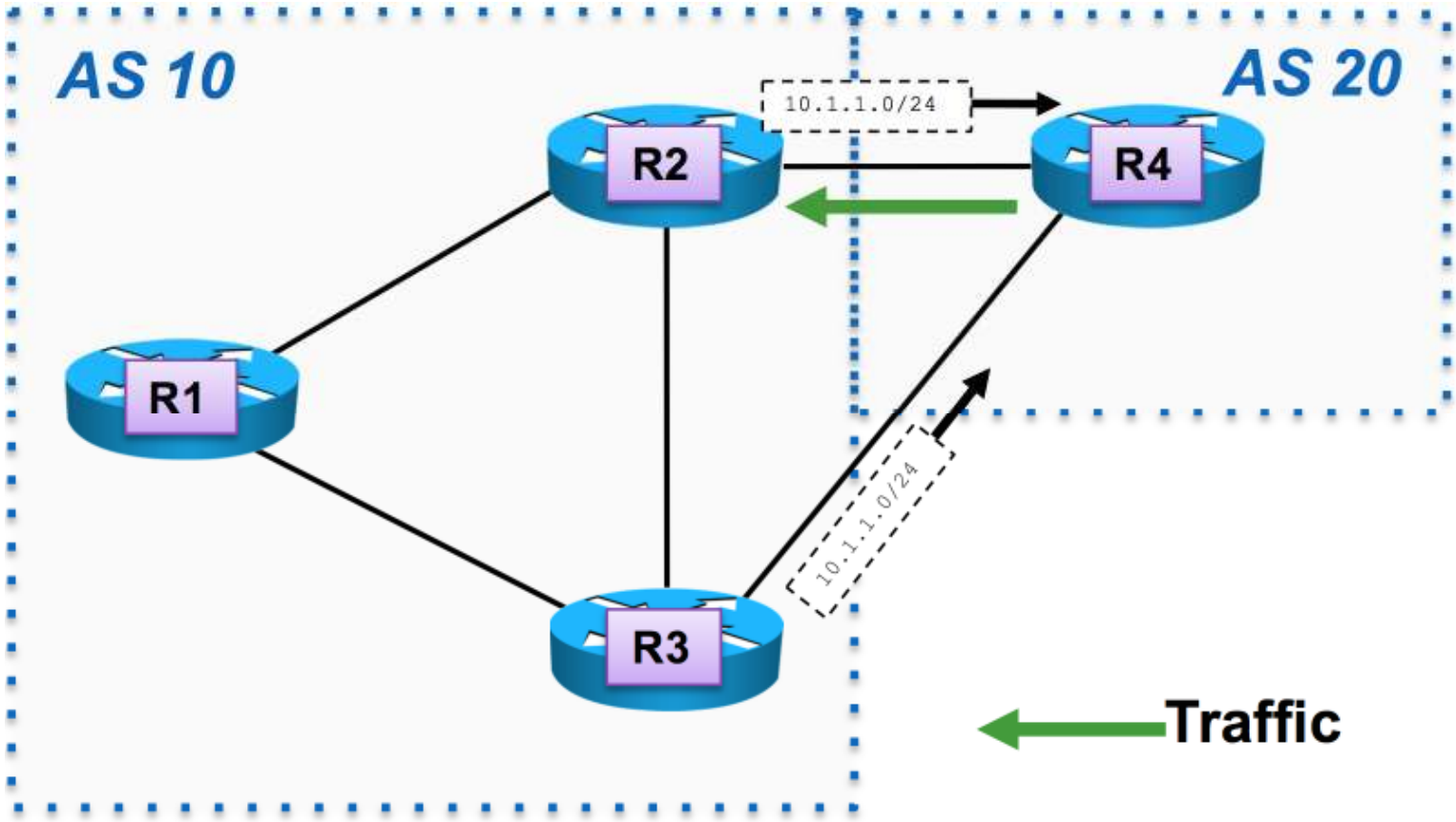
– You do not have ultimate control of how traffic enters your AS
– Your peers may have outbound policies that will override all of your attempts to influence inbound traffic

what are your options?

– MED
– AS-PATH Prepending
– Community/Local Pref agreement

Leaking Specific Route

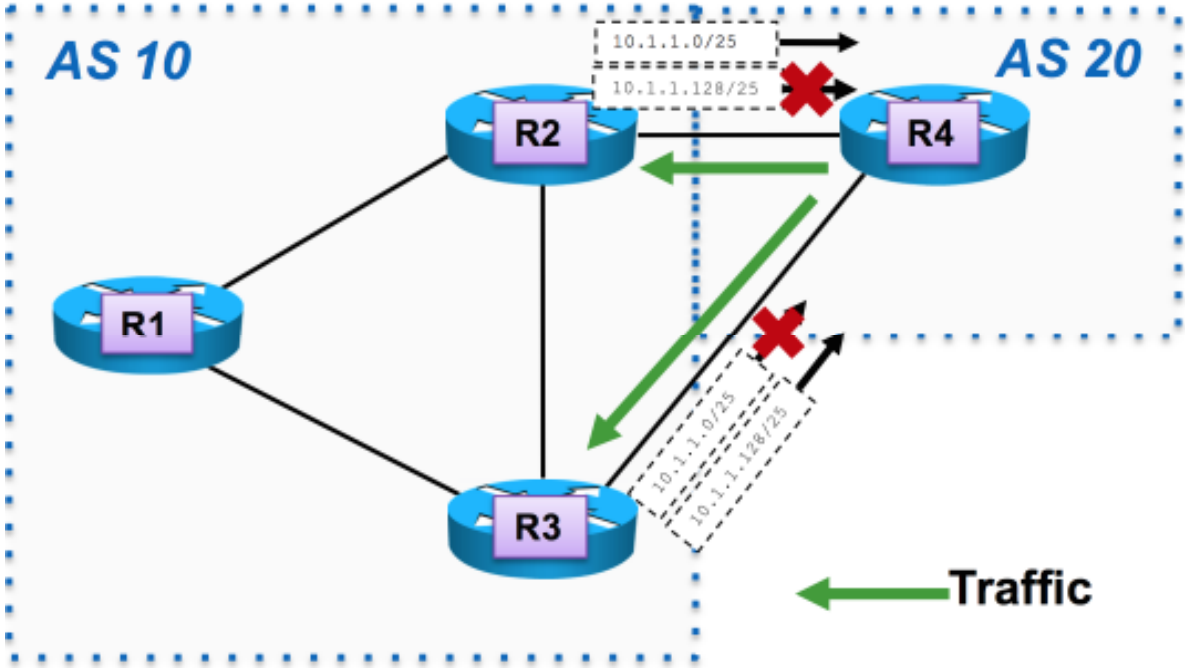
- AS 10 owns 10.1.1.0/24
- AS 20 only uses one link to send traffic to AS 10 (it is not load balance)
- You want to utilize both links



Leaking Specific Route

Split your /24 in two /25s
R2
– advertise 10.1.1.0/25
R3
– advertise 10.1.1.128/25

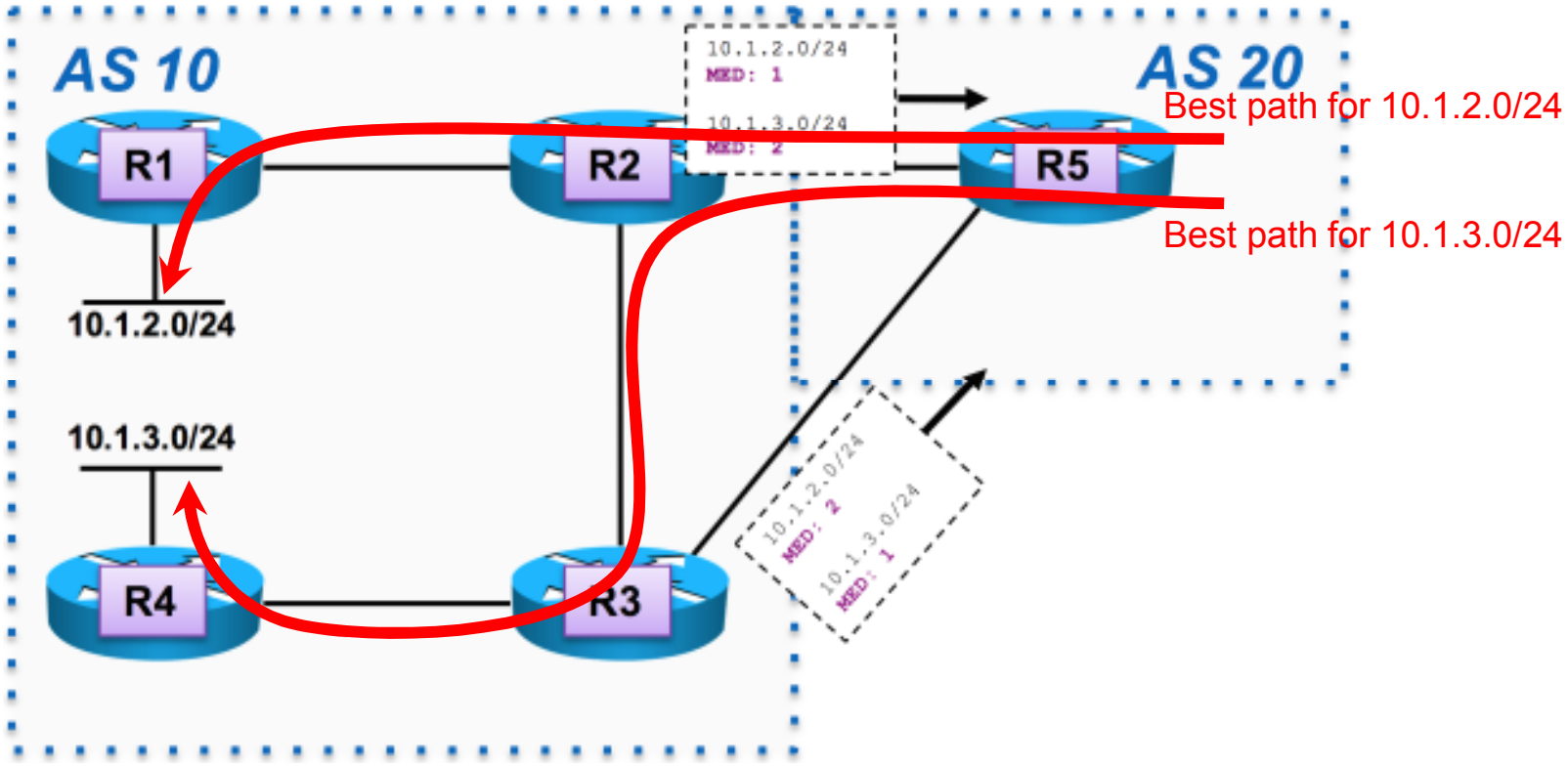
- AS 20 will now send traffic on both links
- Q: What are the problems with this policy ?



KDDI Myanmar

- Officially “Multi Exit Discriminator”
- An attribute used to
  influence inbound traffic
- Lower MED is better
  MED is designed to be a reflection of IGP metrics

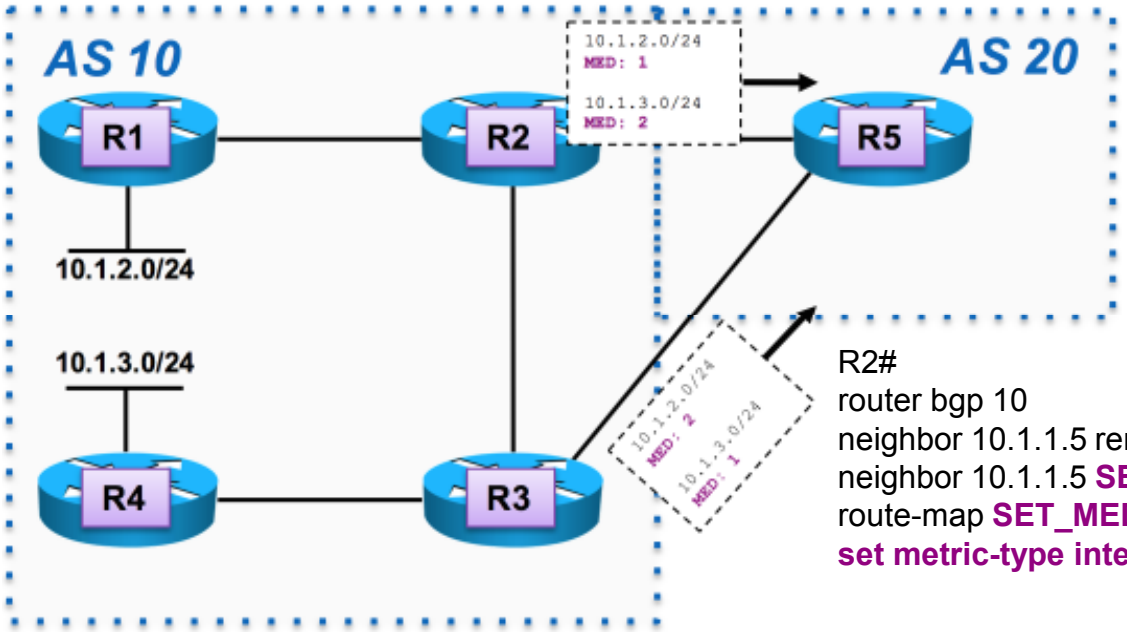
(*)Used to bring traffic into the AS on the eBGP speaker closest to the destination



MED

KDDI Myanmar

- MEDs can be set manually
- “set metric-type internal” sets MED dynamically
- Uses IGP cost to prefix as the MED value
- R2 has an IGP cost of 1 to 10.1.2.0
- R2 has an IGP cost of 2 to 10.1.3.0



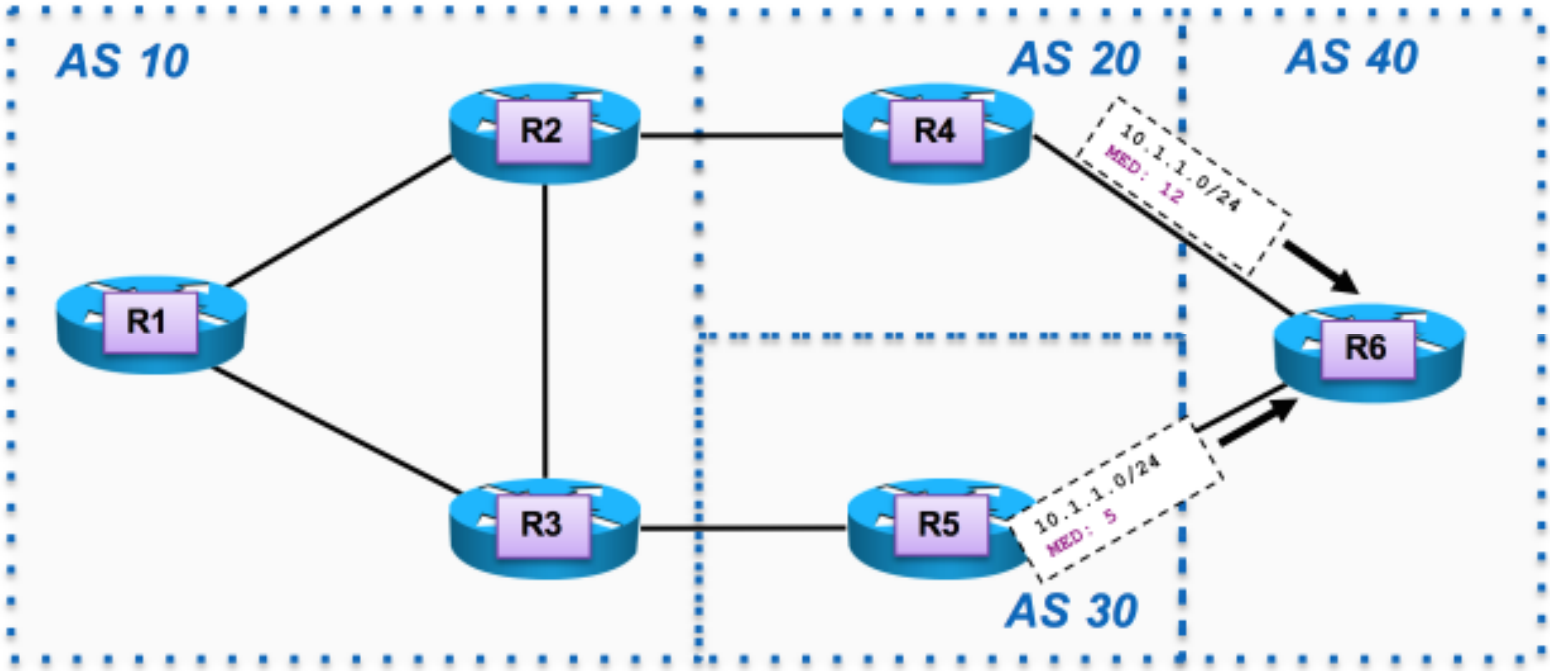
R2#
router bgp 10
neighbor 10.1.1.5 remote-as 20
neighbor 10.1.1.5 **SET_MED** out !
route-map **SET_MED** permit 10
set metric-type internal

- Traffic for 10.1.2.0/24 uses the R2 link
- Traffic for 10.1.3.0/24 uses the R3 link

MED - bgp always-compare-med

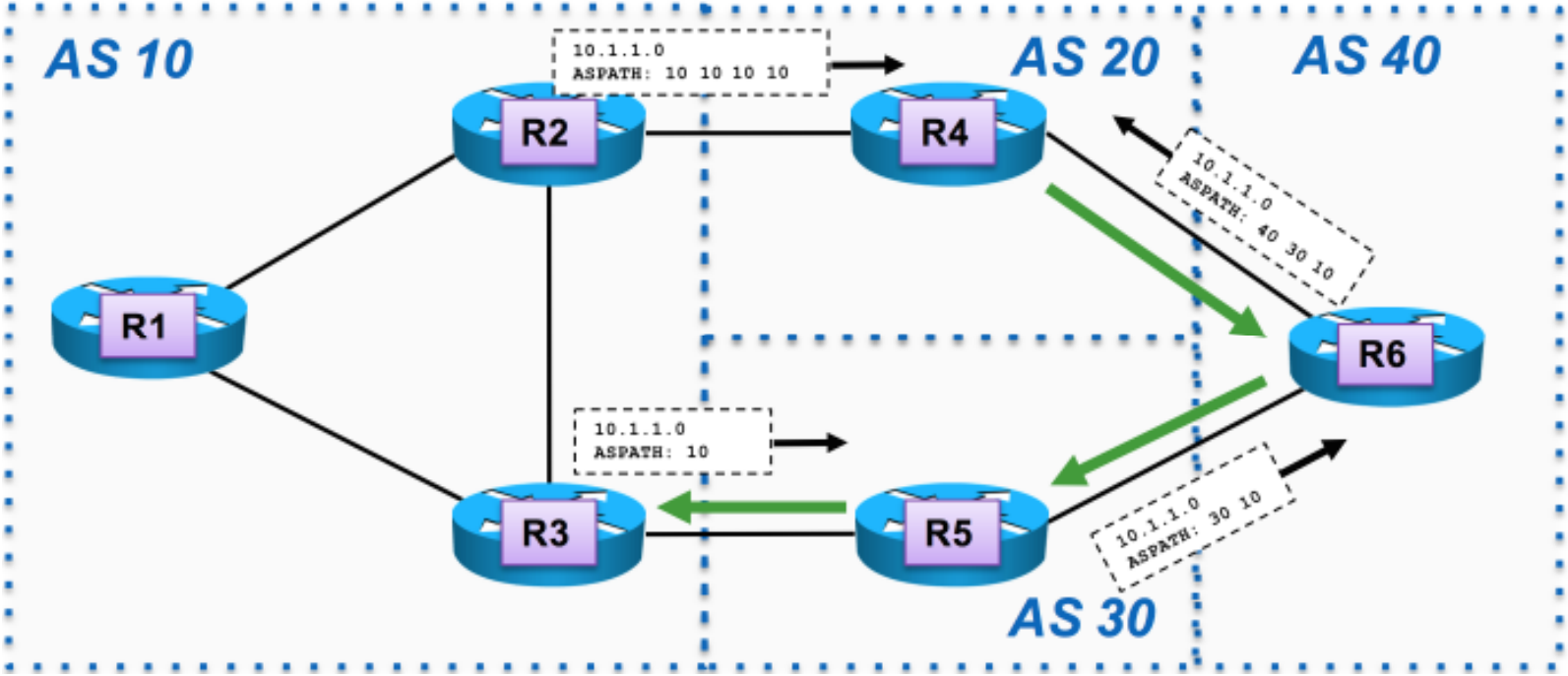
KDDI Myanmar

- MEDs are only compared if received from the same AS
- Makes sense as you can't necessarily compare routing policies across different AS
- R6 does not compare MEDs for the paths received from AS20 and AS30 unless "bgp always-compare-med" is configured



AS-PATH Prepend

- AS 10 can force traffic into R3 by prepending from R2 -> R4
- A shorter ASPATH is preferred



router bgp 10
neighbor 10.1.1.4 remote-as 20
neighbor 10.1.1.4 route-map **PREPEND_3X** out !
route-map **PREPEND_3X** permit 10 **set as-path prepend 10 10 10**

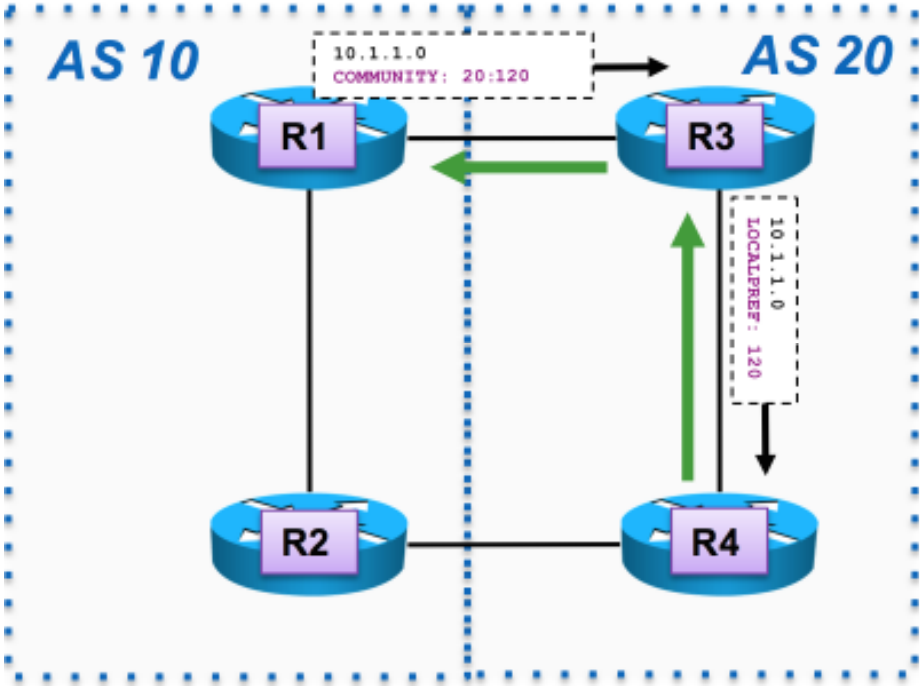
Community/Local Pref Agreement

KDDI Myanmar

- Many providers accept communities from their customers to give customers some control on inbound traffic.
- Customer sends community 20:80, ISP sets the LOCALPREF to 80
- Customer sends community 20:120, ISP sets the LOCALPREF to 120

```
R1#
router bgp 10
neighbor 10.1.1.3 remote-as 20
neighbor 10.1.1.3 route-map SET_COMMUNITY out
neighbor 10.1.1.3 send-community
!
route-map SET_COMMUNITY permit 10
set community 20:120
```

```
R3#
router bgp 20
neighbor 10.1.1.1 remote-as 10
neighbor 10.1.1.1 route-map COMMUNITY_TO_LOCALPREF in
ip community-list standard LP_80 permit 20:80
ip community-list standard LP_120 permit 20:120
route-map COMMUNITY_TO_LOCALPREF permit 10
                match community LP_80
                set local-preference 80
route-map COMMUNITY_TO_LOCALPREF permit 20
                match community LP_120
                set local-preference 120
route-map COMMUNITY_TO_LOCALPREF permit 30
```



BGP Database

KDDI Myanmar

- Neighbor table
  - List of BGP neighbors
  - **-show ip bgp neighbor**
- BGP table (forwarding database)
  - List of all networks learned from each neighbor
  - Can contain multiple paths to destination networks
  - Contains BGP attributes for each path
  - **-show ip bgp**
- IP routing table
  - List of best paths to destination networks
  - **-show ip route**

BGP command (example)

KDDI Myanmar

```
RouterA# show ip bgp summary
BGP router identifier 10.1.1.1, local AS number 65001
BGP table version is 124, main routing table version 124
9 network entries using 1053 bytes of memory
22 path entries using 1144 bytes of memory
12/5 BGP path/bestpath attribute entries using 1488 bytes of memory
6 BGP AS-PATH entries using 144 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 3829 total bytes of memory
BGP activity 58/49 prefixes, 72/50 paths, scan interval 60 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd

10.1.0.2        4 65001      11      11      124    0    0 00:02:28        8
172.31.1.3      4 64998      21      18      124    0    0 00:01:13        6
172.31.11.4     4 64999      11      10      124    0    0 00:01:11        6
```

BGP command (example)

KDDI Myanmar

```
RouterA# show ip bgp
BGP table version is 14, local router ID is 172.31.11.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal, r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop            Metric LocPrf Weight Path
*> 10.1.0.0/24      0.0.0.0                  0         32768 i
* i                 10.1.0.2                 0    100      0 i
*> 10.1.1.0/24      0.0.0.0                  0         32768 i
*>i10.1.2.0/24      10.1.0.2                 0    100      0 i
*> 10.97.97.0/24    172.31.1.3                             0 64998 64997 i
*                   172.31.11.4                            0 64999 64997 i
* i                 172.31.11.4              0    100      0 64999 64997 i
*> 10.254.0.0/24    172.31.1.3               0             0 64998 i
*                   172.31.11.4                            0 64999 64998 i
* i                 172.31.1.3               0    100      0 64998 i
r> 172.31.1.0/24    172.31.1.3               0             0 64998 i
r                   172.31.11.4                            0 64999 64998 i
r i                 172.31.1.3               0    100      0 64998 i
*> 172.31.2.0/24    172.31.1.3               0             0 64998 i
<output omitted>
```

BGP command (example)

KDDI Myanmar

```
RouterA# sh ip bgp neighbors
BGP neighbor is 172.31.1.3,  remote AS 64998, external link
  BGP version 4, remote router ID 172.31.2.3
  BGP state = Established, up for 00:19:10
  Last read 00:00:10, last write 00:00:10, hold time is 180, keepalive
interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received(old & new)
    Address family IPv4 Unicast: advertised and received
  Message statistics:
    InQ depth is 0
    OutQ depth is 0
                         Sent       Rcvd
    Opens:                  7          7
    Notifications:          0          0
    Updates:               13         38
<output omitted>
```

KDDI Myanmar

When establishing a BGP session, BGP goes through the following steps:

- Idle: Router is searching routing table to see if a route exists to reach the neighbor.
- Connect: Router found a route to the neighbor and has completed the three-way TCP handshake.
- Open sent: Open message sent, with the parameters for the BGP session.
- Open confirm: Router received agreement on the parameters for establishing session.

  Alternatively, router goes into Active state if no response to open message
- Established: Peering is established; routing begins.

KDDI Myanmar

- Active: The router has sent out an open packet and is waiting for a response.
- The state may cycle between active and idle. The neighbor may not know how to get back to this router because of the following reasons:
  1. Neighbor does not have a route to the source IP address of the BGP open packet generated by this router
  2. Neighbor peering with the wrong address
  3. Neighbor does not have a **`neighbor`** statement for this router
  4. AS number misconfiguration

```
(*)you can also check ;
```

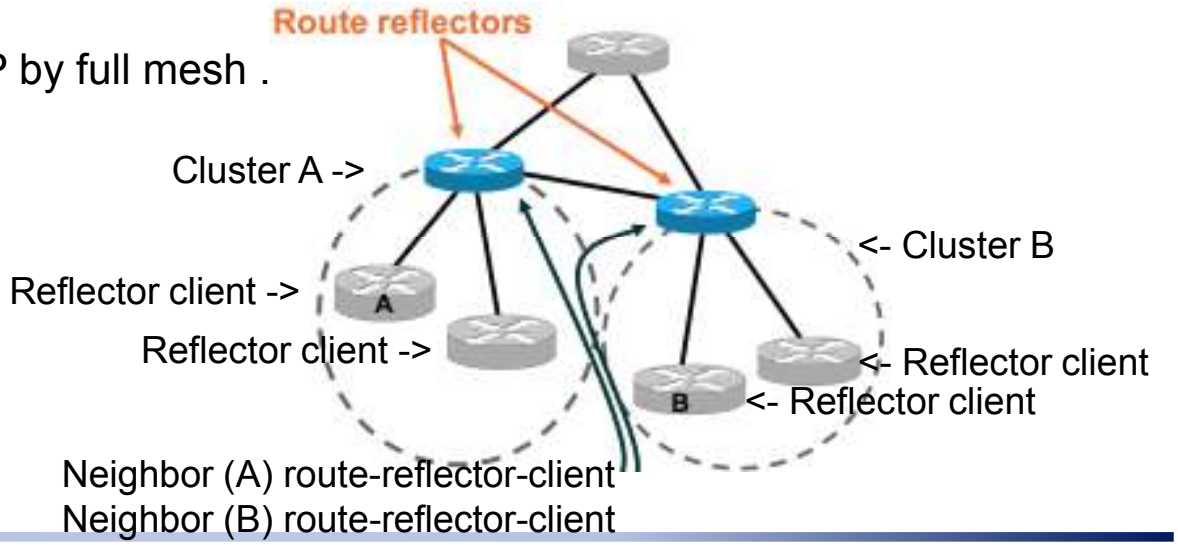
http://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/22166-bgp-trouble-main.html

What is a route Reflectors ?
Do we need it ?
Answer is …

Route Reflector Basics

KDDI Myanmar

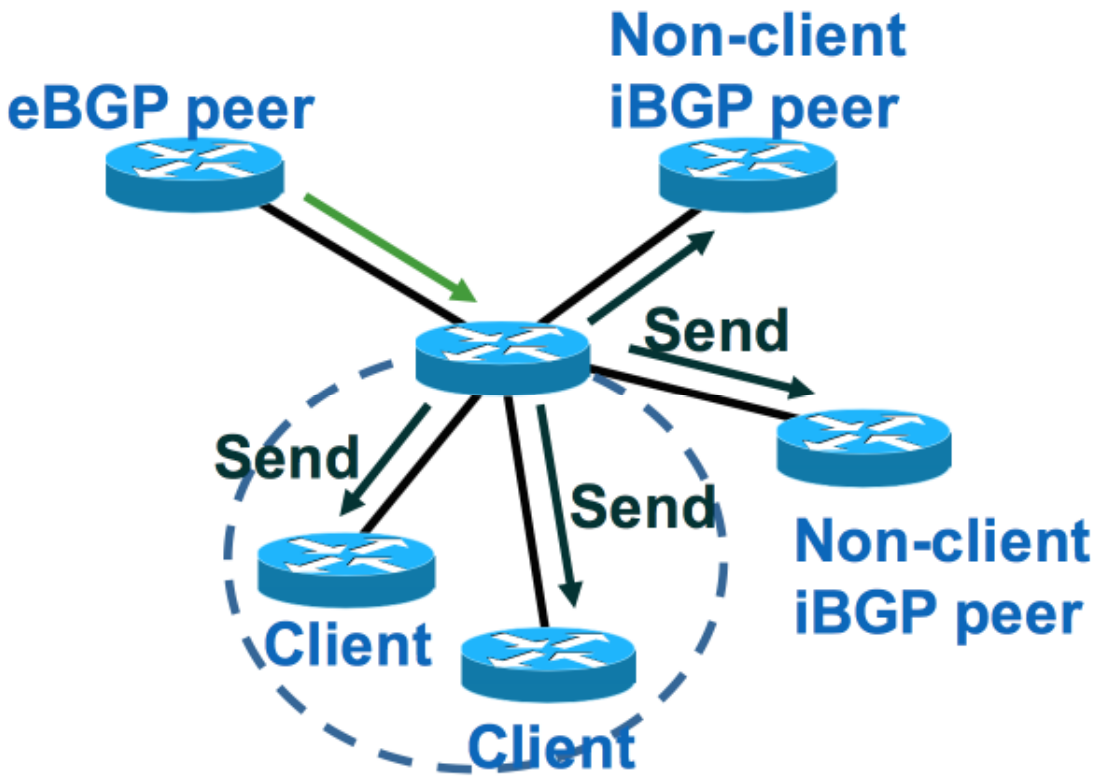
- BGP cannot advertise a path from one iBGP to another.
- eBGP uses ASPATH for loop detection
- Loop detection in iBGP is more difficult ☹
- RFC 2796 defines two attributes for loop detection within an AS:
- Originator ID :
  Set to the Router ID of the client who injects the route into the AS
- Cluster List :
  Each route reflector the route passes through adds their Cluster-ID to this list.
  Cluster-ID = Router ID by default
- A **route reflector** is an iBGP speaker that reflects routes learned from iBGP peers to other iBGP peers
- Route reflectors are designated by configuring some of their iBGP peers as route reflector clients
- Route reflector must have iBGP by full mesh .



Route Reflector - Advertisement Rules

KDDI Myanmar

If a Route Reflector Receives a Route from an eBGP Peer
- Send the route to all clients
- Send the route to all non-clients



Route Reflector - Advertisement Rules

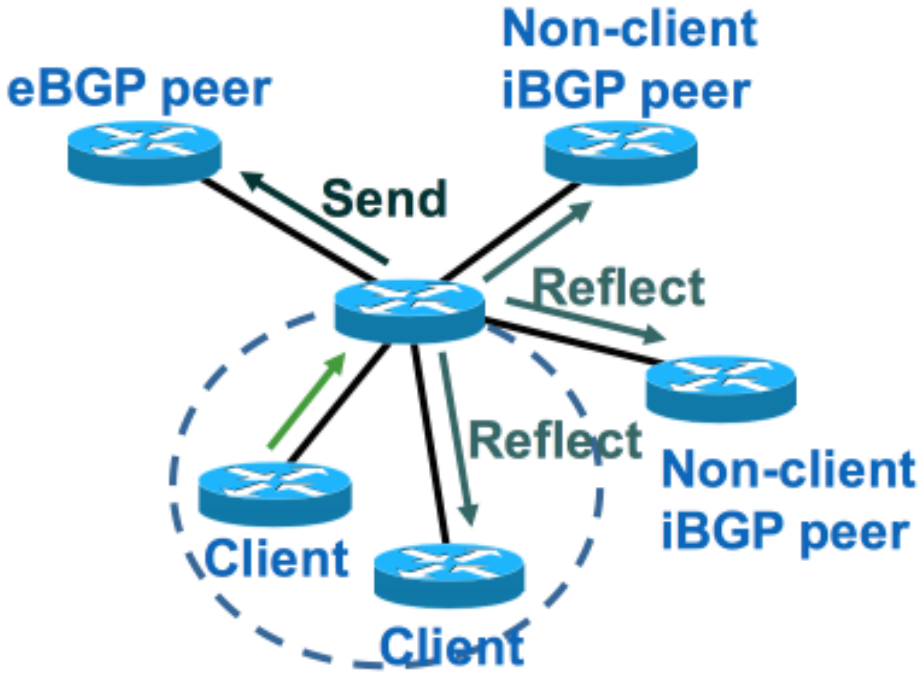
KDDI Myanmar

If a Route Reflector Receives a Route from a Client

- Reflect the route to all clients
  - Unless “no client-to-client reflection”, which is rarely deployed

- Reflect the route to all non-clients
  - Send the route to all eBGP peers

Even route reflector receives a route from non client , behaiver is same ;
  - send the route to all eBGP peer , client ,expect non-client

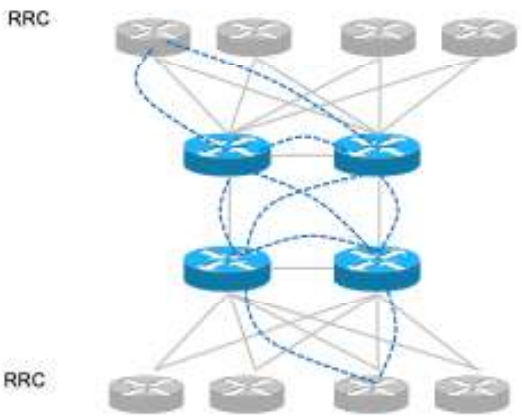


Route Reflector – redundancy (topoligy i.g.)

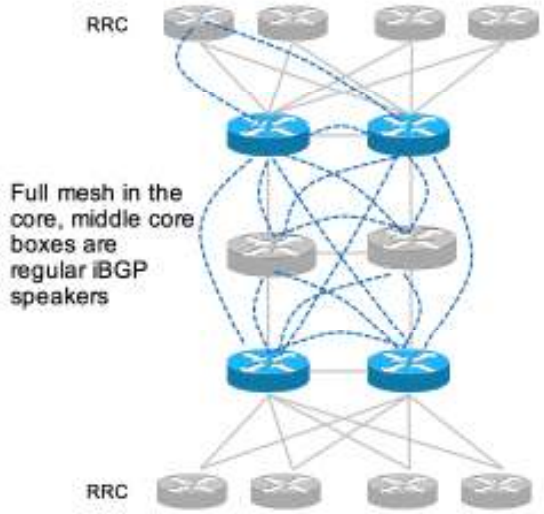
KDDI Myanmar

- Redundancy is needed but....
- Too much burns memory on RRCs because the client learns the same
- Also burns memory on the RRs because they learn multiple paths for each route
- **Two route reflectors per client should be plenty...**
- As with everything else...”it depends”
  – PEs, RRs, SLAs, network size, network topology, etc.

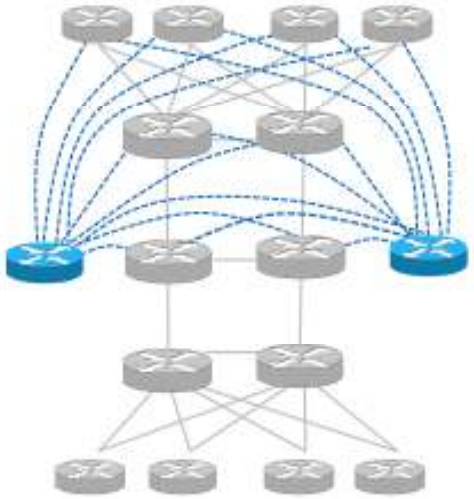
Distribution Routers as RR



Scaling the Core



RR Deployment in MPLS



BGP Convergence

KDDI Myanmar

Most dangerous period … actually
(if keep a high cpu ,network will be died …)

BGP Convergence

KDDI Myanmar

Two general convergence situations
– Initial startup
– Reaction to network failure events

Initial convergence happens when:
– A router boots
– RP failover
– clear ip bgp *

How long initial convergence takes is a factor of the amount of work to be done and the router/network's ability to do this fast and efficiently

Initial convergence can be stressful...if you are approaching BGP scalability limits this is when you will see issues with high cpu , memory .

-> your network is ok ?
-> especially ……

KDDI Myanmar

What work needs to be done?

1) Accept routes from all peers
   Not too difficult

2) Calculate bestpaths
   This is easy

3) Install bestpaths in the RIB
   Also easy

4) Advertise bestpaths to all peers
   This can be difficult and may take several minutes

KDDI Myanmar

BGP Variables
– The number of routes
– The number of peers
– The number of update-groups
– The ability to advertise routes to each update-group efficiently

Router Variables
– CPU horsepower
– Code version
– Outbound Interface Bandwidth

-> your network is ok ?
-> especially …

BGP Convergence – Update packing

KDDI Myanmar

- UPDATE contains a set of Attributes and a list of prefixes (NLRI)

- BGP starts an UPDATE by building an attribute set

- BGP then packs as many destinations (NLRIs) as it can into the UPDATE

- Only NLRI with a matching attribute set can be placed in the UPDATE

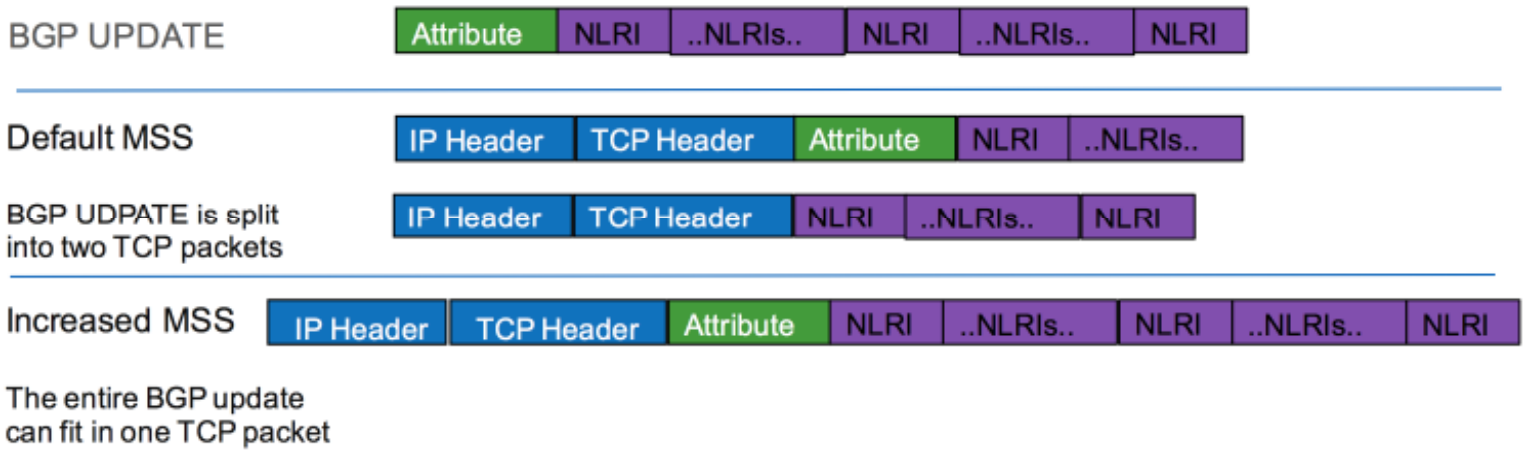
- NLRI are added to the UPDATE until it is full (4096 bytes max)

- The fewer attribute sets you have the better
  - -More NLRI will share an attribute set
  - -Fewer UPDATEs to converge
- Things you can do to reduce attribute sets
  - -next-hop-self for all iBGP sessions
  - -Don’t accept/send communities you don’t need
- To see how many attribute sets you have
  - -show ip bgp summary

BGP Convergence – TCP MSS

KDDI Myanmar

TCP MSS (max segment size) is also a factor in convergence times. The larger the MSS the fewer TCP packets it takes to transport the BGP updates. Fewer packets means less overhead and faster convergence.



(*) NLRI : Network Layer Reachability information

KDDI Myanmar

MSS – Max Segment Size
- Limit on packet size for a TCP socket
- 536 bytes by default

Path MTU Discovery
- Finds smallest MTU between R and R
- Subtract 40 bytes for TCP/IP overhead
- Enabled by default for BGP (at least in recent releases)
- In older releases enable via global cmd “ip tcp path-mtu-discovery”

To find the MSS

sh ip bgp neighbors

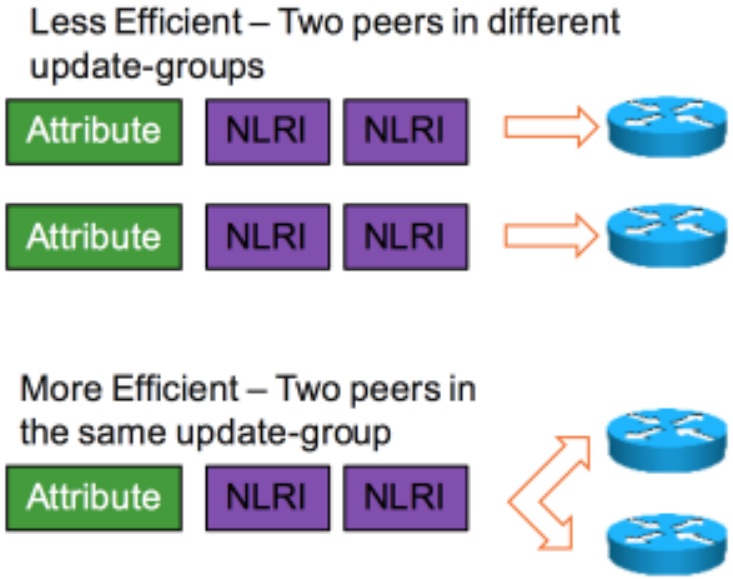
BGP neighbor is 2.2.2.2, remote AS 3, external link Datagrams (max data segment is 1460 bytes):

KDDI Myanmar

- BGP must create updates based on the policies towards each peer
- Peers with a common outbound policy are members of the same update-group
  - iBGP vs. eBGP
  - Outbound route-map, prefix-lists, etc

- UPDATEs are generated for one member of an update-group and then replicated to the other members

- Back in the old days, these “update-groups” had to be created specifically, using “peer-groups”. They’re still widely deployed...



Service efficiency …

IGP vs. BGP Convergence

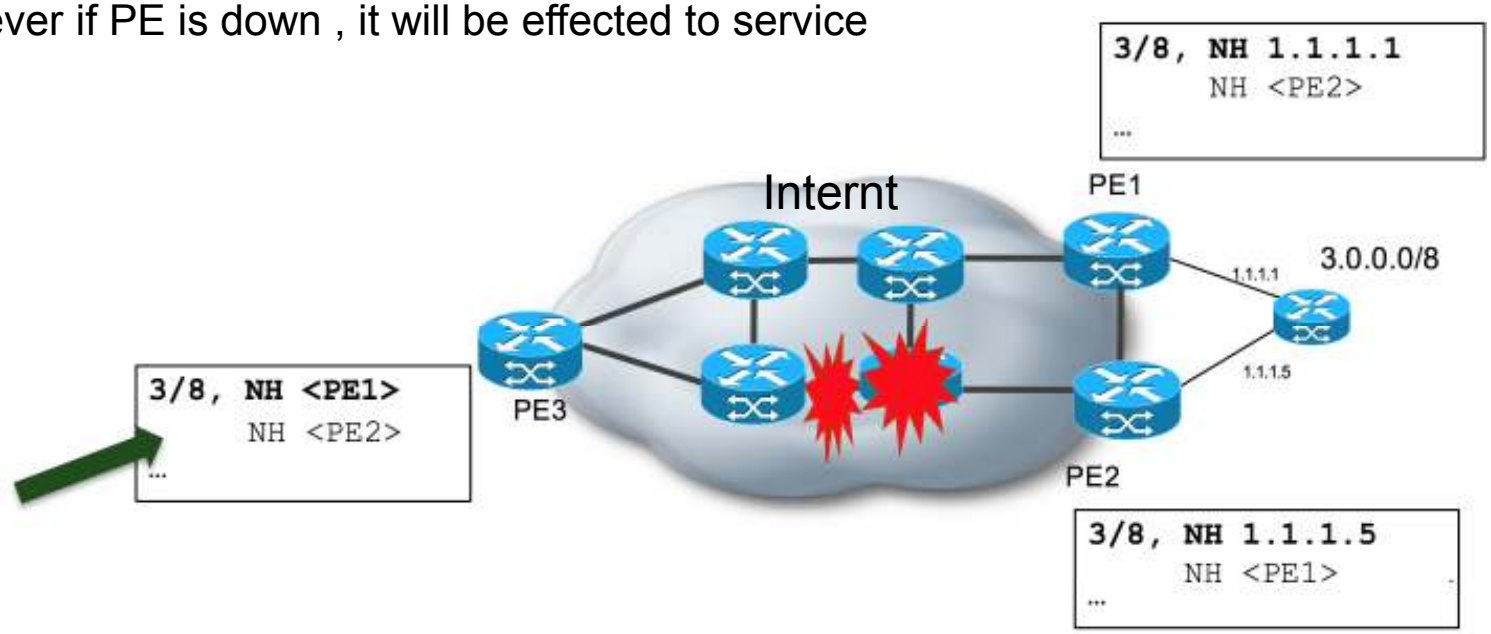
KDDI Myanmar

- IGP (OSPF/ISIS) deals with hundreds routes
  - Max a few thousands, but only a few hundreds are really important/relevant
- EGP is designed to carry millions of routes
  - and a few large customers carry that amount of prefixes!
- We can tune IGPs to converge in << 1 second
- How about BGP ? , it is depend on No# of route information and No# of peers

BGP Control-Plane Convergence Components ;

1. Internet Link or node goes down
2. IGP notices failure, computes new paths to PE1/PE2
3. IGP notifies BGP that a path to a next-hop has changed
4. PE3 identifies affected paths, runs best path, path to PE2 no longer as good as the one to PE1
5. Updates RIB/FIB, traffic continues

(*) however if PE is down , it will be effected to service



Failure Detection

KDDI Myanmar

- Problem: Detect an eBGP neighbour failure

Available Methods

– Fast External Fallover – monitors line protocol for directly connected neighbours (default behaviour)

– Fast Session Deactivation (FSD), monitors routing table for reachability of next-hop address (eBGP multi-hop)

– “Hello”-type protocols: BFD and BGP Hello

```
router bgp …
 [no] bgp fast-external-fallover
interface …
 ip bgp fast-external-fallover {permit|deny}


router bgp …
 neighbor x.x.x.x fall-over


router bgp …
 timers bgp <hello> <hold>
 neighbor <..> timers <hello> <hold>


 neighbor <..> fall-over bfd
```

BGP Next-hop Tracking (NHT)

KDDI Myanmar

- BGP Next Hop Tracking

- Enabled by default
  [no] bgp nexthop trigger enable

- Default trigger delay: 5 seconds
  bgp nexthop trigger delay <seconds>

- BGP registers all nexthops with Address Tracking Feature (ATF)
  Hidden command will let you see a list of nexthops
  show ip bgp attr nexthop

- ATF will let BGP know when a route change occurs for a nexthop

- ATF notification will trigger a lightweight “BGP Scanner” run
  – Bestpaths will be calculated
  – None of the other “Full Scan” work will happen

BGP convergence summary

KDDI Myanmar

- BGP Control Plane convergence can achieve fast convergence for small-ish deployments only (a few 1000s of prefixes)

- Achieving fast convergence for large deployments (including full Internet routes) requires

- BGP Convergence always require the underlying IGP (OSPF, ISIS, EIGRP) to be tuned for fast convergence

HA (High Availability)

How can we get a HA ? How to design a network ?

What is routing HA ?

KDDI Myanmar

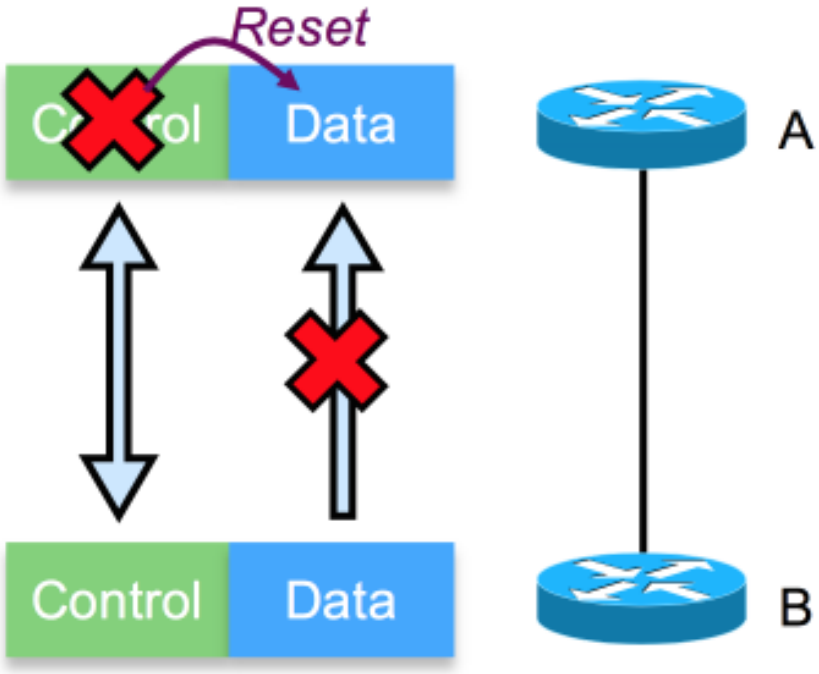
Routing HA

- Set of technologies & features to enable traffic to continue to flow **through** a device during a fault
- Routing HA **maintains** the logical network topology while the faulty device **recovers**
- Routing HA helps to address failures within the **control plane** of a routing device
- Routing HA increases the **resiliency** of a single **system**

Withhout Non-Stop Forwarding (NSF)

KDDI Myanmar

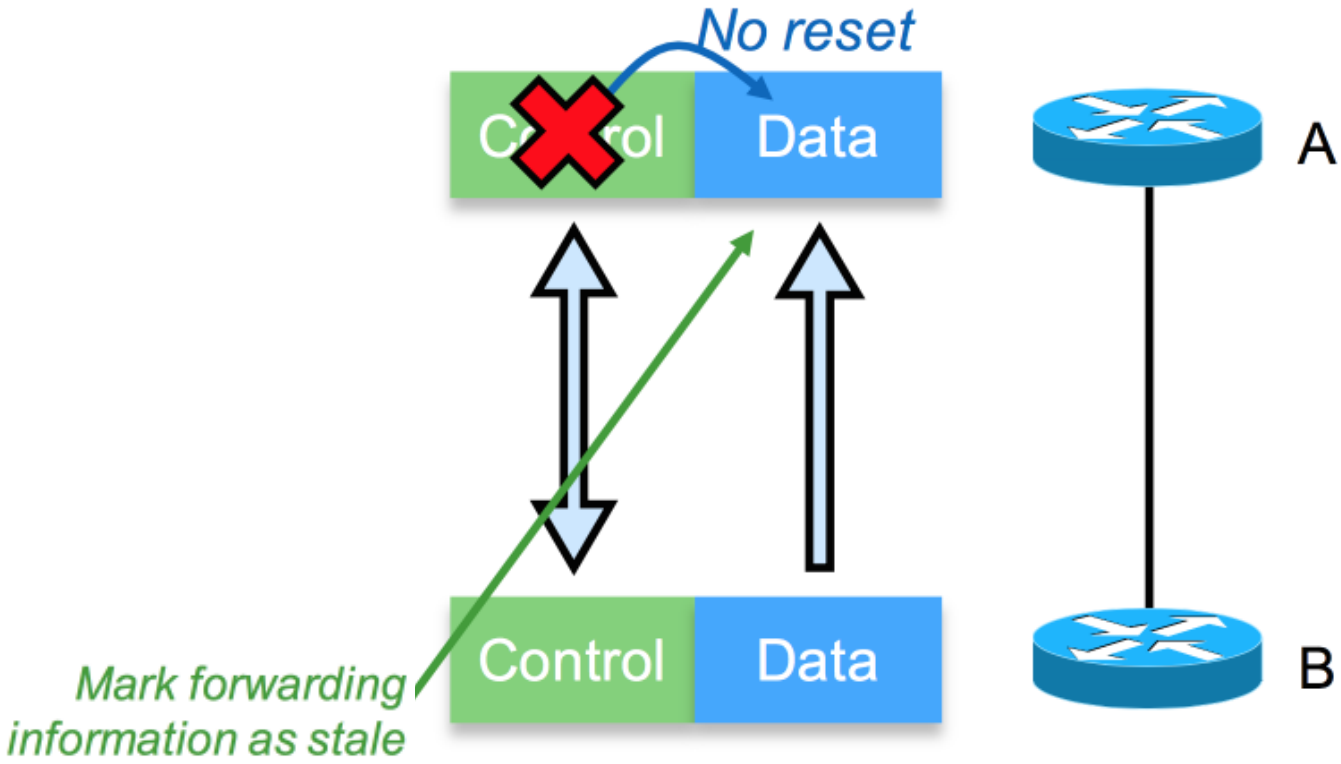
- Router A loses its control plane for some period of time
- It will take some time for Router B to recognize this failure, and react to it
- During the time that A has failed, and B has not detected the failure, B will continue forwarding traffic through A
- Once the control plane resets, the data plane will reset as well, and this traffic will be dropped
- NSF reduces or eliminates the traffic dropped while A's control plane is down



Graceful-Restart/NSF

KDDI Myanmar

- If A is NSF capable, the control plane will not reset the data plane when it restart
- Instead, the forwarding information in the data plane is marked as stale
- Any traffic B sends to A will still be switched based on the last known forwarding information
- This is the *Non-Stop Forwarding* behaviour



Non-stop Routing - NSR

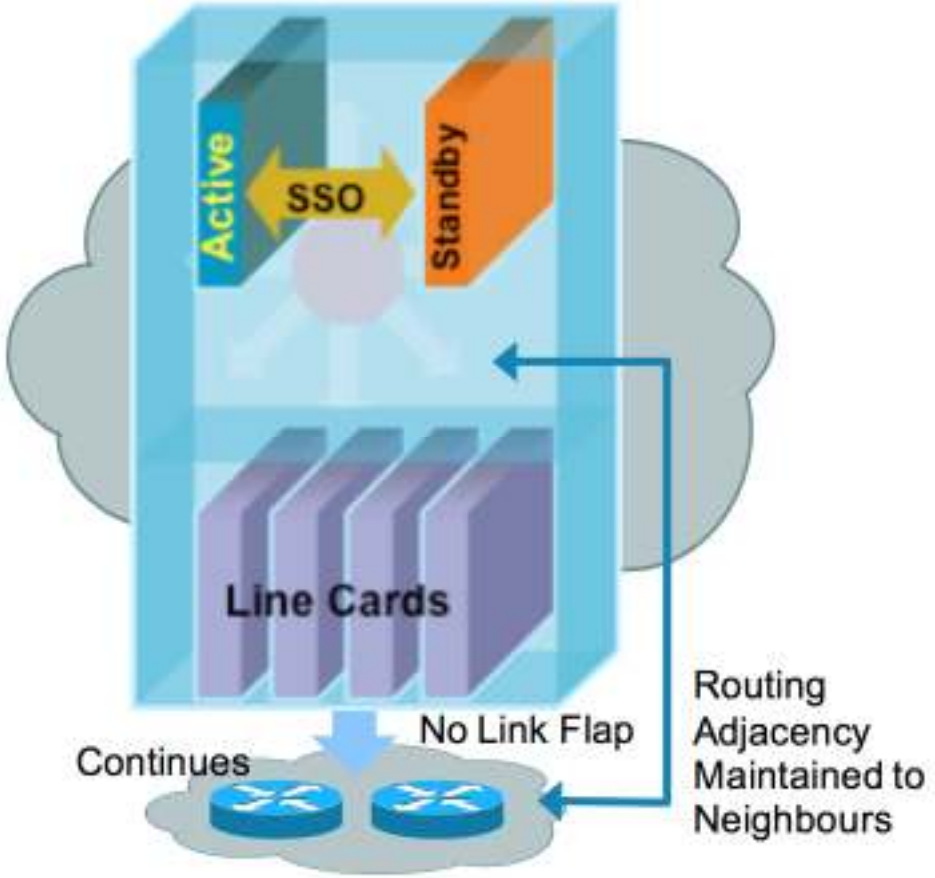
KDDI Myanmar

- Restarting RP could pick up right where the primary left off
- No need to refresh any information, no need for the neighbour to know that anything happened
- Absolutely need to avoid anything to let the neighbour know

- BGP RIB, TCP Session state is synced to standby RP/process
- Router can restart without neighbour's interaction
- No upgrade/changes on peers required

```
router bgp …
 bgp graceful-restart
 address-family ipv4 vrf ..
  neighbor x.x.x.x ha-mode sso
 ....
```

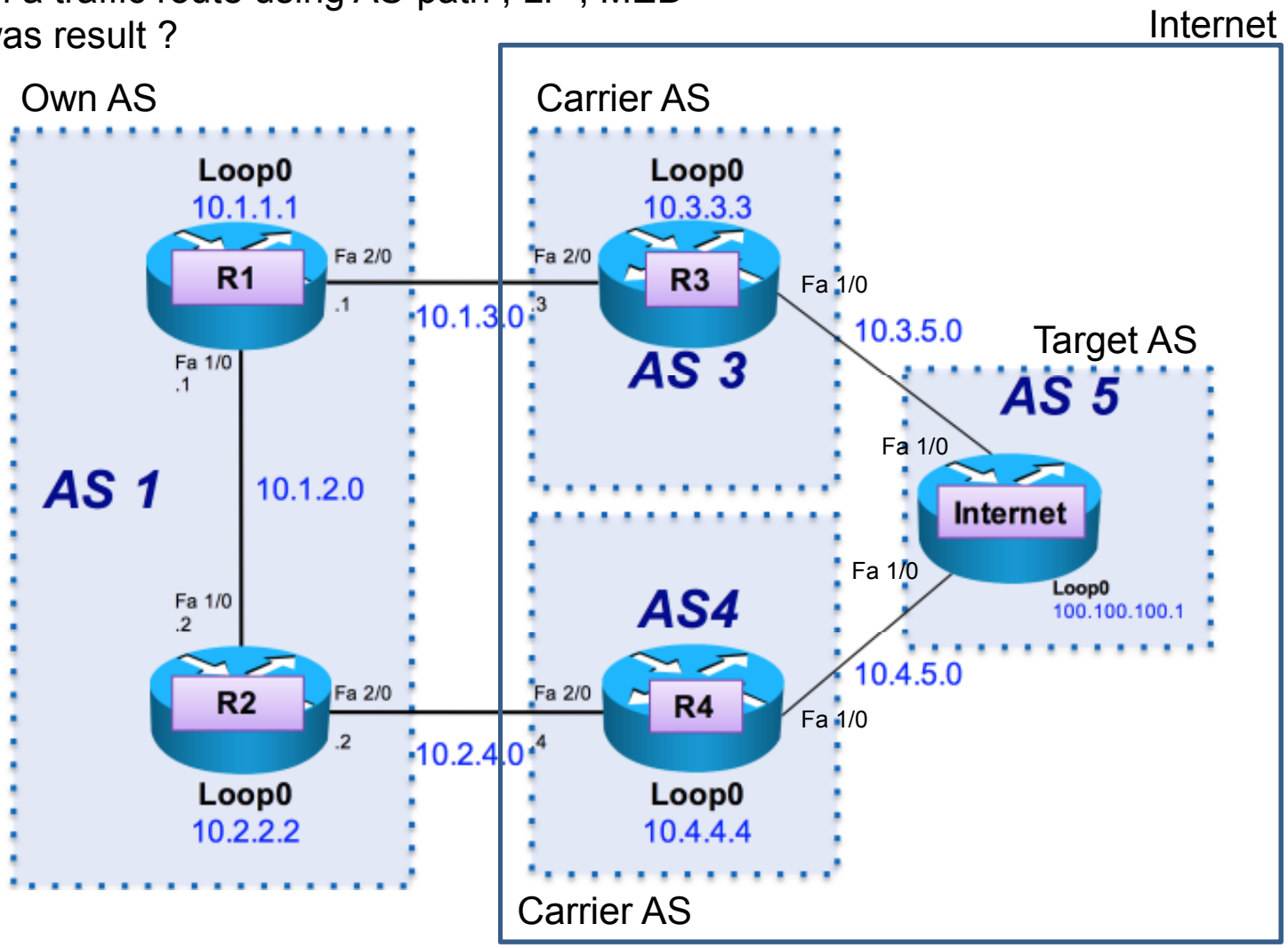
```
# show ip bgp vpnv4 all sso summary
# show tcp ha connections
```



Lets try a configuration

Lets configure a following network ;

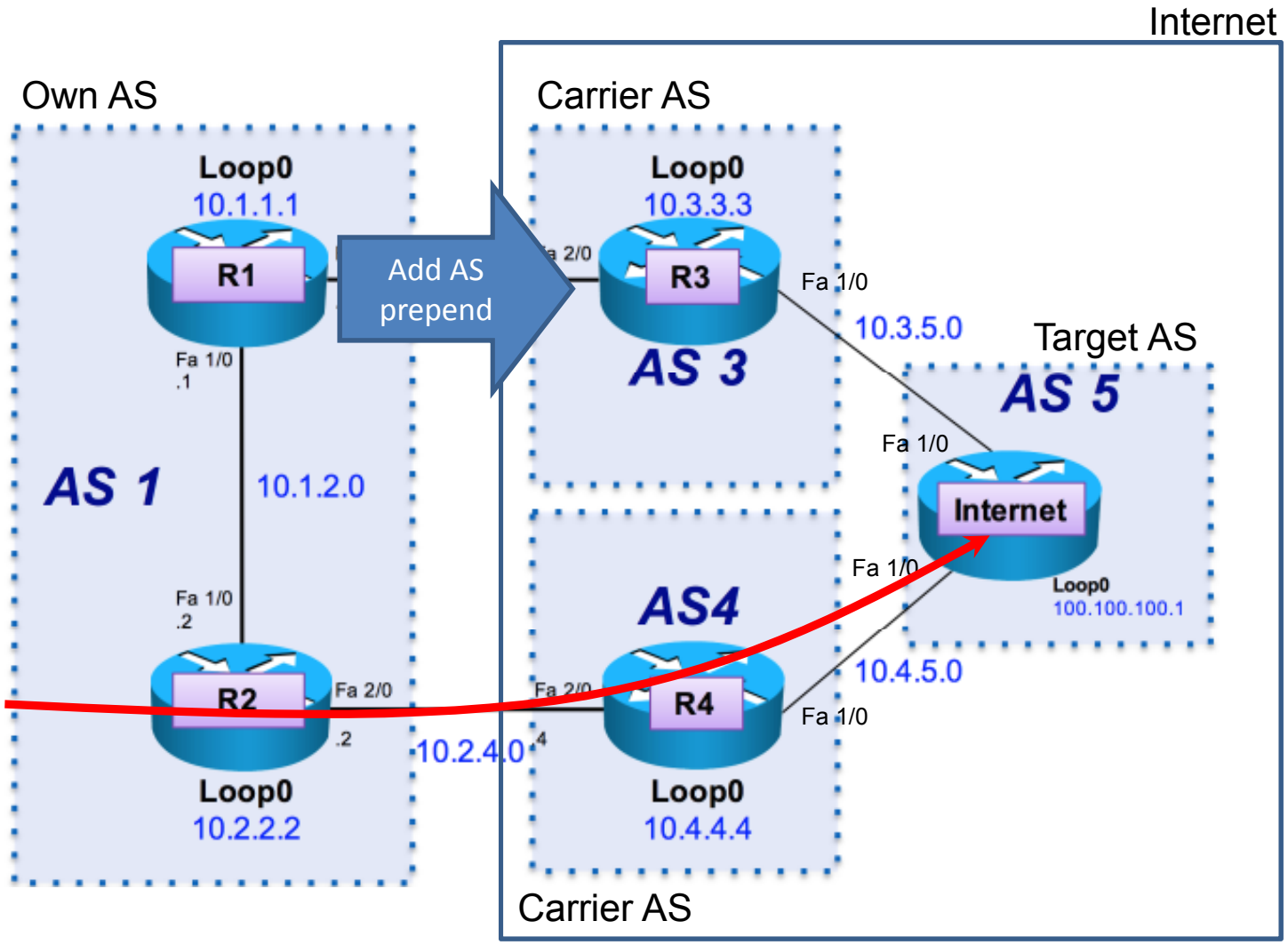
- Build this network first
- Control a traffic route using AS-path , LP , MED
- How was result ?



Case study #1 : Control by AS-path prepend

KDDI Myanmar

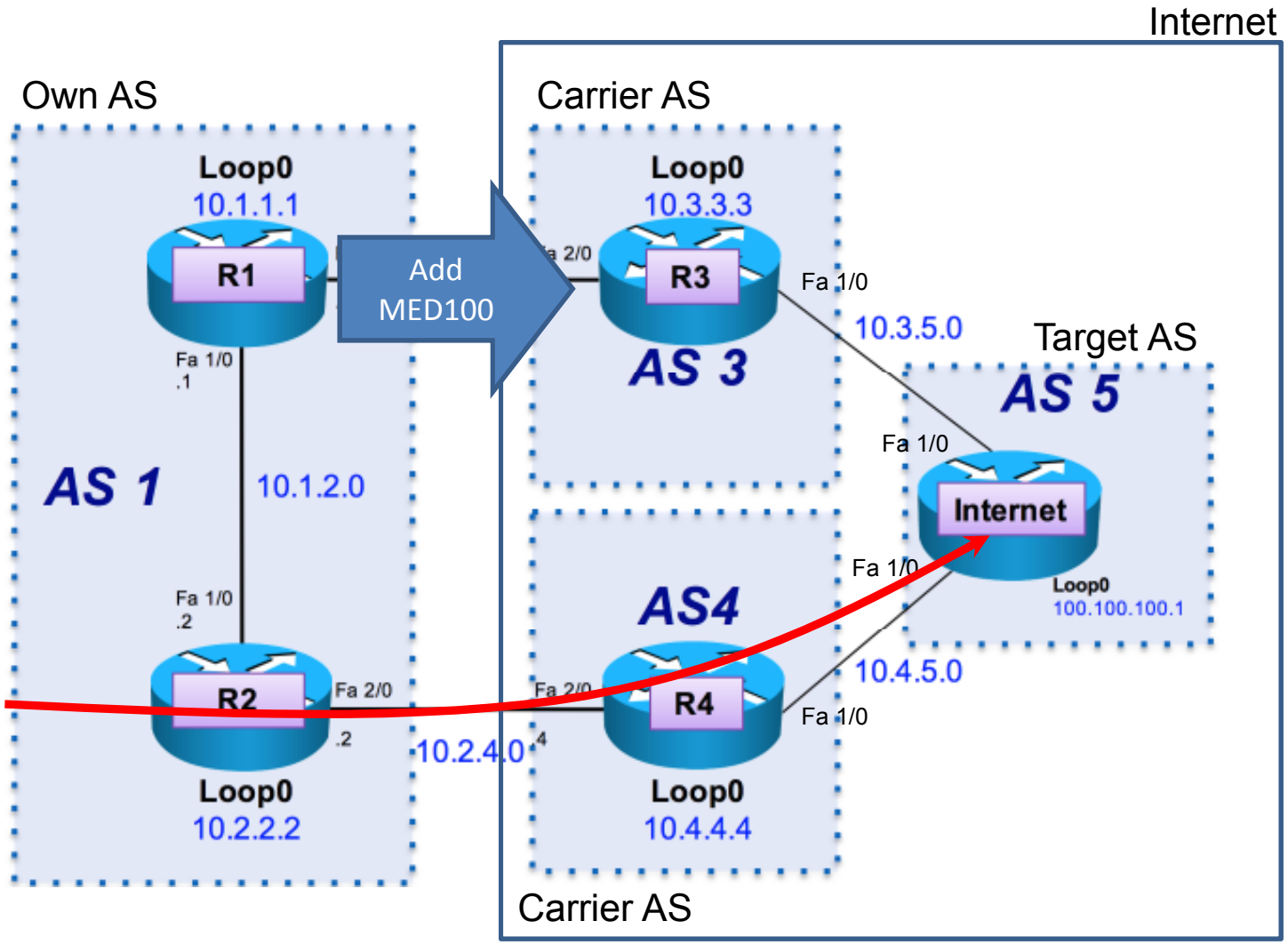
- Control route between R2 and R4 is higher than between R1 and R3 (by AS-path prepend)



Case study #2 : Control by MED

KDDI Myanmar

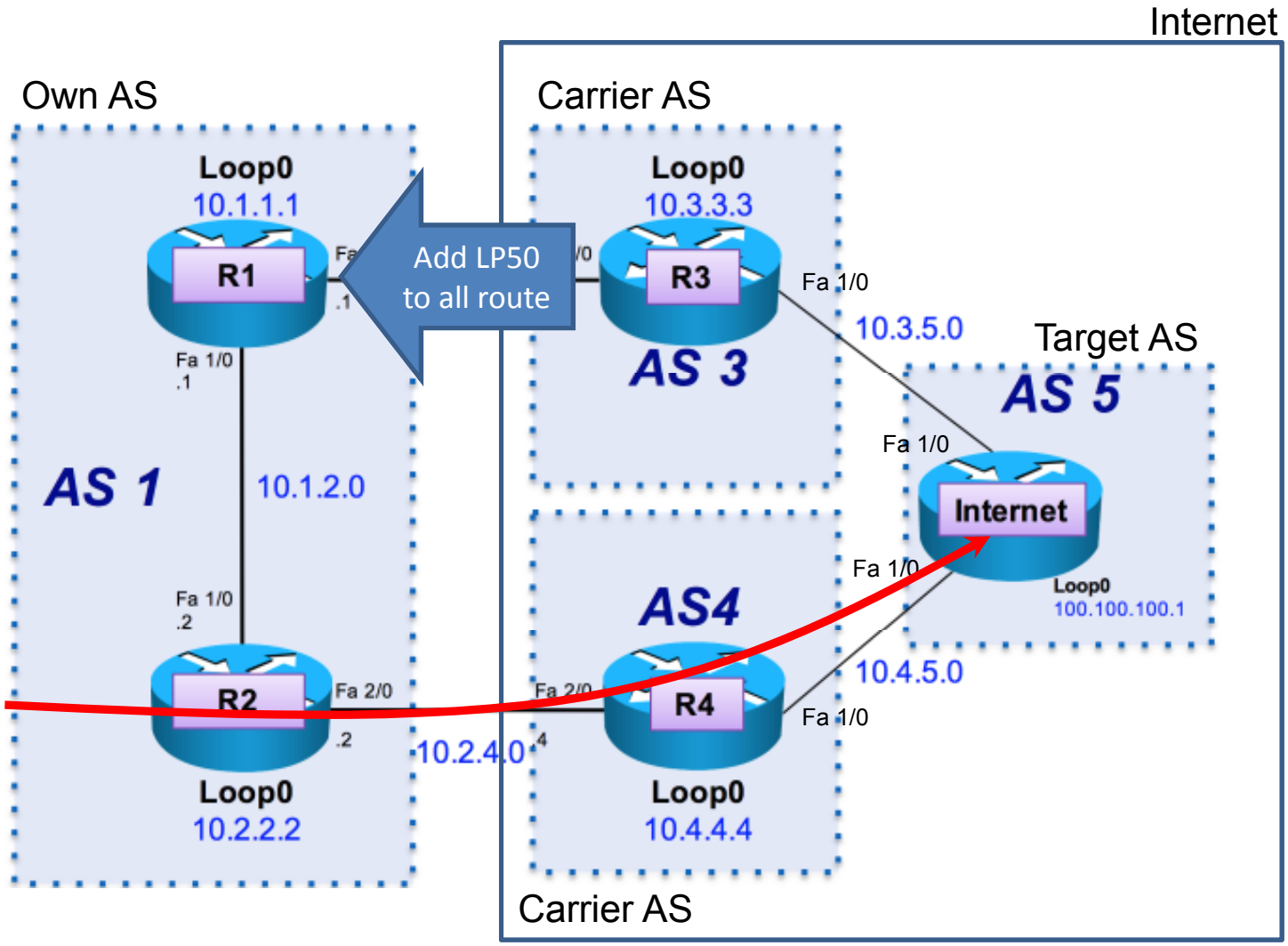
- Add MED 100 from AS#1 to AS3



Case study #3 : Control by LP (local preference)

KDDI Myanmar

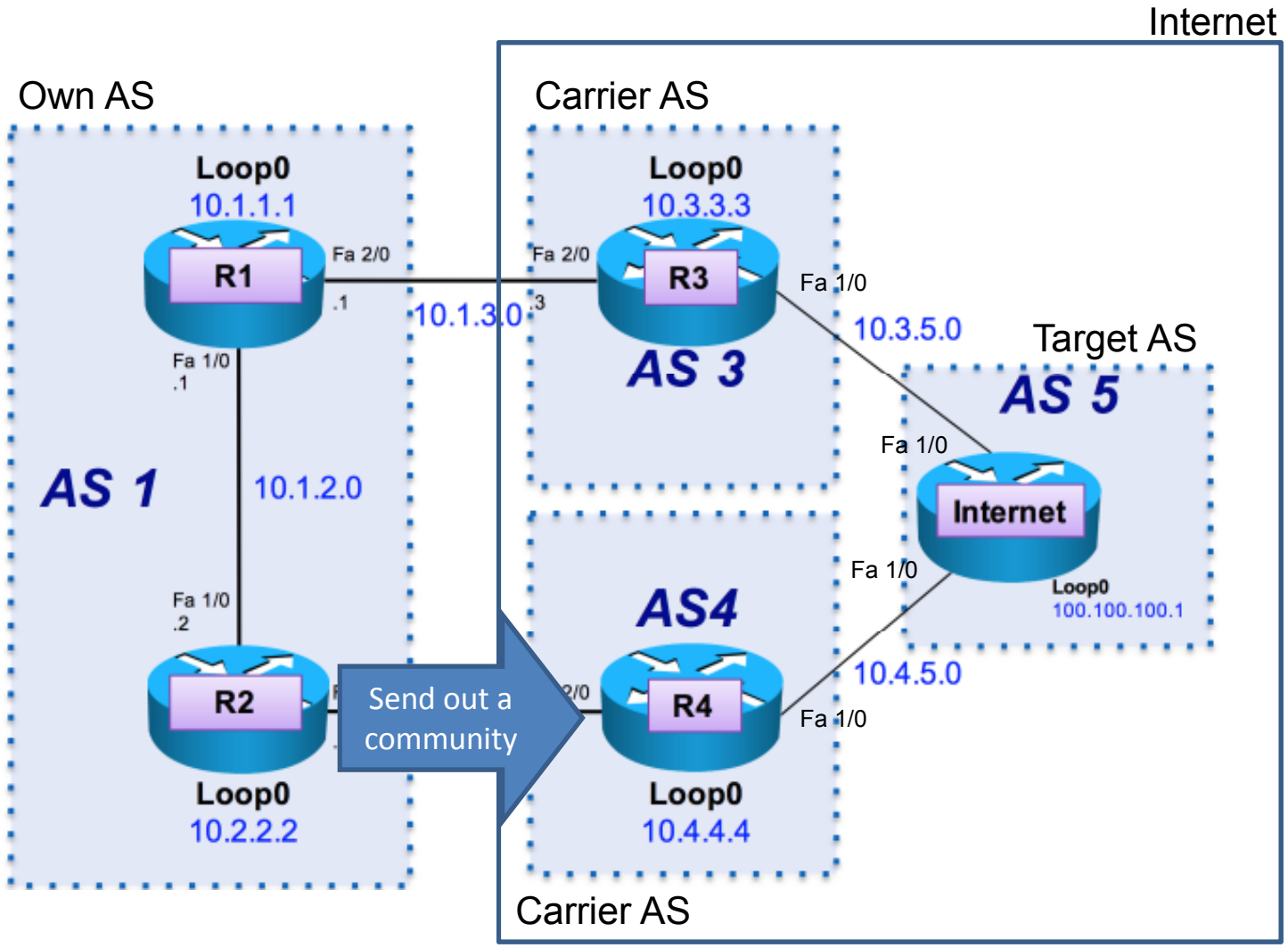
- Add LP 50 to all route which is received from AS3



Case study #3 : Control by Community

KDDI Myanmar

- Add community 1:1 and send out to AS4



KDDI Myanmar

XR basic commands

KDDI Myanmar

System redundancy

- Power redundancy
- Fabric redundancy
- RP redundancy
- NSR
- NSF
- Routing process placement and failure protection.

```
RP/0/RP0/CPU0:CRS1#show redundancy
Wed Apr 15 06:18:20.400 UTC
Redundancy information for node 0/RP0/CPU0:
==========================================
Node 0/RP0/CPU0 is in ACTIVE role
Partner node (0/RP1/CPU0) is in STANDBY role
Standby node in 0/RP1/CPU0 is ready
Standby node in 0/RP1/CPU0 is NSR-ready
```

KDDI Myanmar

Control plane redundancy

```
RP/0/RP1/CPU0:CRS1-MC#show redundancy summary
    Active Node      Standby Node
    -----------      ------------
     0/RP1/CPU0       0/RP0/CPU0 (Node Ready, NSR: Not Configured)
     1/RP0/CPU0       1/RP1/CPU0 (Node Ready, NSR: Not Configured)
     2/RP0/CPU0       2/RP1/CPU0 (Node Ready, NSR: Not Configured)
     3/RP0/CPU0       3/RP1/CPU0 (Node Ready, NSR: Not Configured)
     4/RP0/CPU0       4/RP1/CPU0 (Node Ready, NSR: Not Configured)
     5/RP0/CPU0       5/RP1/CPU0 (Node Ready, NSR: Not Configured)
     6/RP0/CPU0       6/RP1/CPU0 (Node Ready, NSR: Not Configured)
     7/RP1/CPU0       7/RP0/CPU0 (Ready, NSR: Ready)
RP/0/RP1/CPU0:CRS1-MC#
```

KDDI Myanmar

Basic Distributed LC programing

```
RP/0/8/CPU0:GSR1#show route 10.7.107.4
Routing entry for 10.7.107.4/30
  Known via "ospf 2", distance 110, metric 10301, type inter area
  Installed Apr 16 05:08:56.308 for 1d12h
  Routing Descriptor Blocks
    10.5.70.1, from 10.122.0.3, via TenGigE0/0/0/0
      Route metric is 10301
  No advertising protos.

RP/0/8/CPU0:GSR1#show cef 10.7.107.4
10.7.107.4/30, version 18, internal 0x4000001 (ptr 0xaec2d2b8) [1], 0x0 (0xae2252f8), 0x0 (0x0)
 Updated Apr 16 05:08:56.311
 local adjacency 10.5.70.1
 Prefix Len 30, traffic index 0, precedence routine (0)
   via 10.5.70.1, TenGigE0/0/0/0, 8 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0
    next hop 10.5.70.1
    local adjacency

RP/0/8/CPU0:GSR1#show cef 10.7.107.4  location 0/1/cpu0
10.7.107.4/30, version 18, internal 0x4000001 (ptr 0x543e6c6c) [1], 0x0 (0x543b49f8), 0x0 (0x0)
 Updated Apr 16 05:09:29.038
 local adjacency 10.5.70.1
 Prefix Len 30, traffic index 0, precedence routine (0)
   via 10.5.70.1, TenGigE0/0/0/0, 8 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0
    next hop 10.5.70.1
    local adjacency
RP/0/8/CPU0:GSR1#
```

KDDI Myanmar

install add Command

Copy image to disk, verify, and unpack

```
RP/0/0/CPU0:GSR-XR(admin)#install add tftp://172.21.116.8/c12k-mcast.pie-4.2.1.3I

Install: The idle timeout on this line will be suspended for synchronous install operations
Install: Starting install operation. Do not insert or remove cards until the operation
completes.
RP/0/0/CPU0:P4(admin)#
Install: Now operating in asynchronous mode. Do not attempt subsequent install operations
until this operation is complete.
Install 3: [  0%] Install operation 'add /tftp://172.21.116.8/c12k-mcast.pie-4.2.1.3I to
disk0:' assigned request id: 3
Install 3: [  1%] Downloading PIE file from /tftp://172.21.116.8/c12k-mcast.pie-4.2.1.3I
Install 3: [  1%]     Transferred 3298994 Bytes
Install 3: [  1%] Downloaded the package to the router
Install 3: [  1%] Verifying the package
Install 3: [  1%] [OK]
Install 3: [  1%] Verification of the package successful [OK]
Install 3: [ 95%] Going ahead to install the package...
Install 3: [ 95%] Add of '/tftp://172.21.116.8/c12k-mcast.pie-4.2.1.3I' completed.
Install 3: [100%] Add successful.
Install 3: [100%] The following package(s) and/or SMU(s) are now available to be activated:
Install 3: [100%]     disk0:c12k-mcast-4.2.1
Install 3: [100%] Please carefully follow the instructions in the release notes when activating any software
Install 3: [100%] Idle timeout on this line will now be resumed for synchronous install operations
```

KDDI Myanmar

install activate Command

Begin executing new software

```
RP/0/0/CPU0:GSR-XR(admin)#install activate disk0:c12k-mcast-4.2.1
Install: The idle timeout on this line will be suspended for synchronous installoperations
Install: Starting install operation. Do not insert or remove cards until the operation...
RP/0/0/CPU0:P4(admin)#
Install: Now operating in asynchronous mode. Do not attempt subsequent install operations
until this operation is complete.
Install 3: [ 0%] Install operation 'activate disk0:c12k-mcast-4.2.1' assigned request id: 3
Install 3: [ 1%] Performing Inter-Package Card/Node/Scope Version Dependency Checks
Install 3: [ 1%] [OK]
Install 3: [ 1%] Checking API compatibility in software configurations...
Install 3: [ 1%] [OK]
Install 3: [ 10%] Updating software configurations.
Install 3: [ 10%] RP,DRP:
Install 3: [ 10%] Activating c12k-mcast-4.2.1
Install 3: [ 10%] Checking running configuration version compatibility with newly activated…
Install 3: [ 10%] No incompatibilities found between the activated software and router…configuration.
RP/0/0/CPU0:Nov 12 14:24:01.249 : instdir[181]: %INSTMGR-6-SOFTWARE_CHANGE_END :Software change
transaction 3 is COMPLETE.
Install 3: [100%] Performing software change
Install 3: [100%] Activation operation successful.
Install 3: [100%] NOTE: The changes made to software configurations will not be
Install 3: [100%] persistent across RP reloads. Use the command 'install commit'
Install 3: [100%] to make changes persistent.
Install 3: [100%] Idle timeout on this line will now be resumed for synchronous install operations
```

KDDI Myanmar

`install commit` Command

Lock in activated software across reload

```
RP/0/0/CPU0:GSR-XR(admin)#install commit
Install: The idle timeout on this line will be suspended for synchronous
install operations
Install 5: [  1%] Install operation 'commit' assigned request id: 5
Install 5: [100%] Committing uncommitted changes in software configurations.
Install 5: [100%] Commit operation successful.
Install 5: [100%] Idle timeout on this line will now be resumed for
synchronous operations
```

Deactivating Packages

```
RP/0/0/CPU0:GSR-XR(admin)#install deactivate disk0:c12k-rp-mgbl-4.2.1
Install: The idle timeout on this line will be suspended for synchronous install
operations
Install: Starting install operation. Do not insert or remove cards until the operation completes.
RP/0/0/CPU0:P5(admin)#
Install: Now operating in asynchronous mode. Do not attempt subsequent install operations
until this operation is complete.
Install 8: [ 0%] Install operation 'deactivate disk0:c12k-mgbl-4.2.1' assigned
request id: 8
Install 8: [ 1%] Package 'disk0:c12k-mgbl-4.2.1' is not active and cannot be deactivated.
Install 8: [ 1%] Idle timeout on this line will now be resumed for synchronous
install operations
```

Package features no longer available
Package still installed
Package can be reactivated

KDDI Myanmar

Identifying packages

```
RP/0/RP0/CPU0:CR1-CRS#show install active summary
Tue Apr  1 03:51:10.322 UTC
  Active Packages:
    disk0:hfr-mini-px-4.2.4
    disk0:hfr-doc-px-4.2.4
    disk0:hfr-k9sec-px-4.2.4
    disk0:hfr-mpls-px-4.2.4
    disk0:hfr-px-4.2.4.CSCue55783-1.0.0
    disk0:hfr-mgbl-px-4.2.4
    disk0:hfr-mcast-px-4.2.4
    disk0:hfr-fpd-px-4.2.4
    disk0:hfr-diags-px-4.2.4
```

KDDI Myanmar

Show version

```
RP/0/RP0/CPU0:CRS1# show version brief
Wed Apr 15 06:24:32.946 UTC

Cisco IOS XR Software, Version 4.2.4[Default]
Copyright (c) 2012 by Cisco Systems, Inc.

ROM: System Bootstrap, Version 2.06(20110916:145933) [CRS ROMMON],

CRS1 uptime is 14 weeks, 5 days, 16 hours, 27 minutes
System image file is "disk0:hfr-os-mbi-4.2.4/0x100008/mbihfr-rp-x86e.vm"

cisco CRS-8/S-B (Intel 686 F6M14S4) processor with 12582912K bytes of memory.
Intel 686 F6M14S4 processor at 1729Mhz, Revision 2.174
CRS-8 Line Card Chassis-enhanced for CRS-8/S-B

2 Management Ethernet
18 DWDM controller(s)
4 FortyGigE
22 TenGigE
10 SONET/SDH
10 Packet over SONET/SDH
14 WANPHY controller(s)
1019k bytes of non-volatile configuration memory.
15801M bytes of hard disk.
11223024k bytes of disk0: (Sector size 512 bytes).
11223024k bytes of disk1: (Sector size 512 bytes).
RP/0/RP0/CPU0:CRS1
```

KDDI Myanmar

IOS-XR CLI: New CLI format

```
RP/0/0/CPU0:CRS1#show ipv4 interface brief

Interface                      IP-Address      Status          Protocol
MgmtEth0/0/CPU0/0              10.23.1.69      Up              Up
MgmtEth0/0/CPU0/1              unassigned      Shutdown        Down
MgmtEth0/0/CPU0/2              unassigned      Shutdown        Down
GigabitEthernet0/2/0/0         100.12.1.1      Up              Up
```

XR basic commands

KDDI Myanmar

```
RP/0/0/CPU0:CRS1#show configuration commit list
SNo. Label/ID    User      Line               Client      Time Stamp
~~~~ ~~~~~~~~    ~~~~      ~~~~                ~~~~~~      ~~~~~~~~~~
1    1000000037  root      con0_0_CPU           CLI         01:39:03 UTC Mon Apr 24 2013
2    1000000036  JChmbr    vty1:node0_RP0_CPU   CLI         01:18:10 UTC Mon Apr 24 2013
3    1000000035  Mhmali    vty1:node0_RP0_CPU   CLI         01:00:54 UTC Mon Apr 24 2013
```

```
RP/0/0/CPU0:iox-CL11#rollback configuration to 1000000033
Loading Rollback Changes.
Loaded Rollback Changes in 1 sec
Committing.
3 items committed in 1 sec (2)items/sec
Updating.RP/0/0/CPU0:Apr 24 01:01:07.143 : config_rollback[65691]: %MGBL-CONFIG-6-DB_COMMIT :
Configuration committed by user 'root'. Use 'show configuration commit changes 1000000035' to
view the changes.

Updated Commit database in 1 sec
Configuration successfully rolled back to '1000000033'.
RP/0/0/CPU0:CRS1#
```

KDDI Myanmar

Static Routes

IOS

```
ip route 192.1.1.0 255.255.255.0 gi4/0
ip route 223.255.254.0 255.255.255.0 10.13.0.1
ipv6 route 5301::1111/128 fec0::1
```

IOS XR

```
router static
 address-family ipv4 unicast
  43.43.44.0/24 Serial0/5/3/3/0:2
  192.1.1.0/24 GigabitEthernet0/4/0/0
  223.255.254.254/32 MgmtEth0/1/CPU0/0
 !
 address-family ipv6 unicast
  5301::1111/128 Serial0/5/3/3/0:0
 !
!
```

KDDI Myanmar

OSPF Configuration

IOS

```
router ospf 1
  area 0 authentication message-digest
  area 2 authentication message-digest
  network 10.100.1.0 0.0.0.7 area 0
  network 10.200.1.0 0.0.0.15 area 2

interface gi0/0
 ip ospf network point-to-point
 ip ospf message-digest-key 1 md5 CISCO
 ip ospf cost 100
!
interface gi0/1
 ip ospf network point-to-point
 ip ospf message-digest-key 1 md5 CISCO
 ip ospf cost 100
!
interface gi0/2
 ip ospf network point-to-point
 ip ospf message-digest-key 1 md5 CISCO
 ip ospf cost 2000
!
interface gi0/3
 ip ospf network point-to-point
 ip ospf message-digest-key 1 md5 CISCO
 ip ospf cost  9999
```

IOS XR

```
router ospf 1
  authentication message-digest
  message-digest-key 1 md5 CISCO
  network point-to-point
area 0
   cost 100
   interface gi0/0/0/0
   !
   interface gi0/0/0/1
   !
 area 2
cost 2000
   interface gi0/0/0/2
   !
   interface gi0/0/0/3
     cost 9999
```

show run router ospf
show run router ospf 1 area 2

KDDI Myanmar

BGP

IOS

```
router bgp 65000
 bgp log-neighbor-changes
 no bgp default ipv4-unicast
 neighbor 172.16.2.5 remote-as 65111
 neighbor 192.168.1.2 remote-as 65000
 neighbor 192.168.1.2 update-source Loopback 0
!
 address-family ipv4
  network 192.168.1.1 mask 255.255.255.255
  neighbor 172.16.2.5 activate255.255.255.255
  neighbor 192.168.1.2 activate

  maximum-paths 8
```

IOS XR

```
router bgp 65000
 address-family ipv4 unicast
 network 192.168.1.1/32
 maximum-paths 8
 !
 neighbor 192.168.1.2
  remote-as 65000
  update-source Loopback 0
  address-family ipv4 unicast
 !
 neighbor 172.16.2.5
  remote-as 65111
  address-family ipv4 unicast
   route-policy PASS-ALL in
   route-policy PASS-ALL out
```

Will not accept routes from EBGP peers by default

BGP: Show commands

```
RP/0/1/CPU0:CRS1# show bgp ipv4 unicast summary
BGP router identifier 2.2.2.2, local AS number 300
BGP generic scan interval 60 secs
BGP table state: Active
BGP main routing table version 101
BGP scan interval 60 secs
BGP is operating in STANDALONE mode.

Process       RecvTblVer   bRIB/RIB  LabelVer  ImportVer  SendTblVer
Speaker          101          101       101        101         101

Neighbor        Spk    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  St/PfxRcd
192.1.1.2       0     400    2451    2453      101     0    0   00:24:33      100
```

RPL Examples

- Basic conditional statement

```
if med eq 150 then
    drop
endif
```

- Branching options

```
if med eq 150 then
     set local-preference 10
elseif med eq 200 then
     set local-preference 60
else
     set local-preference 0
endif
```

Comparison operator

Named and Inline Set (Same behavior)

```
route-policy USE_INLINE
  if as-path in (ios-regex '_42$', ios-regex '_127$') then
    pass
endif
end-policy
```

```
as-path-set NAMED_SET
  ios-regex '_42$',
  ios-regex '_127$'
end-set

route-policy USE_NAMED
  if as-path in NAMED_SET then
    pass
endif
end-policy
```

RPL Examples

- Bad RPL Logic

```
Route-policy BAD-RPL
        if med eq 150 then
                set local-preference 10
        endif
        set local-preference 0
End-policy
```

Overwrites Setting

- Good RPL Logic

```
Route-policy GOOD-RPL
        if med eq 150 then
                set local-preference 10
                done
        endif
        set local-preference 0
End-policy
```

Stops all processing on matched prefixes

KDDI Myanmar

RPL Actions
The route policy requires a "ticket" for the route to be accepted or dropped.
These are the different operatators

- **Pass** – prefix allowed if not later dropped

  pass grants a ticket to defeat default drop

  Execution continues after pass
- **Set** – value changed, prefix allowed if not later dropped

  Any set at any level grants a ticket

  Execution continues after set

  Values can be set more than once
- **Done** – prefix allowed, stop execution
- **Drop** – prefix is discarded

  Explicit drop stops policy execution

  Implicit drop (if policy runs to end without getting a ticket)

RPL Show Commands

Only display prefixes matching policy – filter show command

```
RP/0/0/1:CRS1#show bgp route-policy SAMPLE
BGP router identifier 172.20.1.1, local AS number 1820
BGP main routing table version 729
Dampening enabled
BGP scan interval 60 secs
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
* 10.13.0.0/16 192.168.40.24 0 1878 704 701 200 ?
* 10.16.0.0/16 192.168.40.24 0 1878 704 701 i
```

KDDI Myanmar

RPL Show Commands

```
RP/0/0/CPU0:CRS1#show rpl route-policy states

ACTIVE -- Referenced by at least one policy which is attached
INACTIVE -- Only referenced by policies which are not attached
UNUSED -- Not attached (directly or indirectly) and not referenced

The following policies are (ACTIVE)
----------------------------------------
..
The following policies are (INACTIVE)
----------------------------------------
None found with this status.

The following policies are (UNUSED)
----------------------------------------
…
```

KDDI Myanmar

Trace functionality

```
RP/0/RP1/CPU0:CRS1#show ospf trace
 OSPF Trace Summary (2, RP/1/RP0/CPU0:CRS1, 0M)

    Trace Name      Size      Count  Description
    ------------ ------- ---------- -------------------------
 1. adj            65536       6291  adjacency
 2. adj_cycle      65536     893383  dbd/flood events/pkts
 3. config          2048        486  config events
 4. errors          8192     868816  errors
 5. events          4096        255  mda/rtrid/bfd/vrf
 6. ha              8192        485  startup/HA/NSF
 7. hello           2048    3982447  hello events/pkts
 8. idb             8192        973  interface
 9. pkt             2048    1927767  I/O packets
10. rib            65536      52190  rib batching
11. spf            65536      93138  spf/topology
12. spf_cycle      65536     352143  spf/topology detail
13. te              4096       3893  mpls-te
14. test            1024      20052  testing info
15. mg             65536          5  message queue info
RP/0/RP0/CPU0:CRS1#show ospf trace hello
Traces for OSPF 2 (Wed Jan 22 08:55:38)
Traces returned/requested/available:  2048/2048/2048
Trace buffer: hello

1    Jan 22 08:49:45.305* ospf_send_hello: area 0.0.0.80 intf MADJ: BE1008 from 0.0.0.0
2    Jan 22 08:49:45.546  ospf_rcv_hello: intf BE1009 area 0.0.0.74 from 10.1.0.9 10.1.9.2
3    Jan 22 08:49:45.546  ospf_check_hello_events: intf MADJ: BE1009 area 0.0.0.74 from 0.0.0.0
4    Jan 22 08:49:45.573* ospf_send_hello: area 0.0.0.74 intf MADJ: BE1008 from 0.0.0.0
5    Jan 22 08:49:45.845* ospf_rcv_hello: intf BE1009 area 0.0.0.80 from 10.1.0.9 10.1.9.2
6    Jan 22 08:49:45.845* ospf_check_hello_events: intf MADJ: BE1009 area 0.0.0.80 from 0.0.0.0
7    Jan 22 08:49:45.917* ospf_send_hello: area 0.0.0.80 intf Te0/5/0/7 from 10.1.80.1
```

Monitor interface

```
RP/0/RP1/CPU0:CRS1#monitor interface Bundle-ether 1008
CRS1              Monitor Time: 00:00:18          SysUptime: 246:02:20

Bundle-Ether1008 is up, line protocol is up
Encapsulation ARPA

Traffic Stats:(2 second rates)                                 Delta
  Input  Packets:                    6489005                     14
  Input  pps:                              8
  Input  Bytes:                   1507217455                   1274
  Input  Kbps (rate):                      5                 (  0%)
  Output Packets:                    7079943                     15
  Output pps:                              9
  Output Bytes:                   1490126647                   2024
  Output Kbps (rate):                      8                 (  0%)

Errors Stats:
  Input  Total:                            0                      0
  Input  CRC:                              0                      0
  Input  Frame:                            0                      0
  Input  Overrun:                          0                      0
  Output Total:                            0                      0
  Output Underrun:                         0                      0

Quit='q', Freeze='f', Thaw='t', Clear='c', Interface='i',
Next='n', Prev='p'

Brief='b', Detail='d', Protocol(IPv4/IPv6)='r'
```

KDDI Myanmar

CPU

```
RP/0/RP0/CPU0:CRS1#show proc cpu | ex " 0%  ."
CPU utilization for one minute: 2%; five minutes: 2%; fifteen minutes: 2%

PID    1Min    5Min    15Min Process
131105   1%      1%       1% ce_switch
131106   1%      1%       1% eth_server
RP/0/RP0/CPU0:CRS1#
```

'Show memory compare' command

- Process how to use the command:
  1. **show memory compare start**
     Takes the initial snapshot of heap usage
  2. **show memory compare end**
     Takes the second snapshot of heap usage
  3. **show memory compare report**
     Displays the heap memory comparison report

```
RP/0/RP1/CPU0:CRS1#show memory compare start
Successfully stored memory snapshot /harddisk:/malloc_dump/memcmp_start.out
RP/0/RP1/CPU0:CRS1#show memory compare end
Successfully stored memory snapshot /harddisk:/malloc_dump/memcmp_end.out
RP/0/RP1/CPU0:CRS1#show memory compare report

JID    name                 mem before    mem after     difference mallocs restart
---    ----                 ----------    ---------     ---------- ------- -------
57     i2c_server           11756         11916         160        1
121    bgp                  2522256       2522208       -48        -1
234    lpts_pa              408536        407632        -904       -14
224    isis                 3089108       3087900       -1208      0
314    tcp                  247196        245740        -1456      -9
241    netio                808136        806464        -1672      -46
```

KDDI Myanmar

Process Restartability

```
RP/0/RP1/CPU0:CRS1#process shutdown snmpd

RP/0/RP1/CPU0:CRS1#show processes snmpd
                  Job Id: 288
                     PID: 143532
         Executable path: /disk0/hfr-base-
4.2.1/bin/snmpd
             Instance #: 1
             Version ID: 00.00.0000
                Respawn: ON
          Respawn count: 1
  Max. spawns per minute: 12
           Last started: Mon May  9 15:32:22 2005
          Process state: Killed (last exit status: 15)
          Package state: Normal
      Registered item(s): cfg/gl/snmp/
                          cfg/gl/udpsnmp/
                          cfg/gl/mibs/
                    core: TEXT SHAREDMEM MAINMEM
                Max. core: 0
            startup_path: /pkg/startup/snmpd.startup
                   Ready: 11.636s
```

JID# remains constant, PID# changed on restart

Respawn counter incremented with process restart

Process state reported as 'killed'

```
RP/0/RP1/CPU0:CRS1#process restart snmpd

RP/0/RP1/CPU0:CRS1#show processes snmpd
                  Job Id: 288
                     PID: 8528114
         Executable path: /disk0/hfr-base-
4.2.1/bin/snmpd
             Instance #: 1
             Version ID: 00.00.0000
                Respawn: ON
          Respawn count: 2
  Max. spawns per minute: 12
           Last started: Thu May 12 11:46:38 2005
          Process state: Run (last exit status : 15)
          Package state: Normal
      Started on config: cfg/gl/snmp/admin/community/ww
                    core: TEXT SHAREDMEM MAINMEM
                Max. core: 0
            startup_path: /pkg/startup/snmpd.startup
                   Ready: 6.657s
        Process cpu time: 0.721 user, 0.145 kernel,
0.866 total
```

KDDI Myanmar

How to provisioning a new circuit for IPTransit or Customer ?

Step #1 : configure interface ip (ipv4 and iopv6)

```
interface TenGigE0/0/0/6
 ipv4 address 203.208.177.94 255.255.255.252
 ipv6 address 2001:c10:80:2::43e/126
 ipv6 enable
```

Step #2 : configure BGP peer to partner ISP (IP Transit or Peering)

```
router bgp 45558
 neighbor 203.208.177.93
 remote-as 7473
 address-family ipv4 unicast
  send-community-ebgp ( if using a community )
```

KDDI Myanmar

Step #3 : route-policy configuration if need

```
route-policy ST7473-10G-OUT
  if (as-path in aspath_2_p1_permit) then
    drop
  elseif (destination in pfx_MPT_203.81.64_21_p1_permit) then
    set med 500
    prepend as-path 45558 5
  elseif (destination in pfx_MPT_203.81.72_21_p1_permit) then
    set med 500
    prepend as-path 45558 5
  elseif (destination in pfx_MPT_203.81.80_21_p1_permit) then
    set med 500
  elseif (destination in pfx_MPT_203.81.88_21_p1_permit) then
    set med 500
  elseif (destination in pfx_MPT_103.25.12_22_p1_permit) then
    set med 500
    prepend as-path 45558 3
  elseif (destination in pfx_MPT_45.112.176_22_p1_permit) then
    set med 500
  elseif (destination in pfx_MPT_103.52.12_22_p1_permit) then
    set med 500
  elseif (destination in pfx_MPT_103.47.184_24_p1_permit) then
    set med 600
  elseif (destination in pfx_NGW_p1_permit) then
    set med 500
  elseif (destination in pfx_TELEPORT_all_p1_permit) then
    set med 500
  elseif (destination in pfx_TELENOR_p1_permit) then
    set med 500
  elseif (destination in pfx_CDN_p1_permit) then
    set med 500
  endif
end-policy
```

```
as-path-set aspath_1_p1_permit
  ios-regex '^4657_',
  ios-regex '^3491_',
  ios-regex '^6762_',
  ios-regex '^4788_',
  ios-regex '^4651_',
  ios-regex '^4809_',
  ios-regex '^7713_',
  ios-regex '^9304_',
  ios-regex '^2914_'
end-set

prefix-set pfx_MPT_203.81.64_21_p1_permit
 103.52.12.0/22 le 24
end-set

.
.
.
.
.
```

```
router bgp 45558
 neighbor 203.208.177.93
 route-policy ST7473-10G-OUT out
```

(*)you can set up a in-policy as well

KDDI Myanmar

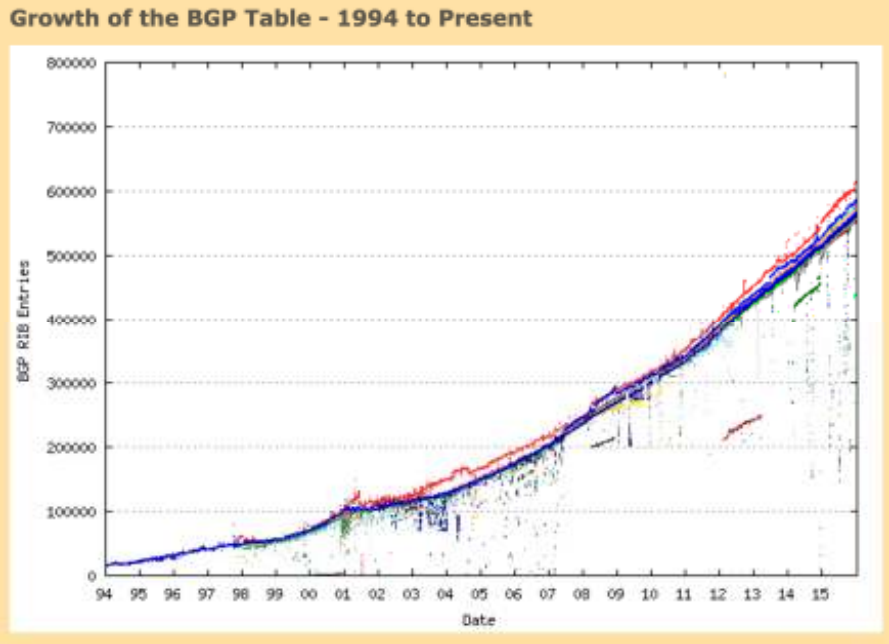
IPv6 Summary

Why need IPv6

KDDI Myanmar

- IETF IPv6 WG began in early 90s, to solve addressing growth issues
- but CIDR, NAT,…were developed !
- IPv4 32 bit address = 4 billion hosts
- IP is everywhere now
  Data, voice, audio and video integration …
- So, only compelling reason: More IP addresses

(*) Then after stand up IPv6 , we also need to consider Router Spec



-> what dose it meaning ?
-> Xday will be … ??

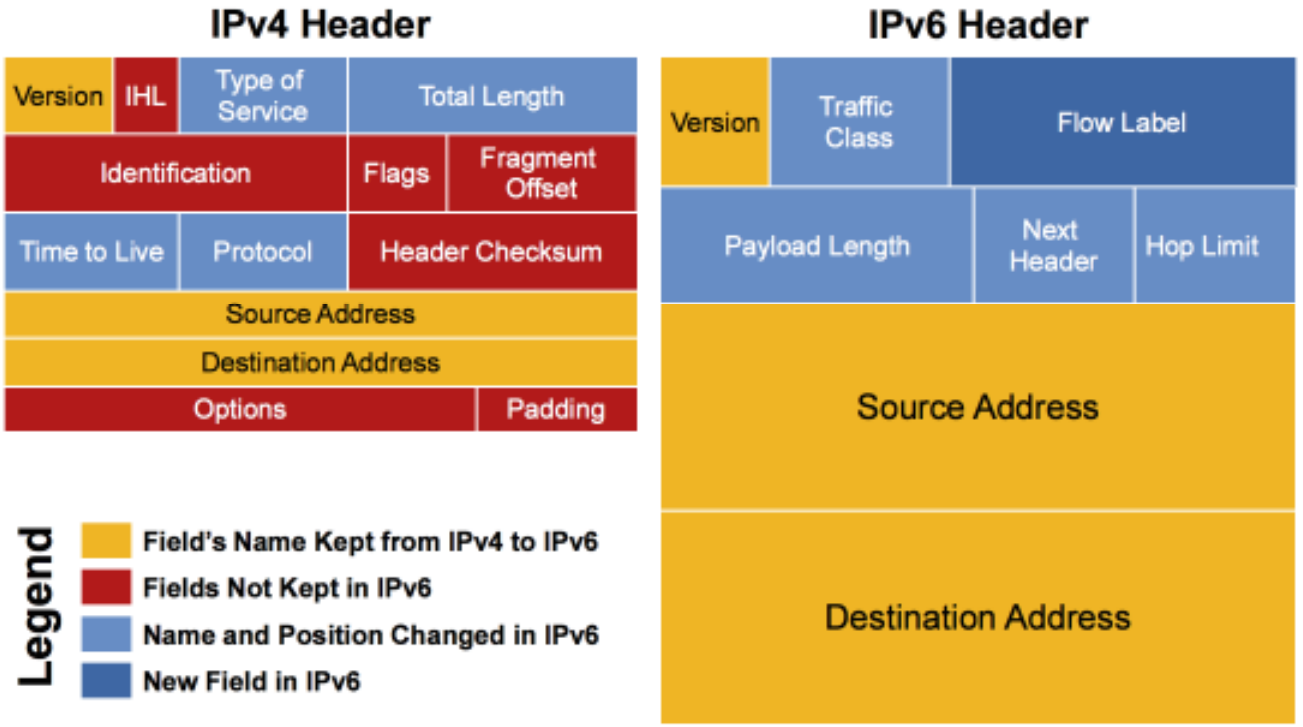
Why not NAT ?

KDDI Myanmar

- It was created as a temp solution
- NAT breaks the end-to-end model
- Growth of NAT has slowed down growth of transparent applications
- No easy way to maintain states of NAT in case of node failures
- NAT break security
- NAT complicates mergers, double NATing is needed for devices to communicate with each other

KDDI Myanmar

IPv4 and IPv6 Header Comparison



- 20-Bit Flow Label Field to Identify Specific Flows need special QOS

KDDI Myanmar

IPv4 32-bits

IPv6 128-bits

2^{32} = 4,294,967,296

2^{128} = 340,282,366,920,938,463,463,374,607,431,768,211,456

$2^{128} = 2^{32} * 2^{96}$

2^{96} = 79,228,162,514,264,337,593,543,950,336 times the number of possible IPv4 Addresses

KDDI Myanmar

Type	Binary	Hex
Unspecified	000…0	::/128
Loopback	000…1	::1/128
Global Unicast Address	0010	2000::/3
Link Local Unicast Address	1111 1110 10	FE80::/10
Unique Local Unicast Address	1111 1100 1111 1101	FC00::/7
Multicast Address	1111 1111	FF00::/8

Addressing model

KDDI Myanmar

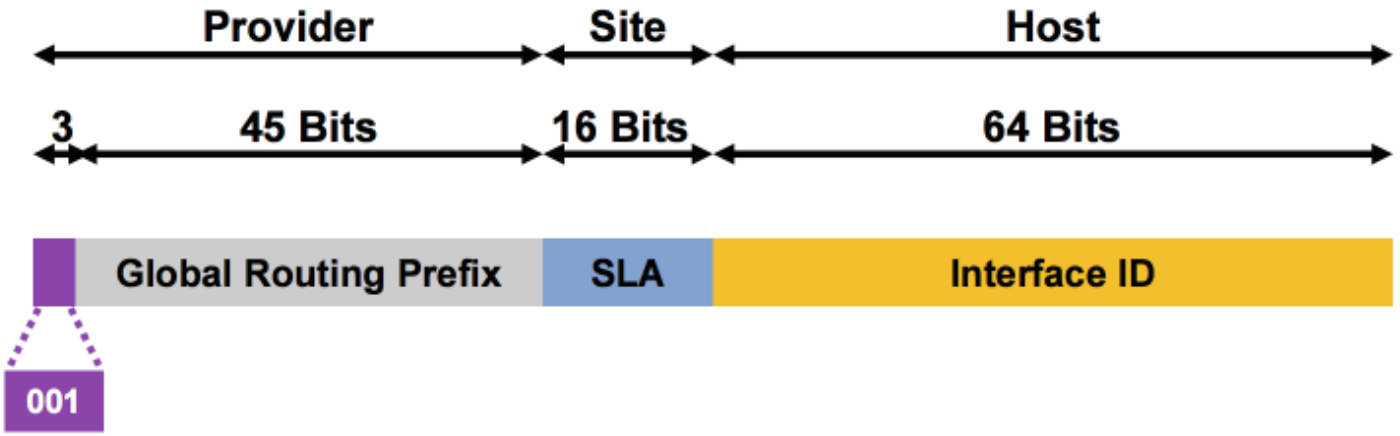
- Unicast
  - Address of a single interface .
  - One-to-one delivery to single interface !
- Multicast
  - Address of a set of interfaces.
  - One-to-many delivery to all
- Anycast
  - Address of a set of interfaces.
  - One-to-one-of-many delivery to a single interface in the set that is closest
- No more broadcast addresses

IPv6—Addressing Model

- Addresses are assigned to interfaces
  - Change from IPv4 mode:
- Interface "expected" to have multiple addresses
- Addresses have scope
  - Link Local
  - Unique Local
  - Global

KDDI Myanmar

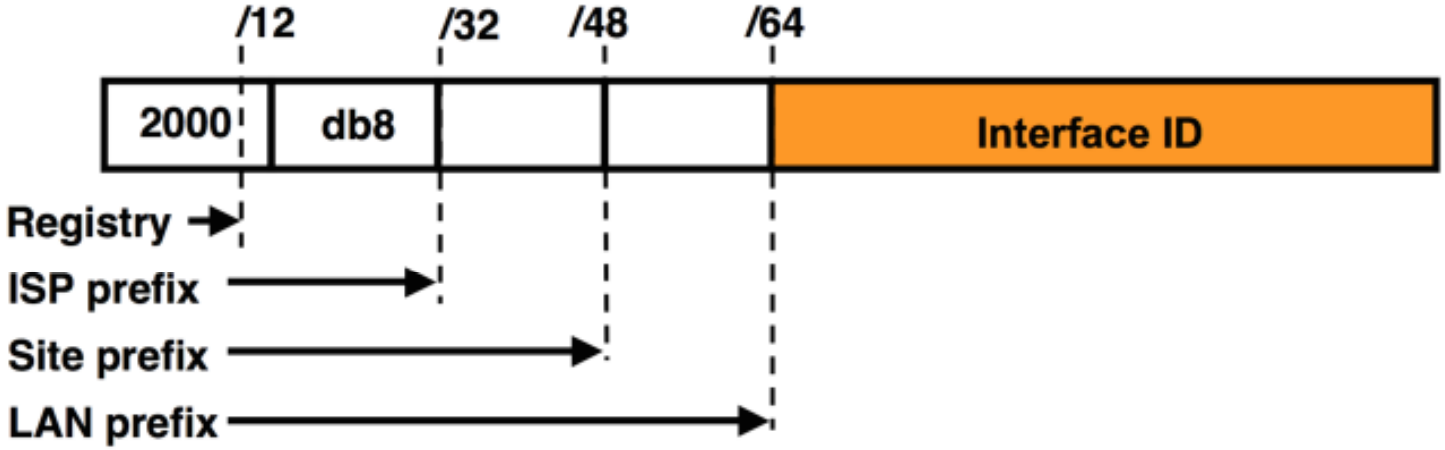
Aggregatable Global Unicast Addresses



Aggregatable Global Unicast Addresses Are:

- Addresses for generic use of IPv6
- Structured as a hierarchy to keep the aggregation

KDDI Myanmar



- The allocation process is:
  - The IANA is allocating out of 2000::/3 for initial IPv6 unicast use
  - Each registry gets a /12 prefix from the IANA
  - Registry allocates a /32 prefix (or larger) to an IPv6 ISP
  - Policy is that an ISP allocates a /48 prefix to each end customer

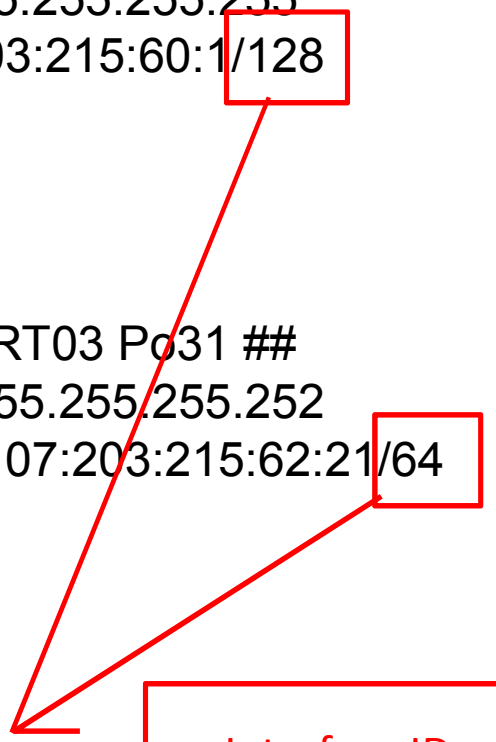
KDDI Myanmar

- 64 bits reserved for the interface ID
  - Possibility of 2^{64} hosts on one network LAN
  - In theory 18,446,744,073,709,551,616 hosts
  - Arrangement to accommodate MAC addresses within the IPv6 address
- 16 bits reserved for the end site
  - Possibility of 2^{16} networks at each end-site
  - 65536 subnets equivalent to a /12 in IPv4 (assuming a /28 or 16 hosts per IPv4 subnet)

Let's your current config …

KDDI Myanmar

```
interface Loopback0
ip address 203.215.60.1 255.255.255.255
ipv6 address 2401:F200::203:215:60:1/128
ipv6 enable
ipv6 ospf 45558 area 0

interface Port-channel31
description ## YGN-INGW-RT03 Po31 ##
ip address 203.215.62.21 255.255.255.252
ipv6 address 2401:F200:0:107:203:215:62:21/64
ipv6 enable
ipv6 ospf 45558 area 0
```



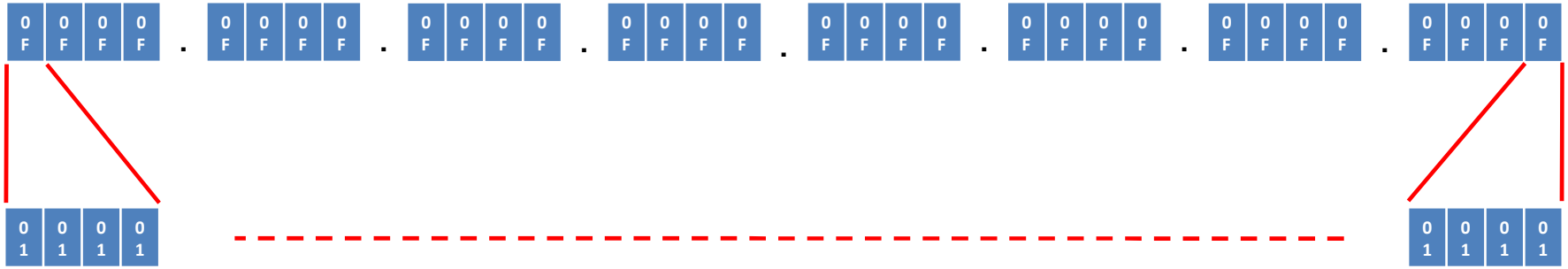
- Interface ID range is too big
- Just fill in /size , normally use /64 as interface ID

How to calculate IPv6 address ?

KDDI Myanmar

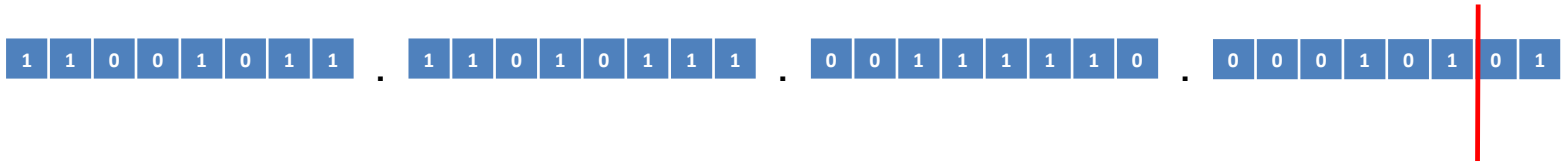
IPv4 (32bit):
(*) Binary numbers

IPv6(128bit):
(*) Hexadecimal number

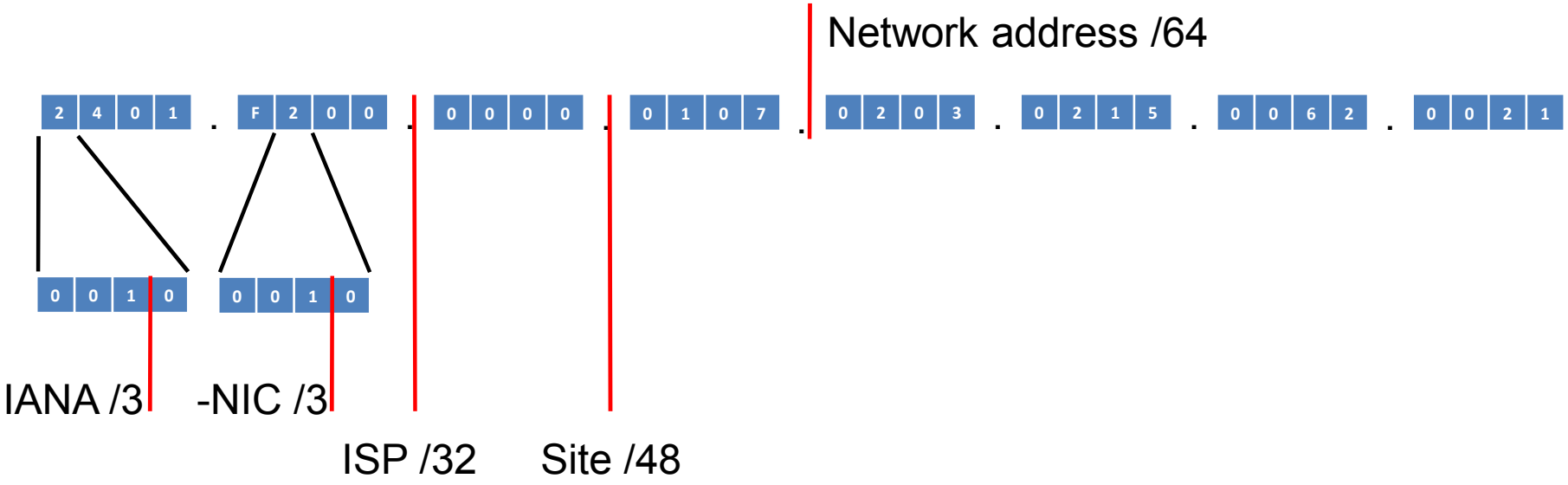


How to calculate IPv6 address ?

203.215.62.21 255.255.255.252 mean ?

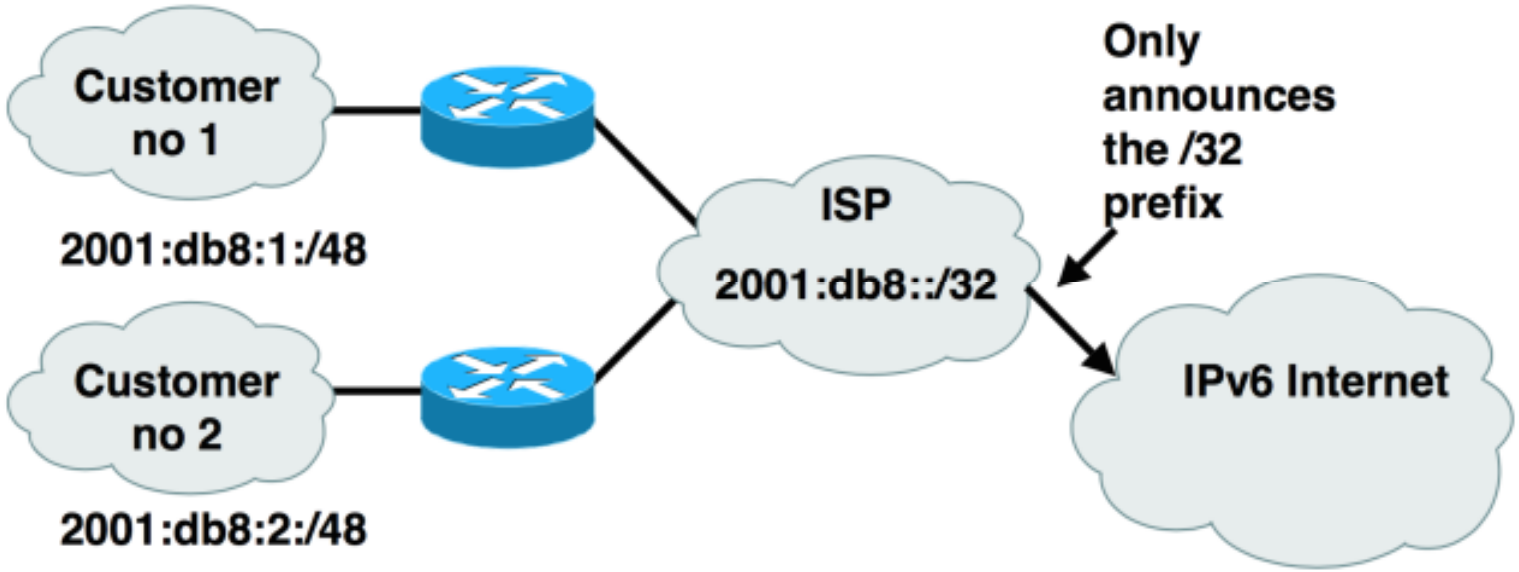


2401:F200:0:107:203:215:62:21/64 mean ?



Simple designing of IPv6 (ISP to internet)

KDDI Myanmar



- Larger address space enables aggregation of prefixes announced in the global routing table

- Idea was to allow efficient and scalable routing

KDDI Myanmar

- IP multicast address has a prefix FF00::/8 (1111 1111); the second octet defines the lifetime and scope of the multicast address

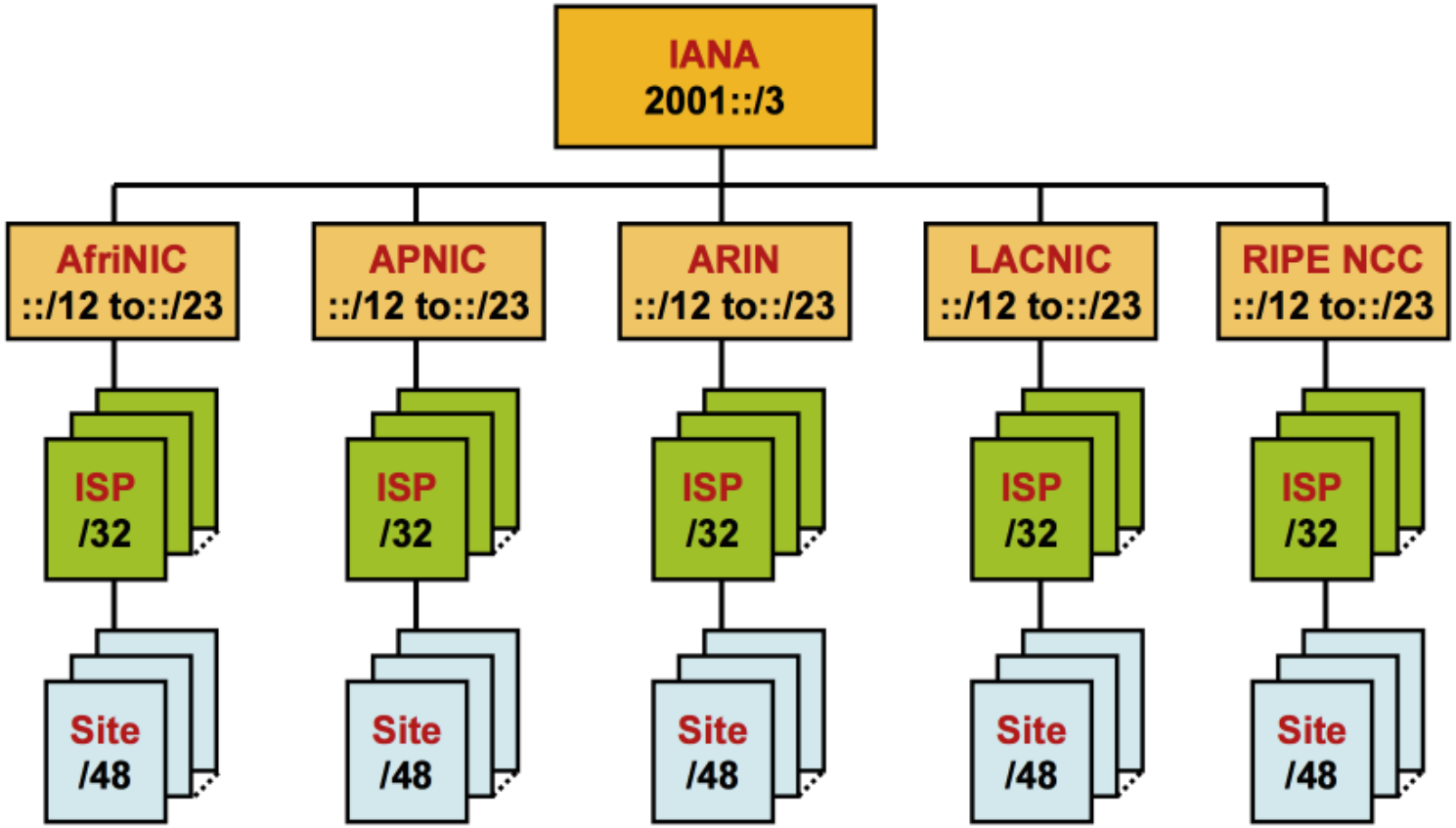
8-bit	4-bit	4-bit	112-bit
1111 1111	Lifetime	Scope	Group-ID

Lifetime	
0	If Permanent
1	If Temporary

Scope	
1	Node
2	Link
5	Site
8	Organization
E	Global

IPv6 Address policy

KDDI Myanmar



IPv6 MTU size

KDDI Myanmar

- Minimum link MTU for IPv6 is 1280 octets
- (68 octets for IPv4)
- Implementations are expected to perform path MTU discovery to send packets bigger than 1280

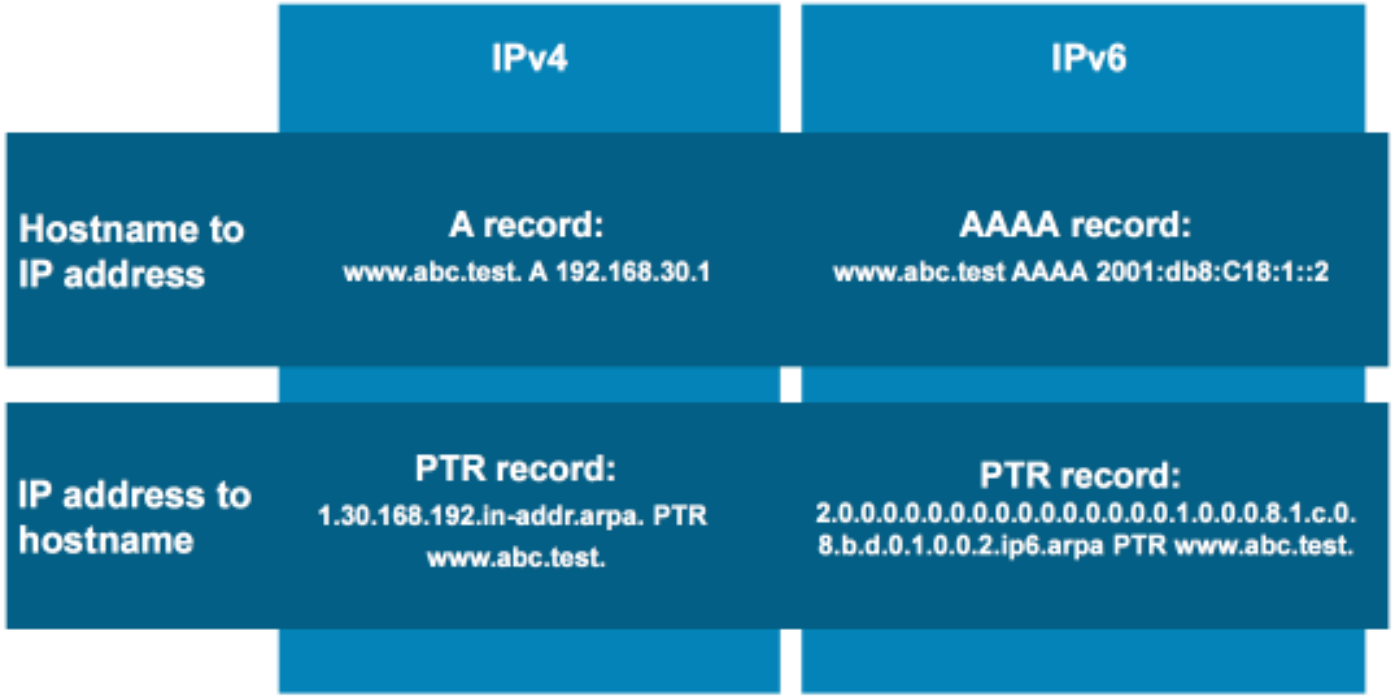
Neighbor Discovery

- Protocol built on top of ICMPv6 (RFC 4443)
- combination of IPv4 protocols (ARP, ICMP, IGMP,...)
- Fully dynamic, interactive between Hosts & Routers defines 5 ICMPv6 packet types:

  Router Solicitation
  Router Advertisements
  Neighbour Solicitation
  Neighbour Advertisements
  Redirect

IPv6 and DNS

KDDI Myanmar



IPv6 Tech Summary

KDDI Myanmar

IP Service	IPv4 Solution	IPv6 Solution
Addressing Range	32-bit, Network Address Translation	128-bit, Multiple Scopes
Autoconfiguration	DHCP	Serverless, Reconfiguration, DHCP
Security	IPSec	IPSec Mandated, works End-to-End
Mobility	Mobile IP	Mobile IP with Direct Routing
Quality-of-Service	Differentiated Service, Integrated Service	Differentiated Service, Integrated Service
IP Multicast	IGMP/PIM/Multicast BGP	MLD/PIM/Multicast BGP,Scope Identifier

KDDI Myanmar

- IPsec standards apply to both IPv4 and IPv6
- All implementations required to support authentication and encryption headers (“IPsec”)
- Authentication separate from encryption for use in situations where encryption is prohibited or prohibitively expensive
- Key distribution protocols are not yet defined (independent of IP v4/v6)
- Support for manual key configuration required

KDDI Myanmar

- Two basic approaches developed by IETF:
- “Integrated Service” (int-serv)
  - Fine-grain (per-flow)
  - quantitative promises
  - uses RSVP signaling

- “Differentiated Service” (diff-serv)
  - Coarse-grain (per-class)
  - qualitative promises (e.g., higher priority)
  - no explicit signaling

- Signaled diff-serv (RFC 2998)
  - Uses RSVP for signaling with course-grained qualitative aggregate markings
  - Allows for policy control without requiring per-router state overhead

IPv6 Standard

KDDI Myanmar

- Several key components on standards track…
  - Specification (RFC2460)
  - ICMPv6 (RFC4443)
  - RIP (RFC2080)
  - IGMPv6 (RFC2710)
  - Router Alert (RFC2711)
  - Autoconfiguration (RFC4862)
  - DHCPv6 (RFC3315 & 4361)
  - IPv6 Mobility (RFC3775)
  - GRE Tunnelling (RFC2473)
  - DAD for IPv6 (RFC4429)
  - ISIS for IPv6 (RFC5308)
  - Neighbour Discovery (RFC4861 & 4311)
  - IPv6 Addresses (RFC4291 & 3587)
  - BGP (RFC2545)
  - OSPF (RFC5340)
  - Jumbograms (RFC2675)
  - Radius (RFC3162)
  - Flow Label (RFC3697)
  - Mobile IPv6 MIB (RFC4295)
  - Unique Local IPv6 Addresses (RFC4193)
  - Teredo (RFC4380)
- IPv6 available over:
  - PPP (RFC5072)
  - FDDI (RFC2467)
  - NBMA (RFC2491)
  - Frame Relay (RFC2590)
  - IEEE1394 (RFC3146)
  - Facebook (RFC5514)
  - Ethernet (RFC2464)
  - Token Ring (RFC2470)
  - ATM (RFC2492)
  - ARCnet (RFC2497)
  - FibreChannel (RFC4338)

KDDI Myanmar

- Why nobody do not like to move ipv6 ?
  -> because all is thinking ipv6 is defficult …

But …

- If you understood until last page , you will be able to design and configure IPv6

- Actually almost configuration is same with ipv4 .

- Let's show sample configuration from next page and you will get how to configure and what is different with ipv4 .

However …

- We also need to consider how to combine ipv6 into ipv4 network
- Lets see one by one …

KDDI Myanmar

Let's thinking ;

- **If you want to build a IPv6 network ,**
- **How is your network router capacity ?**
- **Especially if running a BGP as ISP network , need to consider and calculate a total capacity of both ipv4 + ipv6 prefix , CPU , memery ,bandwidth etc …**

KDDI Myanmar

Ask to ;

KDDI Myanmar

- From the RIR
  Receive a /32 (or larger if you have more than 65k /48 assignments)

- From upstream ISP
  Get one /48 from your upstream ISP
  More than one /48 if you have more than 65k subnets

- Use 6to4
  Take a single public IPv4 /32 address
  2002:<ipv4 /32 address>::/48
  becomes your IPv6 address block, giving 65k subnets

KDDI Myanmar

- ISPs should receive /32 from their RIR
- Address block for router loop-back interfaces
  - Generally number all loopbacks out of **one** /64
  - /128 per loopback

- Address block for infrastructure
  - /48 allows 65k subnets
  - /48 per PoP or region (for large networks)
  - /48 for whole backbone (for small to medium networks)
  - Summarise between sites if it makes sense

KDDI Myanmar

For enable a IPv6 on your router ;

Router(config)# ipv6 unicast-routing Also enable IPv6 CEF (not on by default):
Router(config)# ipv6 cef Also disable IPv6 Source Routing (enabled by default):
Router(config)# no ipv6 source-routing

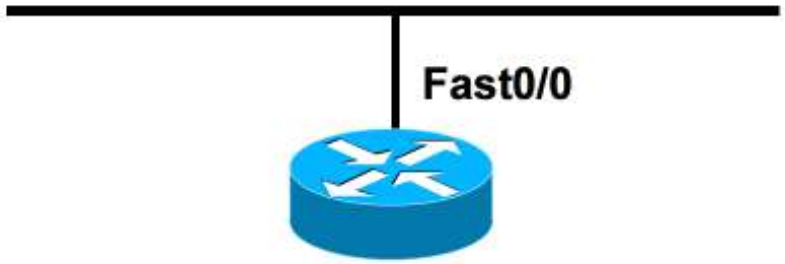
To configure a global or unique-local IPv6 address the following interface command should be entered:

Router(config-if)# ipv6 address X:X..X:X/prefix

To configure an EUI-64 based IPv6 address the following interface command should be entered:

Router(config-if)# ipv6 address X:X::/prefix eui-64
(*) This is not useful on a router and is **not recommended**

IPv6 Addressing to interface

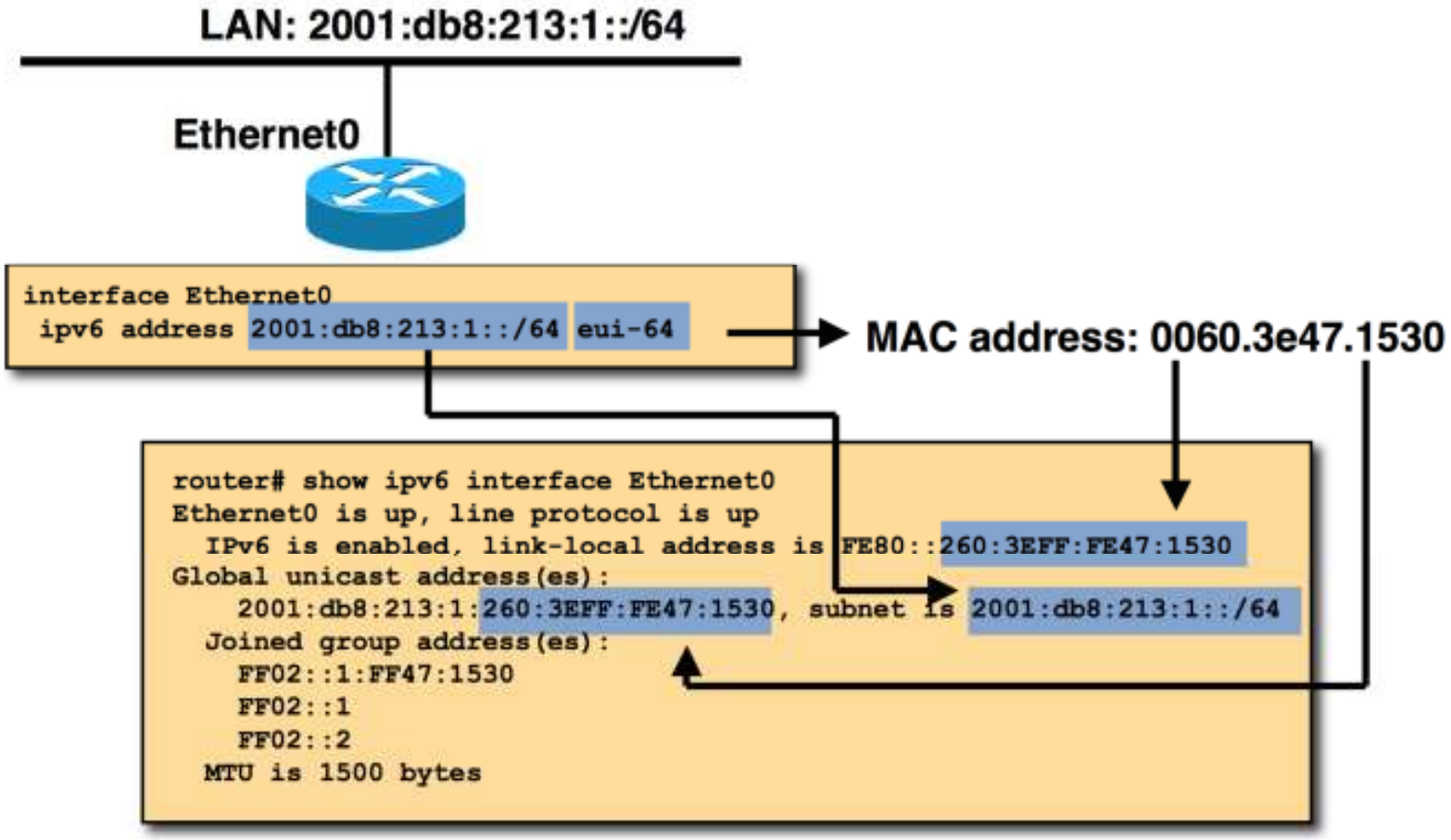



```
ipv6 unicast-routing
!
interface FastEthernet0/0
 ip address 10.151.1.1 255.255.255.0
 ip pim sparse-mode
 duplex auto
 speed auto
 ipv6 address 2006:1::1/64
 ipv6 enable
 ipv6 nd ra-interval 30
 ipv6 nd prefix 2006:1::/64 300 300
!
```

show commands ;

show ipv6 int gi0/0

IPv6 Addressing (eui-64)



KDDI Myanmar

Syntax is:

ipv6 route ipv6-prefix/prefix-length {ipv6-address | interface-type interface-number} [administrative-distance]

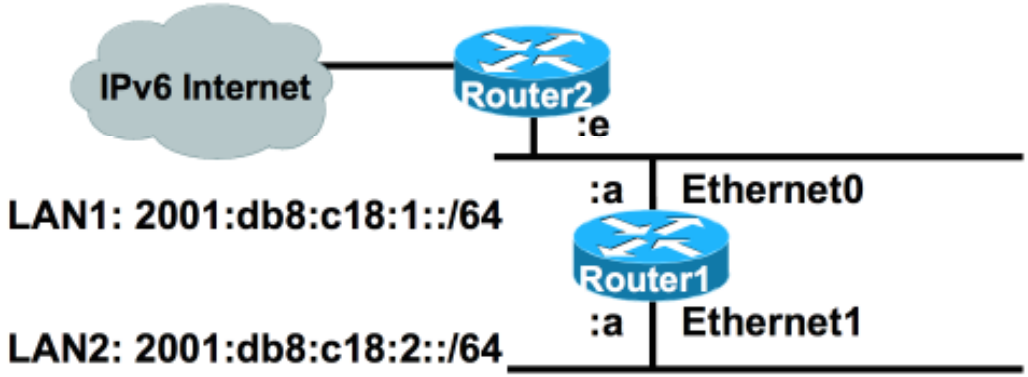
Static Route

```
ipv6 route 2001:db8::/64 2001:db8:0:CC00::1 150
```

Routes packets for network 2001:db8::/64 to a networking device at 2001:db8:0:CC00::1 with an administrative distance of 150

KDDI Myanmar

Default Routing Example




```
ipv6 unicast-routing
!
interface Ethernet0
 ipv6 address 2001:db8:c18:1::a/64
!
interface Ethernet1
 ipv6 address 2001:db8:c18:2::a/64
!
ipv6 route ::/0 2001:db8:c18:1::e
```

Default Route to Router2

KDDI Myanmar

Dynamic Routing in IPv6 is unchanged from IPv4:

- IPv6 has 2 types of routing protocols: IGP and EGP
- IPv6 still uses the longest-prefix match routing algorithm

IGP

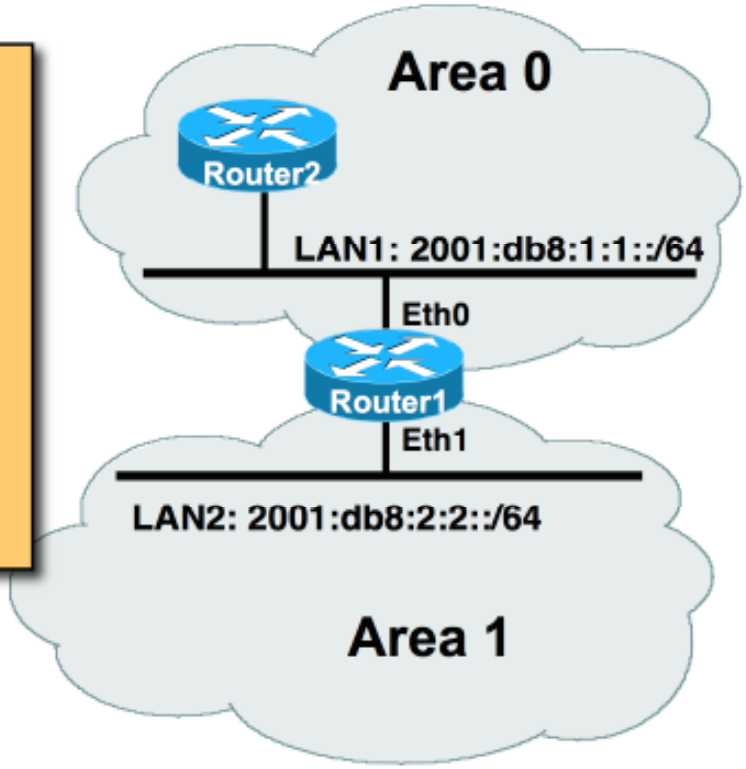
- RIPng (RFC 2080)
- Cisco EIGRP for IPv6
- OSPFv3 (RFC 5340)
- Integrated IS-ISv6 (RFC 5308)

EGP

- MP-BGP4 (RFC 4760 and RFC 2545)

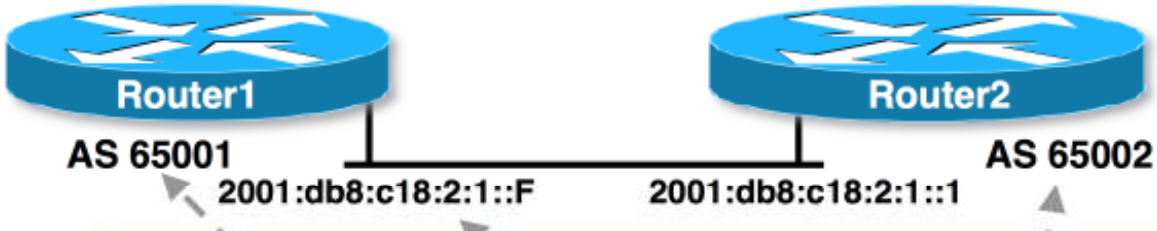
OSPFv3 configuration example

```
Router1#
 interface Ethernet0
  ipv6 address 2001:db8:1:1::1/64
  ipv6 ospf 1 area 0

 interface Ethernet1
  ipv6 address 2001:db8:2:2::2/64
  ipv6 ospf 1 area 1

 ipv6 router ospf 1
  router-id 1.1.1.1
```



KDDI Myanmar

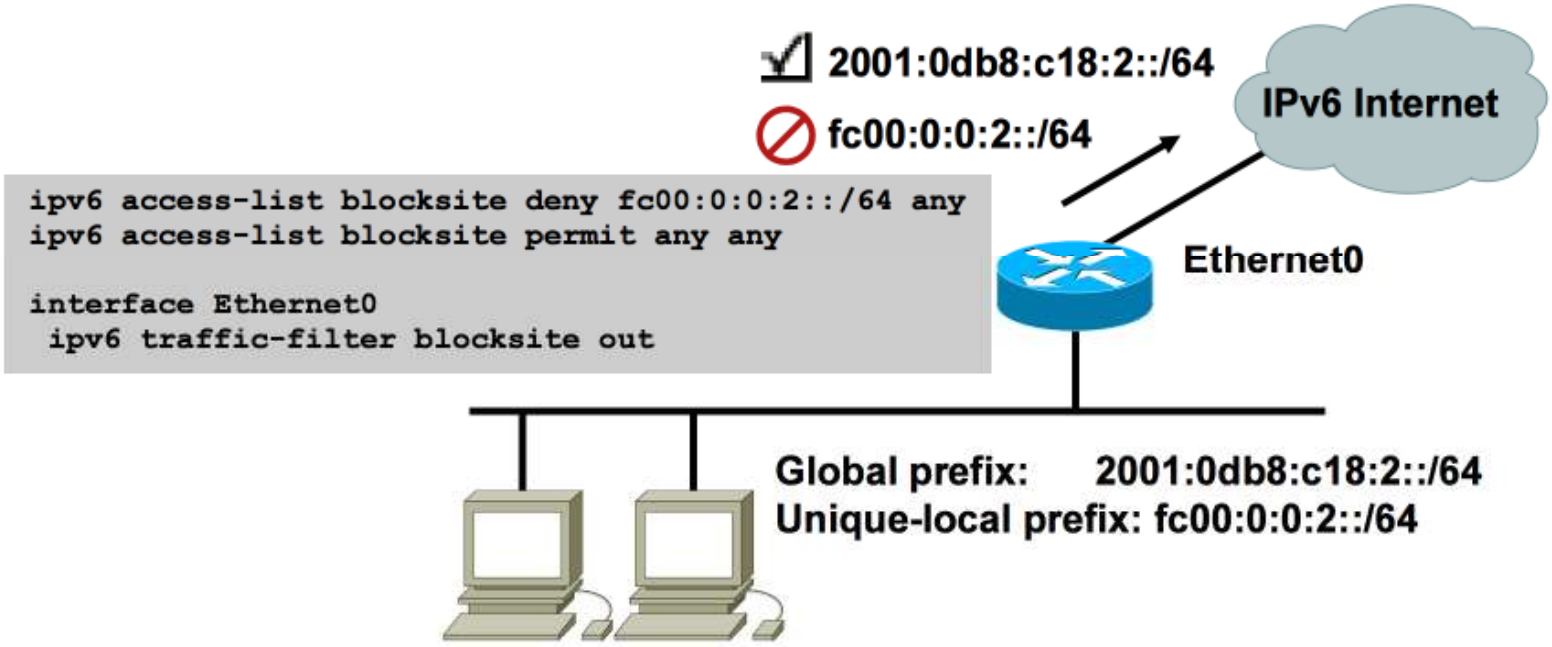
A Simple MP-BGP Session




```
Router1#
interface Ethernet0
 ipv6 address 2001:db8:c18:2:1::F/64
!
router bgp 65001
 bgp router-id 10.10.10.1
 no bgp default ipv4-unicast
 neighbor 2001:db8:c18:2:1::1 remote-as 65002
 address-family ipv6
  neighbor 2001:db8:c18:2:1::1 activate
  neighbor 2001:db8:c18:2:1::1 prefix-list bgp65002in in
  neighbor 2001:db8:c18:2:1::1 prefix-list bgp65002out out
  exit-address-family
```

KDDI Myanmar

- Filtering outgoing traffic from site-local source addresses



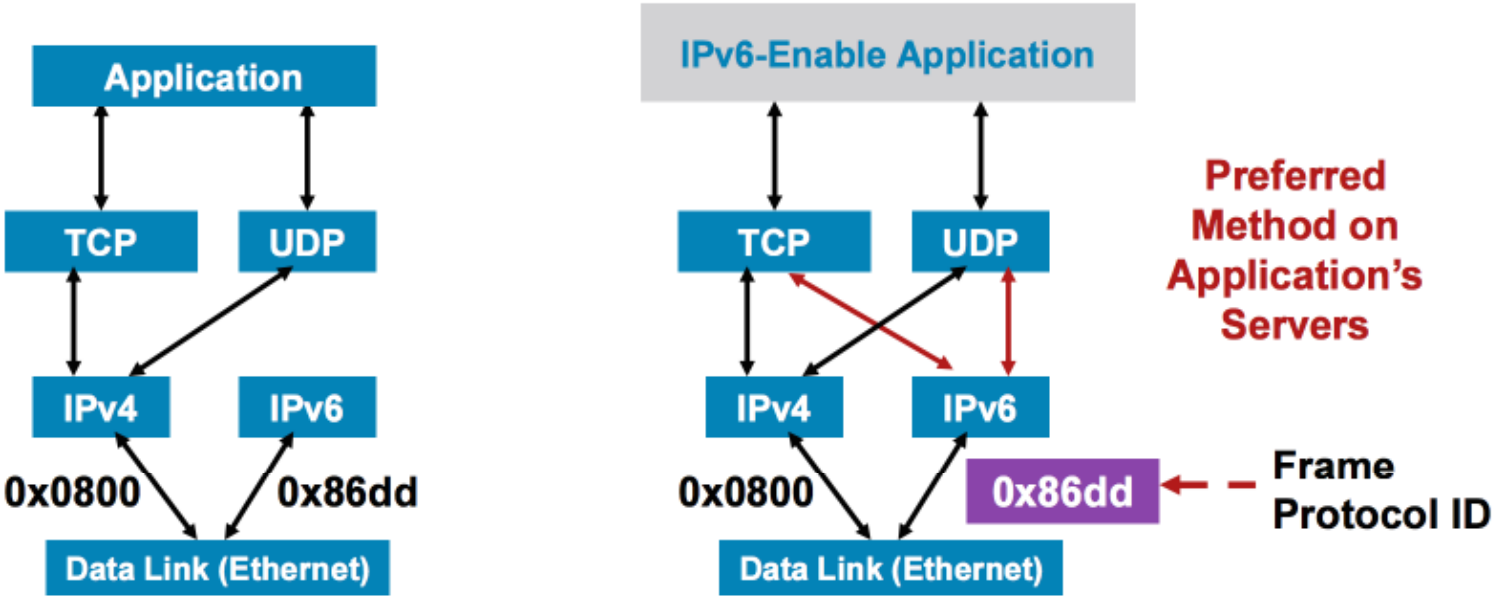
KDDI Myanmar

IPv4-IPv6 Transition/Coexistence

- A wide range of techniques have been identified and implemented, basically falling into three categories:
  1. Dual-stack techniques, to allow IPv4 and IPv6 to co-exist in the same devices and networks
  2. Tunneling techniques, to avoid order dependencies when upgrading hosts, routers, or regions
  3. Translation techniques, to allow IPv6-only devices to communicate with IPv4-only devices
- Expect all of these to be used, in combination

IPv4 - IPv6 Dual Stack Approach

KDDI Myanmar

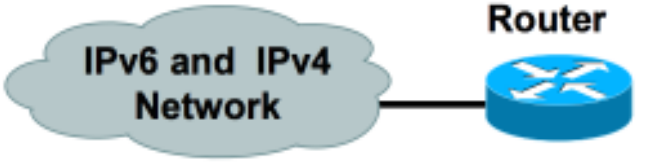


Dual Stack Node Means:

- Both IPv4 and IPv6 stacks enabled
- Applications can talk to both
- Choice of the IP version is based on name lookup and application preference

IPv4 - IPv6 Dual Stack configuration

KDDI Myanmar




```
router#
ipv6 unicast-routing

interface Ethernet0
 ip address 192.168.99.1 255.255.255.0
 ipv6 address 2001:db8:213:1::/64 eui-64
```

IPv4: 192.168.99.1

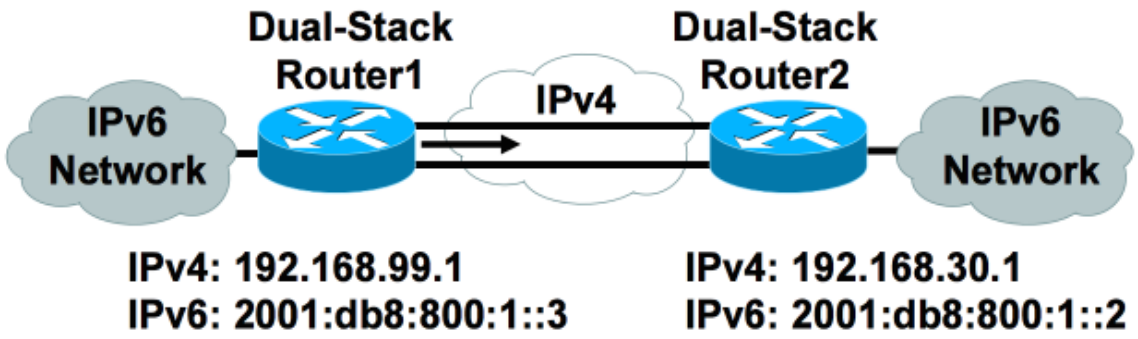
IPv6: 2001:db8:213:1::/64 eui-64

Cisco IOS® Is IPv6-Enable:

- If IPv4 and IPv6 are configured on one interface, the router is dual-stacked
- Telnet, Ping, Traceroute, SSH, DNS client, TFTP, etc.

IPv4 - IPv6 Tunneling

KDDI Myanmar




```
router1#

interface Tunnel0
 ipv6 enable
 ipv6 address 2001:db8:c18:1::3/128
 tunnel source 192.168.99.1
 tunnel destination 192.168.30.1
 tunnel mode gre ipv6
```

```
router2#

interface Tunnel0
 ipv6 enable
 ipv6 address 2001:db8:c18:1::2/128
 tunnel source 192.168.30.1
 tunnel destination 192.168.99.1
 tunnel mode gre ipv6
```

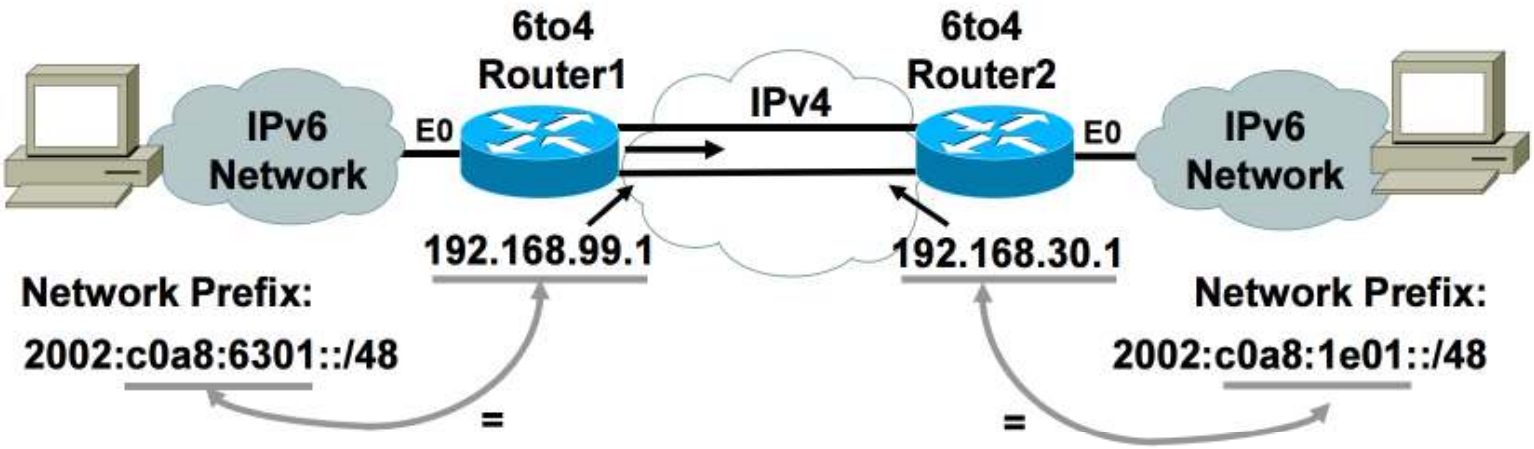
or

```
router1#

interface Tunnel0
 ipv6 enable
 ipv6 address 2001:db8:c18:1::3/127
 tunnel source 192.168.99.1
 tunnel destination 192.168.30.1
 tunnel mode ipv6ip
```

```
router2#

interface Tunnel0
 ipv6 enable
 ipv6 address 2001:db8:c18:1::2/127
 tunnel source 192.168.30.1
 tunnel destination 192.168.99.1
 tunnel mode ipv6ip
```

IPv4 - IPv6 Tunneling

KDDI Myanmar




```
router1#
interface Ethernet0
 ipv6 address 2002:c0a8:6301:1::/64 eui-64
Interface Ethernet1
  ip address 192.168.99.1 255.255.0.0
interface Tunnel0
 ipv6 unnumbered Ethernet0
 tunnel source Ethernet1
 tunnel mode ipv6ip 6to4

ipv6 route 2002::/16 Tunnel0
```

```
router2#
interface Ethernet0
 ipv6 address 2002:c0a8:1e01:1::/64 eui-64
Interface Ethernet1
 ip address 192.168.30.1 255.255.0.0
interface Tunnel0
 ipv6 unnumbered Ethernet0
 tunnel source Ethernet1
 tunnel mode ipv6ip 6to4

ipv6 route 2002::/16 Tunnel0
```

KDDI Myanmar

Do we need a Tunneling technology ?

- **Basically we should not use Tunneling and your network will be more simple by Dual stack especially on ISP network or SP network**
- **However if you have a router which is not support IPv6 on network , then you can not replace to new router due to cost issue or for temporally network , you can also consider Tunneling .**
- **Meaning you can consider all and depend on your requirement or your plan what you want to do .**

KDDI Myanmar

VPN Service

Transport Options for Private IP VPN Service

KDDI Myanmar

Topologies

- Point-point, multi-point
- Full/partial mesh
- Hub/Spoke or Multi-Tier

Media

- Serial, ATM/FR, OC-x
- Dark fiber, Lambda
- Ethernet

"Last Mile" Transport

- L3 – Broadband/WiFi/3G/4G

VPN Offerings

- L2 – Ethernet (p2p, p2mp)
- L3 – Private IP VPN
- L3 – Public IP (Internet)

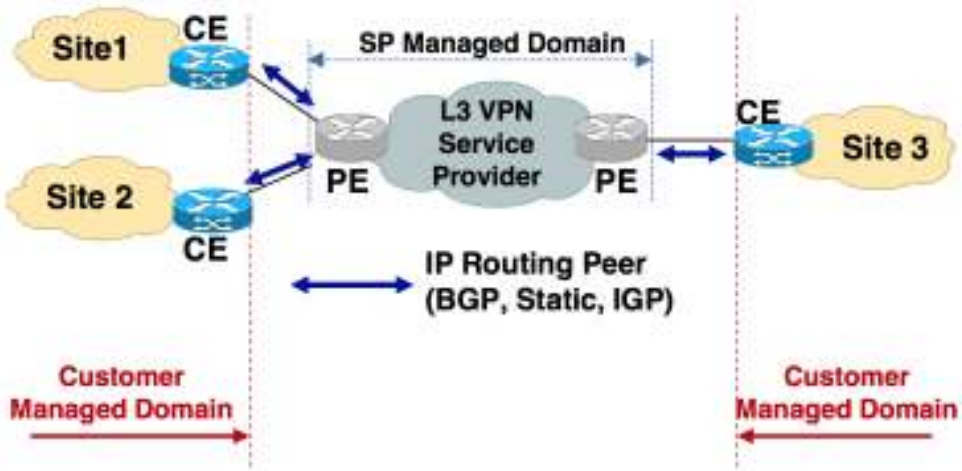
KDDI Myanmar

MPLS VPN Design Trends

- **Single Carrier Designs**:
  - Enterprise will home all sites into a single carrier to provide L3 MPLS VPN connectivity.
  - **Pro:** Simpler design with consistent features
  - **Con:** Bound to single carrier for feature velocity
  - **Con:** Does not protect against MPLS cloud failure with Single Provider
- **Dual Carrier Designs:**
  - Enterprise will single or dual home sites into one or both carriers to provide L3 MPLS VPN connectivity.
  - **Pro:** Protects against MPLS service failure with Single Provider
  - **Pro:** Potential business leverage for better competitive pricing
  - **Con:** Increased design complexity due to Service Implementation Differences (e.g. QoS, BGP AS Topology)
  - **Con:** Feature differences between providers could force customer to use least common denominator features.
- **Variants of these designs and site connectivity:**
  - Encryption Overlay (e.g. IPSec, DMVPN, GET VPN, etc.)
  - Sites with On-demand / Permanent backup links

KDDI Myanmar

SP Managed "IP VPN" Service



- **CE Routers owned by customer**
- **PE Routers owned by SP**
- Customer "peers" to "PE" via IP
  - No labels are exchanged with SP PE
  - No end-to-end visibility of other CE's
- Route exchange with SP done via eBGP/static
- Customer relies on SP to advertise their internal routes to all CE's in the VPN for reachability
- SP can offer multiple services: QoS, multicast, IPv6

Why MPLS ?

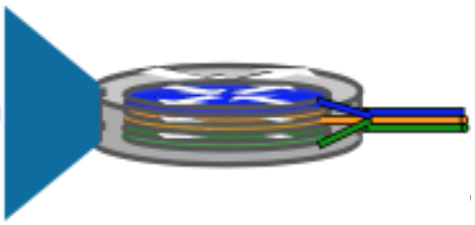
KDDI Myanmar

Large Scale "Virtualization" and Service Enabler in the WAN

- **Layer 3 VPN/Segmentation**
  - VPN (RFC 4364)
  - Provides Any-to-Any connectivity
- Maximize Link Utilization with Selective Routing/Path Manipulation
  - Traffic Engineering
  - Optimization of bandwidth and protection using Fast-ReRoute (FRR)
- **Layer 2 VPN/Transport**
  - AToM (Any Transport over MPLS) i.e. "pseudo-wire"
  - Layer-2 transport: Ethernet, ATM/FR, HDLC/PPP, interworking
  - Layer-2 VPN: VPLS for bridged L2 domains over MPLS
- QoS Capabilities
  - Diffserv, Diffserv aware Traffic Engineering (DS-TE)
- Bandwidth Protection Services
  - Combination of TE, Diffserv, DS-TE, and FRR
- IP Multicast (per VPN/VRF)
- Transport of IPv6 over an IPv4 (Global Routing Table)
- Unified Control Plane (Generalized MPLS)

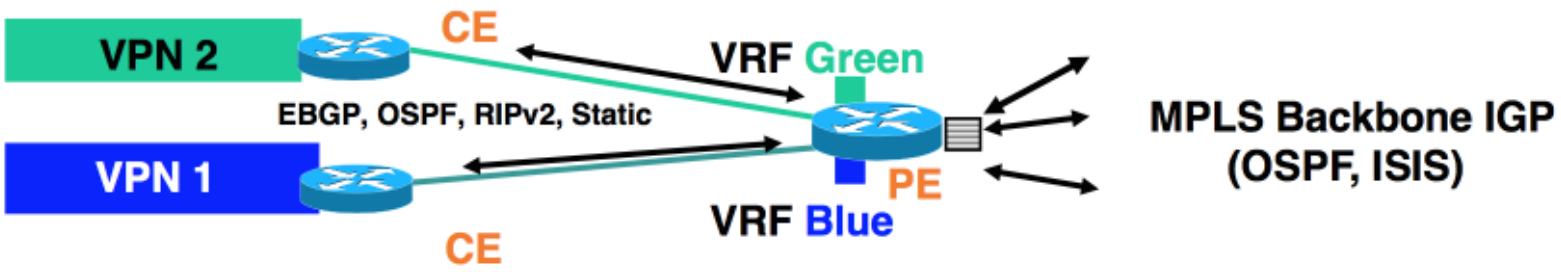
Key virtualization Mechanisms over an IP Infrastructure

- L3 VPN
- L2 VPN (P2P)
- L2 VPN (P2MP)
- QoS
- IPv6, MVPN



KDDI Myanmar

Virtual Routing and Forwarding Instance

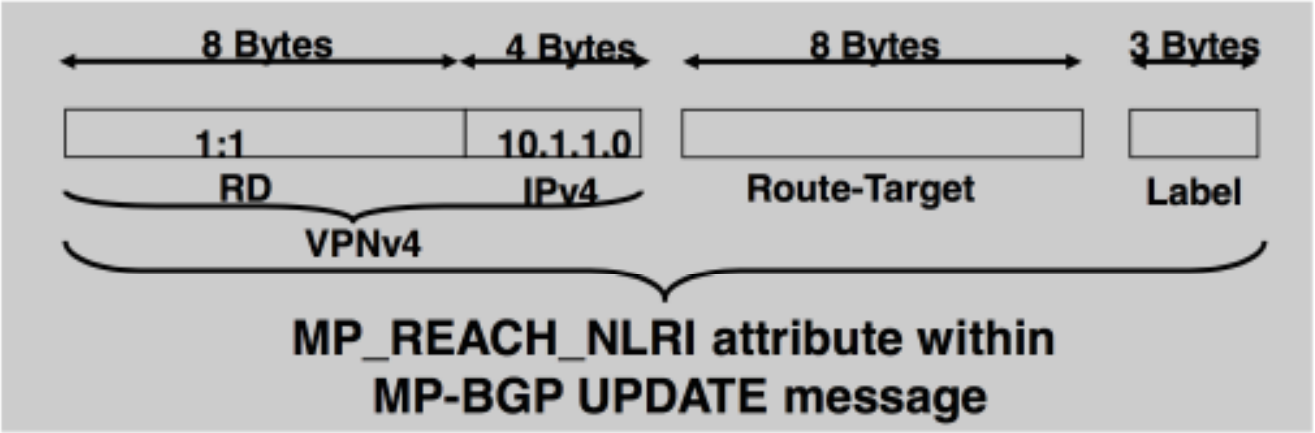


- Associates to one or more interfaces on PE
  - Privatize an interface i.e., color the interface
- Has its own routing table and forwarding table (CEF)
- VRF has its own instance for the routing protocol
  - (static, RIP, BGP, EIGRP, OSPF)
- CE router runs standard routing software
- Allows overlapping address space

KDDI Myanmar

MPLS-VPN Technology

Control Plane – Multiprotocol BGP (MP-BGP)



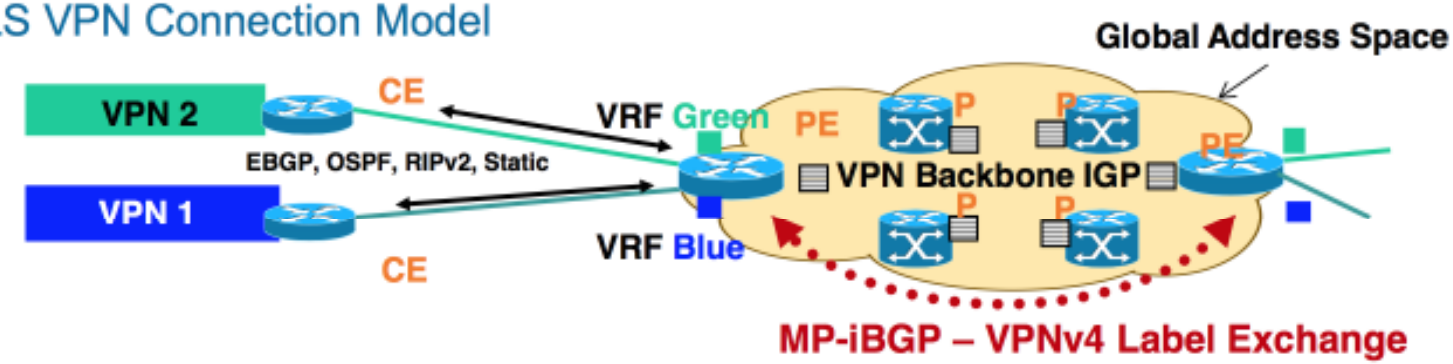
Key Components:

- Route Distinguisher (RD); VPNv4 route
- Route Target (RT)
- Label

KDDI Myanmar

MPLS VPN Technology—Refresher

MPLS VPN Connection Model



CE Routers

- VRF Associates to one or more interfaces on PE
- Has its own routing table and forwarding table (CEF)
- VRF has its own instance for the routing protocol
  (static, RIP, BGP, EIGRP, OSPF)

PE Routers

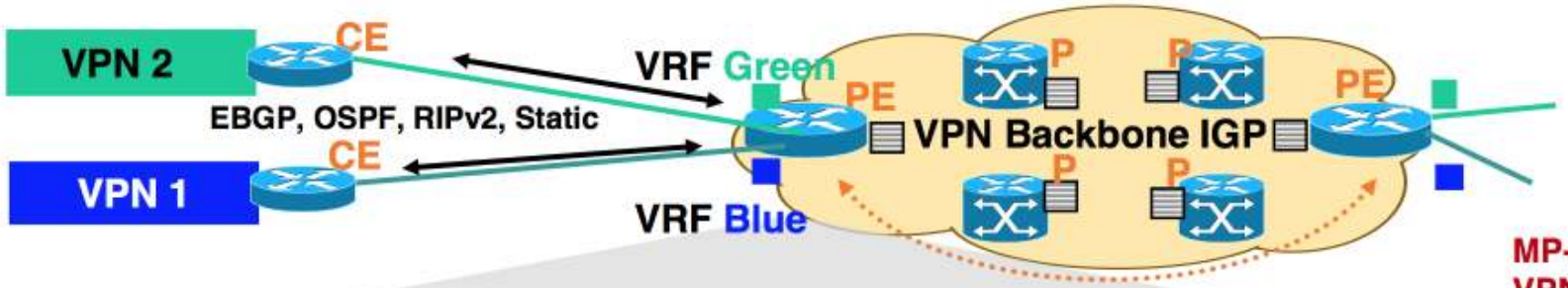
- MPLS Edge routers
- MPLS forwarding to P routers
- IGP/BGP – IP to CE routers
- Distributes VPN information through MP-BGP to other PE routers with VPN-IPv4 updates, VPN labels, etc.

P Routers

- P routers are in the core of the MPLS cloud
- P routers do not need to run BGP
- Do not have knowledge of VPNs
- Switches packets based on labels (push/pop) not IP

Self Deployed MPLS VPN

KDDI Myanmar



MP-iBGP – VPNv4 Label Exchange

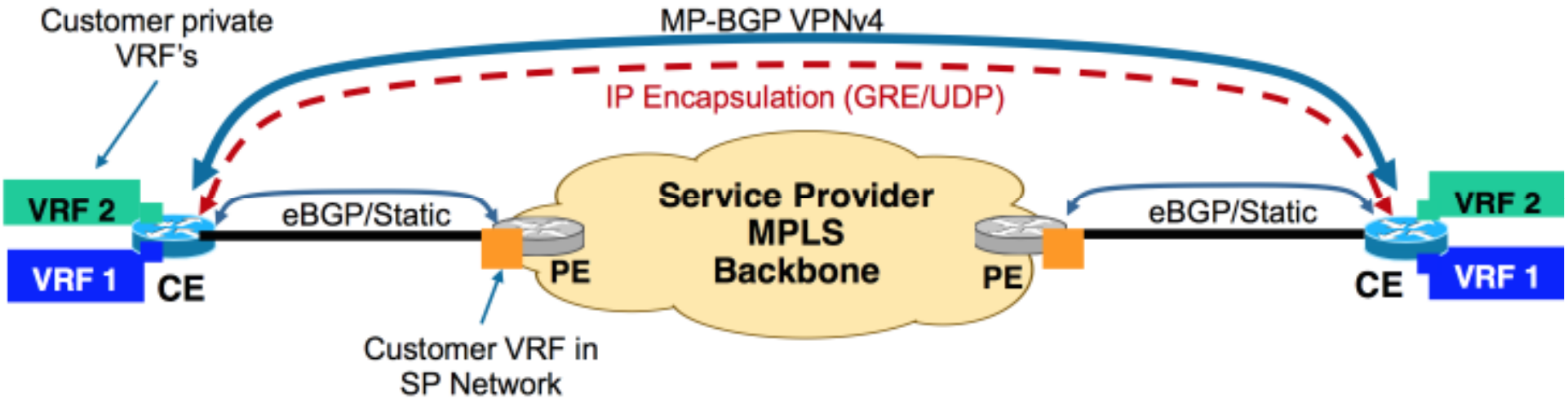
VRF Configuration (PE)

```
! PE Router - Multiple VRFs
ip vrf blue
 rd 65100:10
 route-target import 65100:10
 route-target export 65100:10
ip vrf green
 rd 65100:20
 route-target import 65100:20
 route-target export 65100:20
!
interface GigabitEthernet0/1
 ip vrf forwarding blue
interface GigabitEthernet0/2
 ip vrf forwarding green
```

MP-iBGP Configuration (PE)

```
! PE router
router bgp 65100
 neighbor 192.168.100.4 remote-as 65100
 !
 address-family vpnv4
 neighbor 192.168.100.4 activate
 neighbor 192.168.100.4 send-community extended
 exit-address-family
 !
 address-family ipv4 vrf blue
 neighbor 172.20.10.1 remote-as 65111
 neighbor 172.20.10.1 activate
 exit-address-family
```

KDDI Myanmar



- MPLS VPN or VRF-Lite over IP Encapsulation
- Routing and data forwarding done “Over the Top” of the SP transport
- Enterprise routing - exchanged either inside IP tunnel, and/or over the top (BGP)
- Routing to SP – BGP/static and minimal (typically IP tunnel end-points)
- Multicast can be supported either (1) leveraging the SP service, or (2) inside the IP tunnel

MPLS VPN Technology (L2VPN)

KDDI Myanmar

E-Line (Point-to-Point)

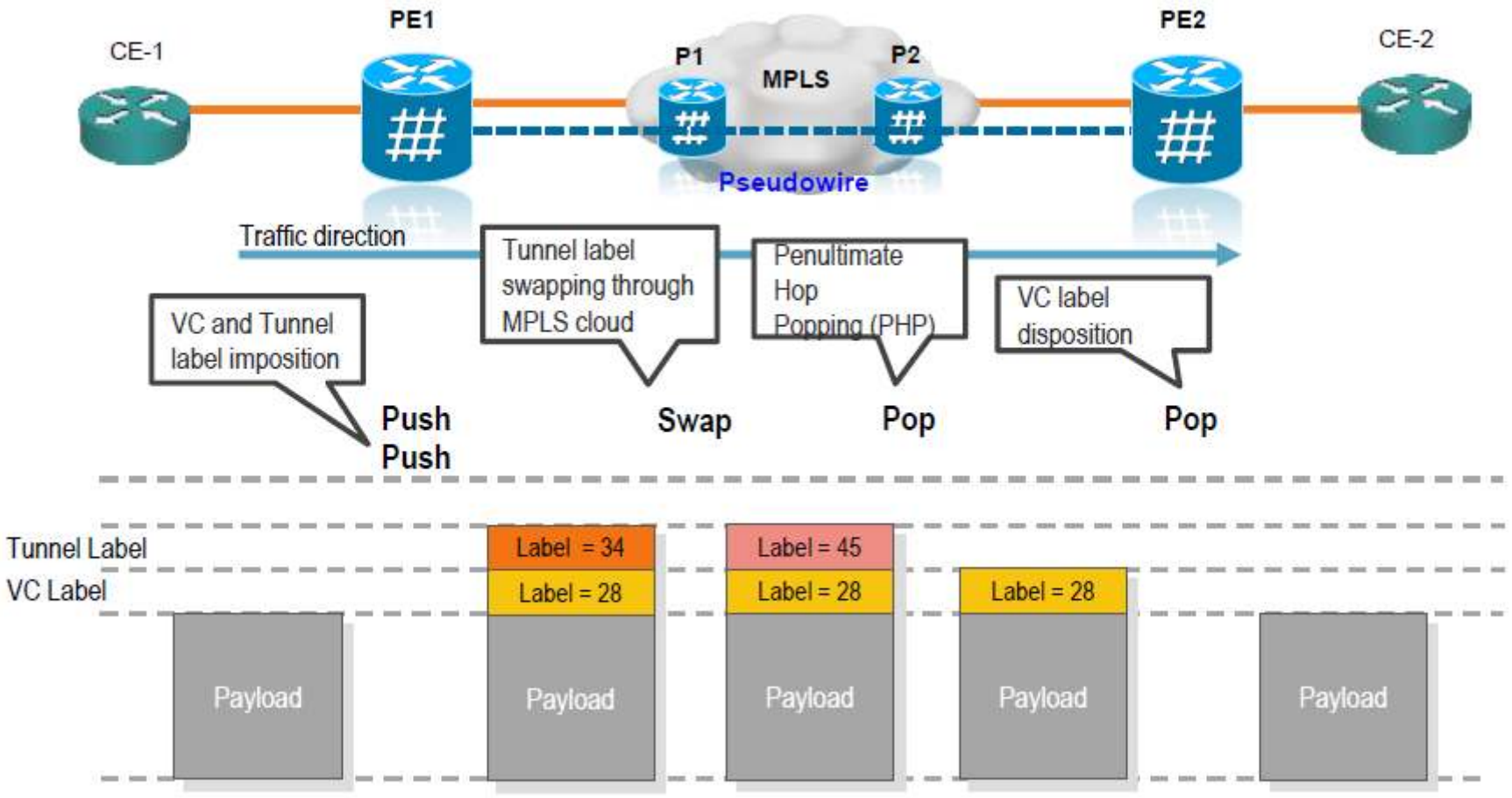
- Replaces TDM private line
- Point-to-point EVCs offer predictable performance for applications
- One or more EVCs allowed per single physical interface (UNI)
- Ideal for voice, video, and real-time data

E-LAN (Point-to-Multipoint)

- Offers point to multipoint for any-to-any connectivity
- Transparent to VLANs and Layer 2 control protocols
- 4 or 6 classes of QoS support
- Ideal for LAN-to-LAN bulk data

MPLS VPN Technology (L2VPN : P2P)

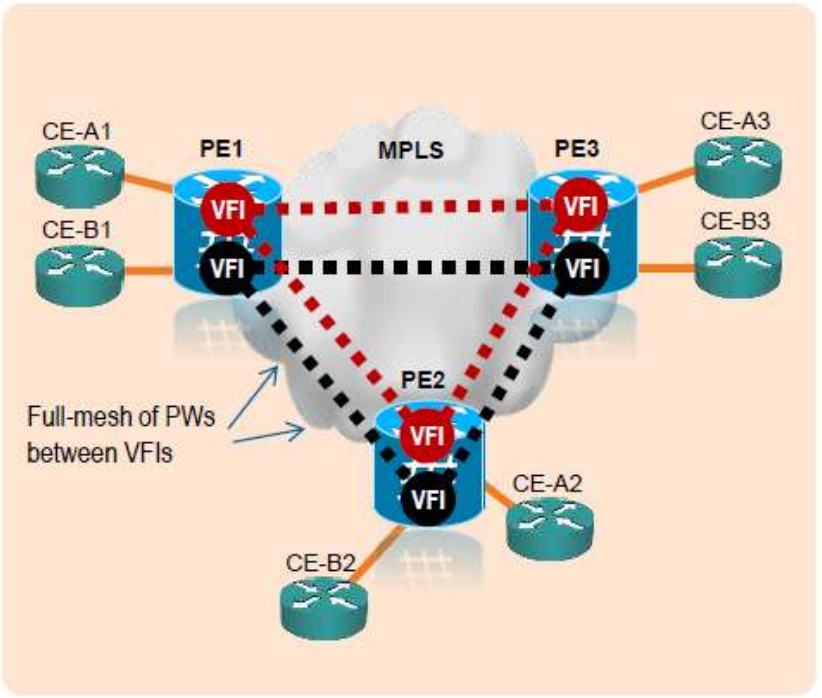
KDDI Myanmar



MPLS VPN Technology (L2VPN : P2M)

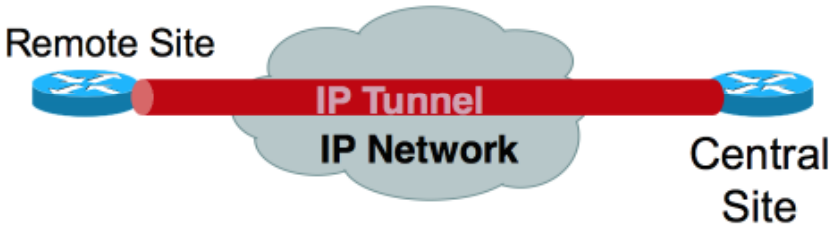
KDDI Myanmar

- VFI (Virtual Forwarding Instance)
  - Also called VSI (Virtual Switching Instance)
  - Emulates L2 broadcast domain among ACs and VCs
  - Unique per service. Multiple VFIs can exist same PE
- AC (Attachment Circuit)
  - Connect to CE device, it could be Ethernet physical or logical port
  - One or multiple ACs can belong to same VFI
- VC (Virtual Circuit)
  - EoMPLS data encapsulation, tunnel label used to reach remote PE, VC label used to identify VFI
  - One or multiple VCs can belong to same VFI
  - PEs must have a full-mesh of PWs in the VPLS core



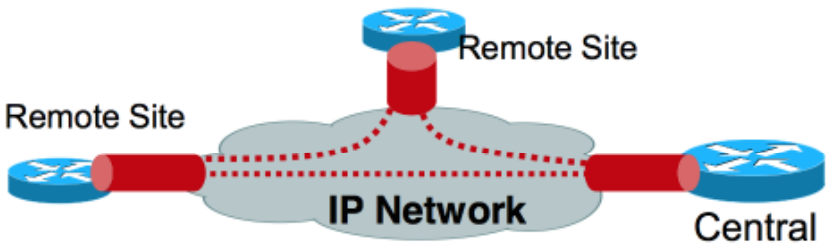
"Stateful" vs. "Stateless" GRE Tunneling

Point-to-Point GRE



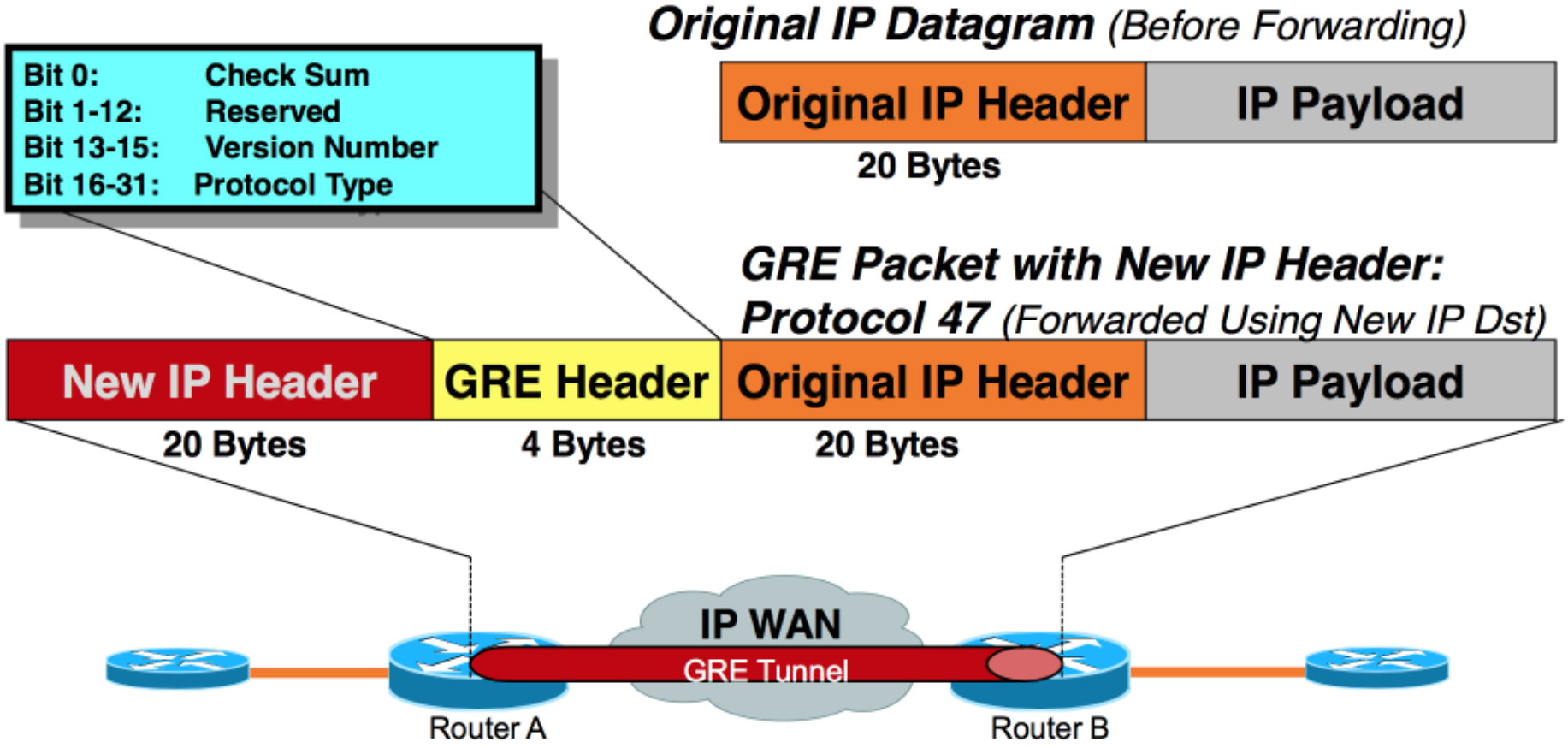
- Source and destination requires manual configuration
- Tunnel end-points are stateful neighbors
- Tunnel destination is explicitly configured
- Creates a logical point-to-point "Tunnel"

Multipoint GRE

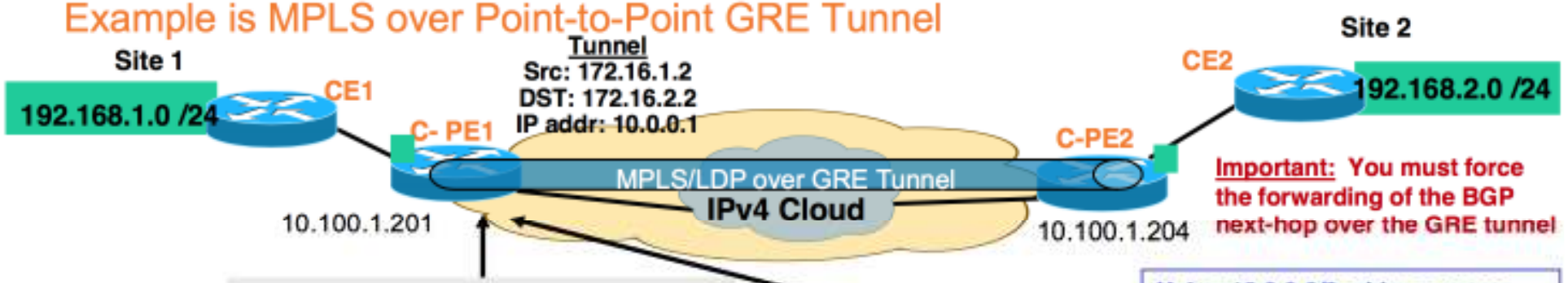


- **Single** multipoint tunnel interface is created per node
- Only the tunnel source is defined
- Tunnel destination is derived dynamically through some signaling mechanism
  - DMVPN – uses NHRP
  - MPLS VPN over mGRE – uses BGP
- Creates an "encapsulation" using IP headers

GRE Tunnel Encapsulation (RFC 2784)



Example is MPLS over Point-to-Point GRE Tunnel




```
ip vrf green
 rd 1:1
 route-target export 1:1
 route-target import 1:1
!
mpls label protocol ldp
!
interface Tunnel0
 ip address 10.0.0.1 255.255.255.0
 ip mtu 1400
 mpls ip
 tunnel source 172.16.1.2
 tunnel destination 172.16.2.2
 tunnel path-mtu-discovery
!
interface Loopback0
 ip address 10.100.1.201 255.255.255.255
```

Enables MPLS/LDP over GRE

```
router eigrp 1
 network 10.0.0.0
 no auto-summary
 !
router bgp 65000
 bgp router-id 10.100.1.201
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 neighbor 10.100.1.204 remote-as 65000
 neighbor 10.100.1.204 update-source Loopback0
!
 address-family vpnv4
 neighbor 10.100.1.204 activate
 neighbor 10.100.1.204 send-community extended
 exit-address-family
```

Using 10.0.0.0/8 address space Forces Loopback 0 learning over GRE Tunnel

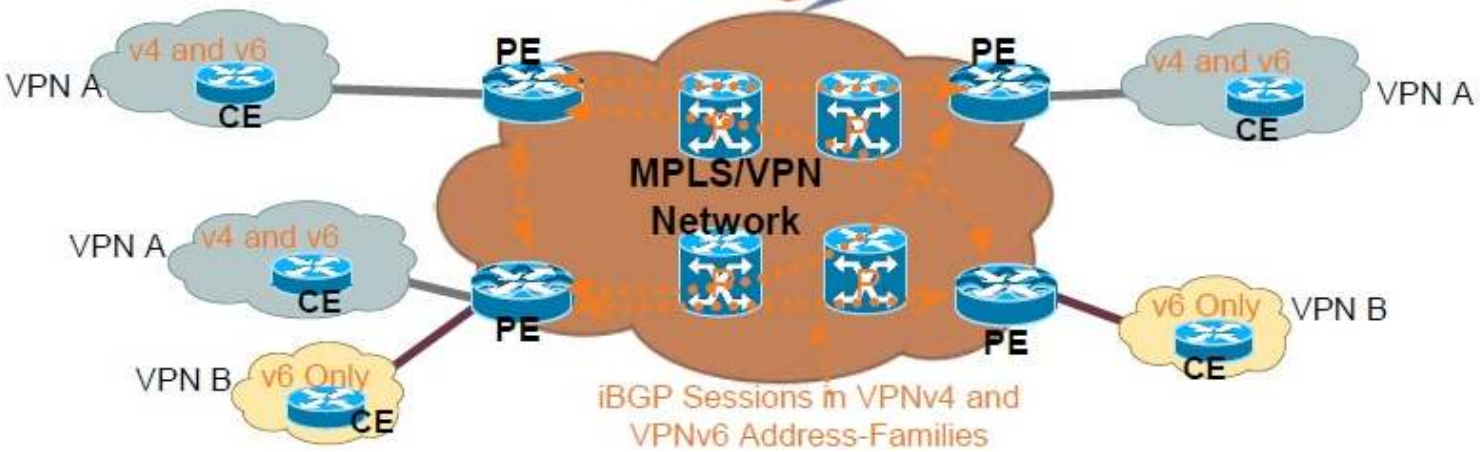
Compare a configuration btw IOS , XR , NXOS

KDDI Myanmar

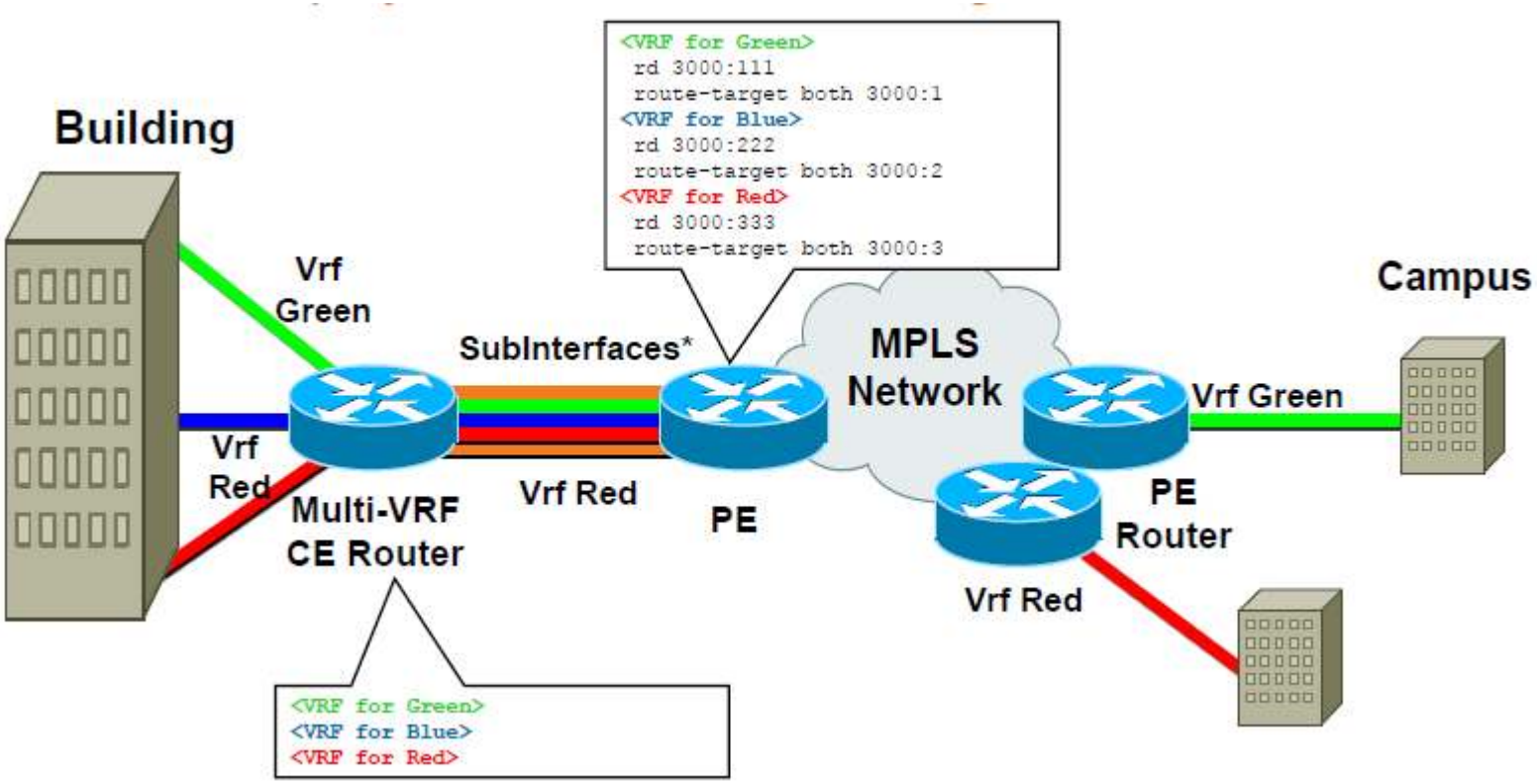
```
IOS_PE#
!
vrf definition v2
 rd 2:2
 !
 address-family ipv6
  route-target export 2:2
  route-target import 2:2
!
router bgp 1
!
 address-family vpnv6
  neighbor 10.13.1.21 activate
  neighbor 10.13.1.21 send-community both
!
 address-family ipv6 vrf v2
  neighbor 200::2 remote-as 30000
  neighbor 200::2 activate
!
```

```
IOS-XR_PE#
!
vrf v2
!
 address-family ipv6 unicast
  route-target export 2:2
  route-target import 2:2
!
router bgp 1
 address-family vpnv6 unicast
 !
 neighbor 10.13.1.21
  remote-as 30000
  address-family vpnv6 unicast
!
 vrf v2
  rd 2:2
  address-family ipv6 unicast
  !
  neighbor 200::2
   remote-as 30000
   address-family ipv6 unicast
   !
```

```
NXOS_PE#
!
vrf context v2
 rd 2:2
 !
 address-family ipv6 unicast
  route-target export 2:2
  route-target import 2:2
!
router bgp 1
  neighbor 10.13.1.21
   remote-as 1
   update-source loopback0
   address-family vpnv6 unicast
      send-community extended
  !
  vrf vpn1
    neighbor 200::2
    remote-as 30000
      address-family ipv6 unicast
 !
```



Case diagram for VPN network



http://mpls-configuration-on-cisco-ios-software.org.ua/1587051990/ch06lev1sec1.html

(*) Can not communicate btw different colors !

MPLS VPN Technology (L2VPN vs L3VPN)

KDDI Myanmar

MPLS Layer 3 Service

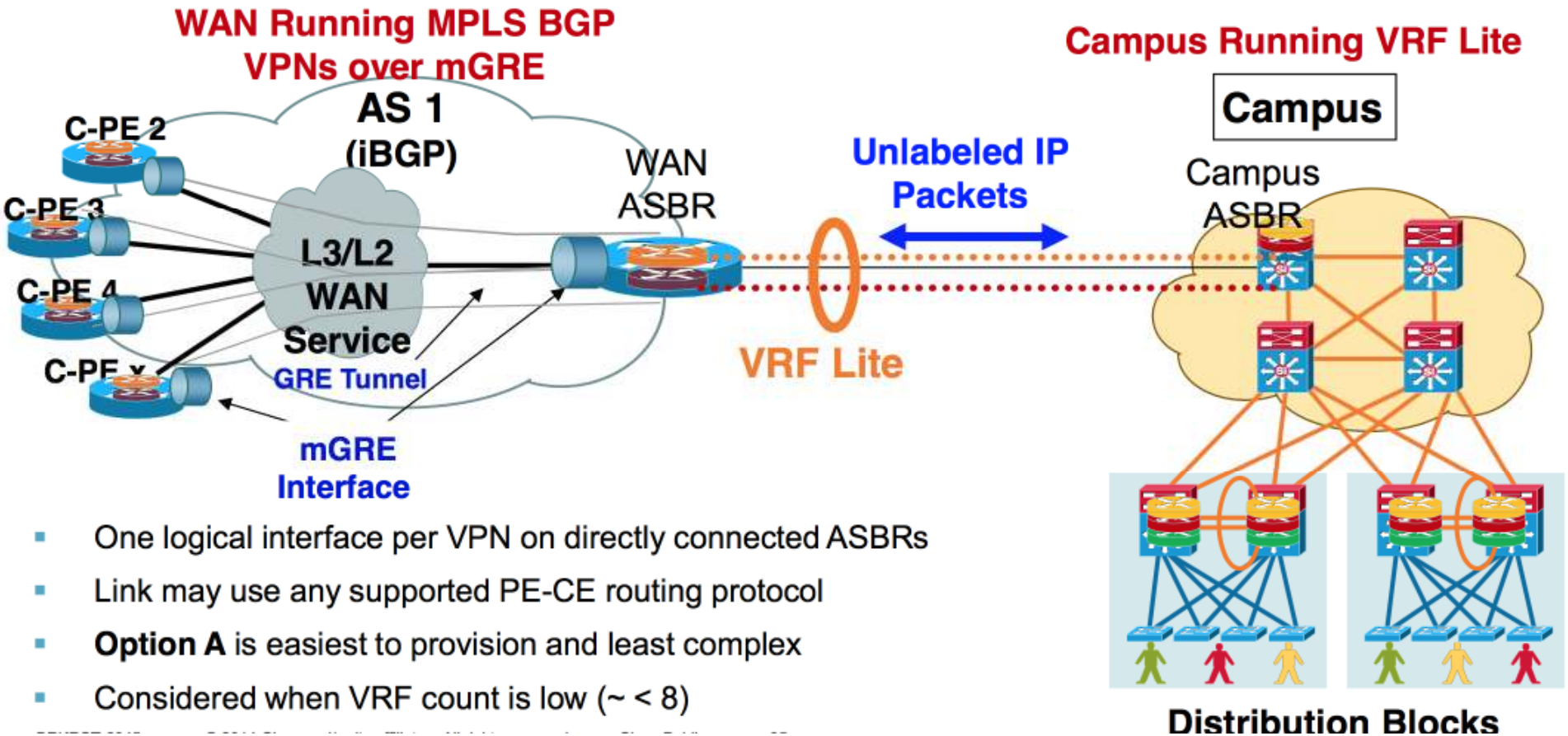
- Routing protocol dependent on the carrier
- Layer 3 capability depends on carrier offering
  - QoS (4 classes/6 classes)
  - IPv6 adoption
- Transport IP protocol only
- Peering with carrier for routing protocol adjacency

MetroE Layer 2 Service

- Routing protocol independent of the carrier
- Customer manages layer 3 QoS
- Capable of transport IP and none-IP traffic.
- Routing protocol scalability in point-to-multipoint topology

KDDI Myanmar

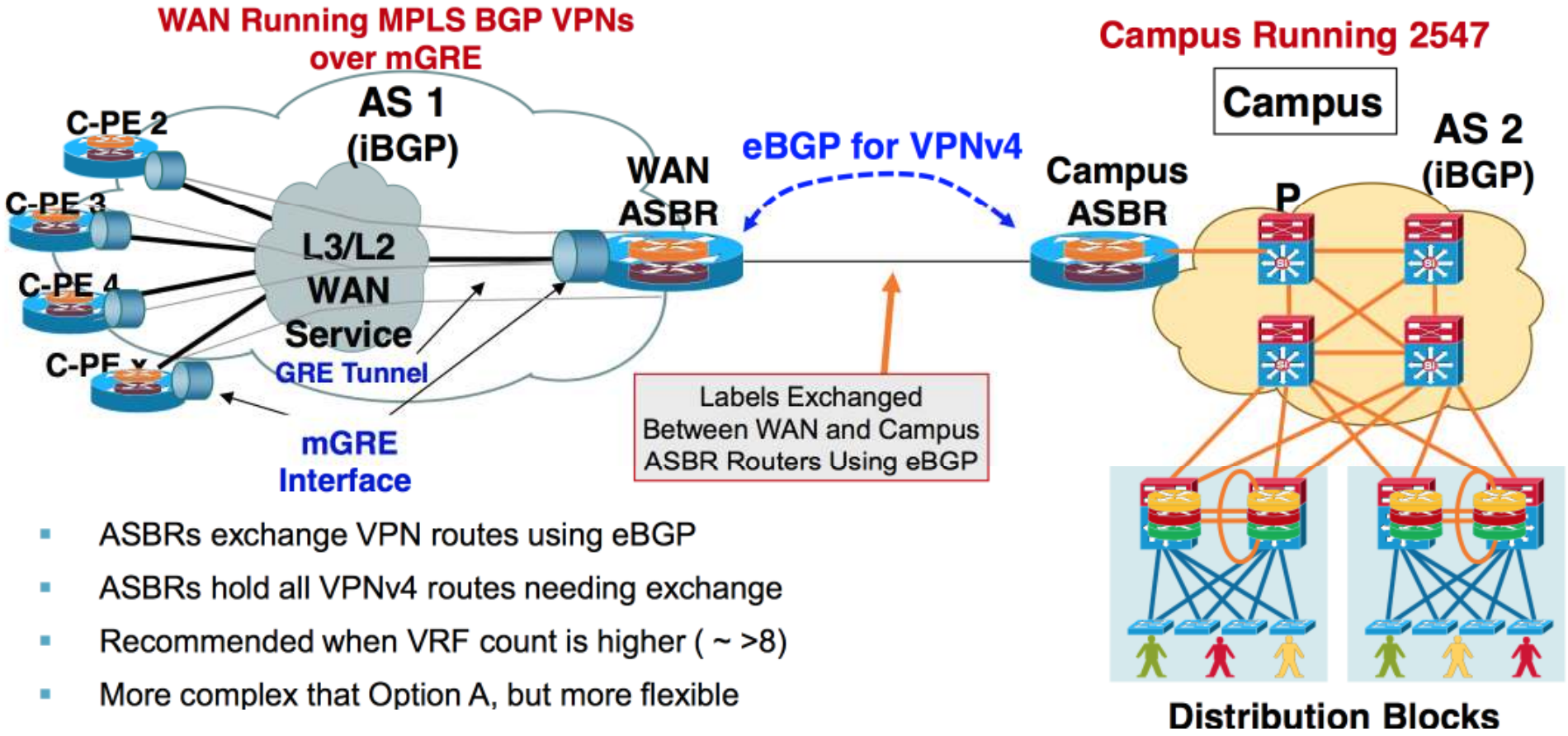
Inter AS Option A (Back to Back VRFs)



- One logical interface per VPN on directly connected ASBRs
- Link may use any supported PE-CE routing protocol
- **Option A** is easiest to provision and least complex
- Considered when VRF count is low (~ < 8)

KDDI Myanmar

Inter AS Option B (Medium/Large VRF Deployments)

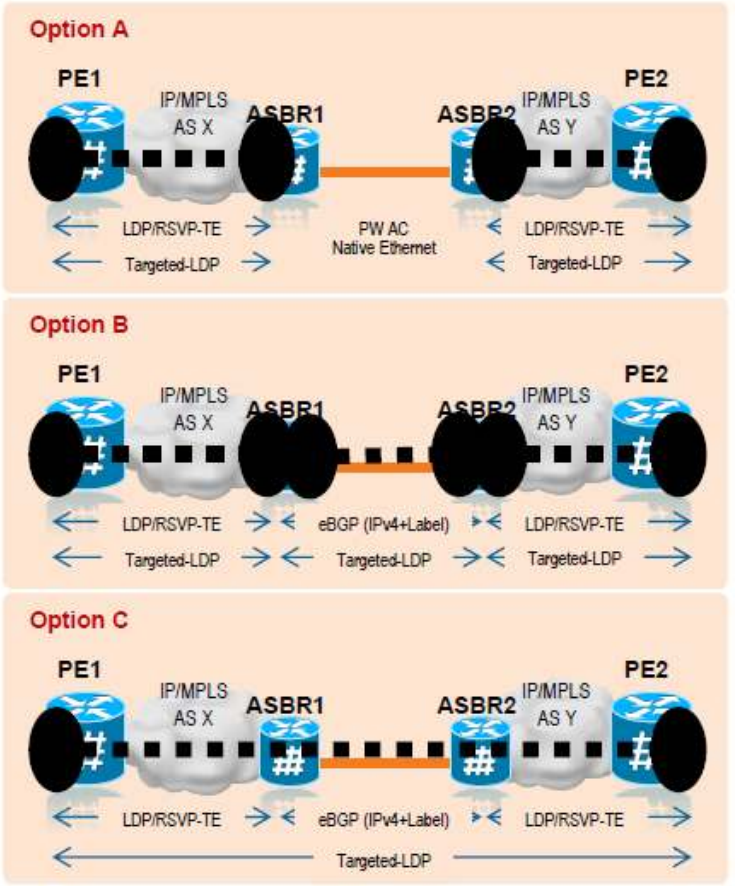


- ASBRs exchange VPN routes using eBGP
- ASBRs hold all VPNv4 routes needing exchange
- Recommended when VRF count is higher (~ >8)
- More complex that Option A, but more flexible

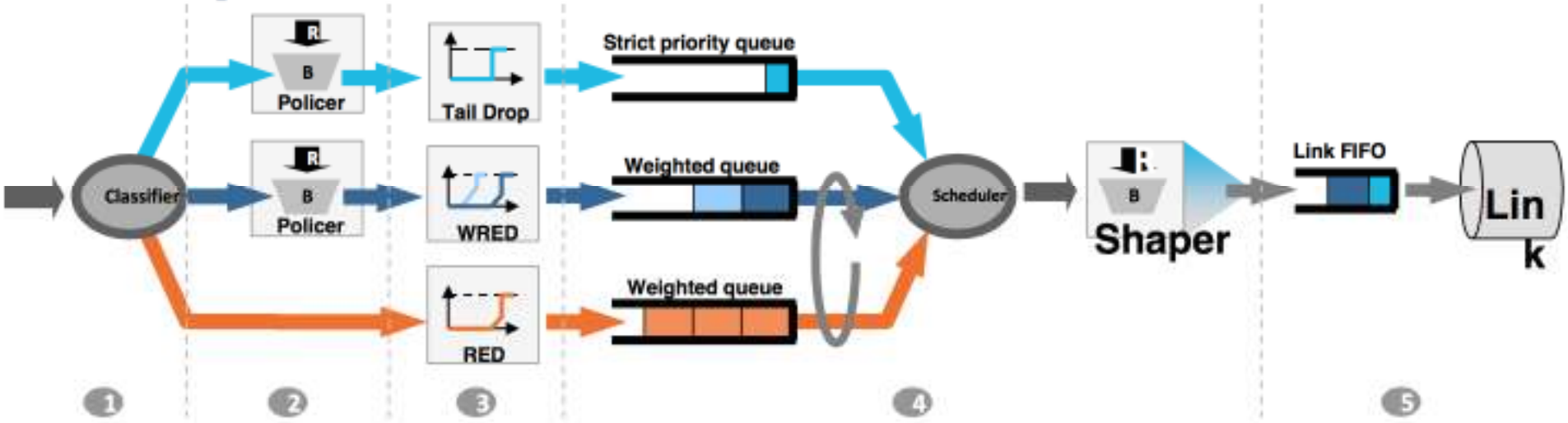
NNI (Inter AS) for L2VPN

KDDI Myanmar

- Three (3) deployment models
- Option A
  - No reachability information shared between AS
- Option B
  - Minimal reachability information shared between AS
  - ASBR configured as S-PEs (multi-segment PWs)
  - eBGP (IPv4 prefix + label) used to build PSN tunnel between AS
- Option C
  - Significant reachability information shared between AS
  - Single-segment PW signaled across AS boundary



Components of QoS

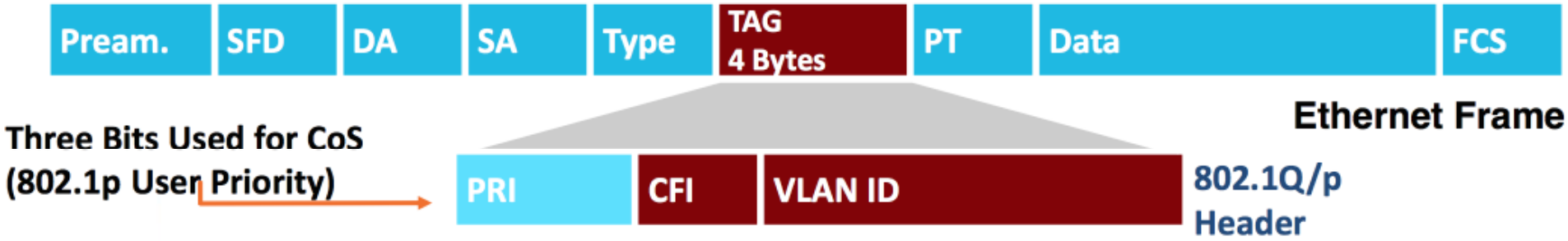


1. Classification and Marking - CoS, DSCP, Port Num, Packet Len, Protocol, VLAN etc
2. Admission Control - Local, Measurement and Resource Based (CAC and RSVP).
3. Policing - Pre Queuing includes Marking, Policing, Dropping (Tail Drop and WRED)
4. Queuing and Scheduling – Priority, Queue Length (Buffers)
5. Shaping – generally outbound, also sharing.

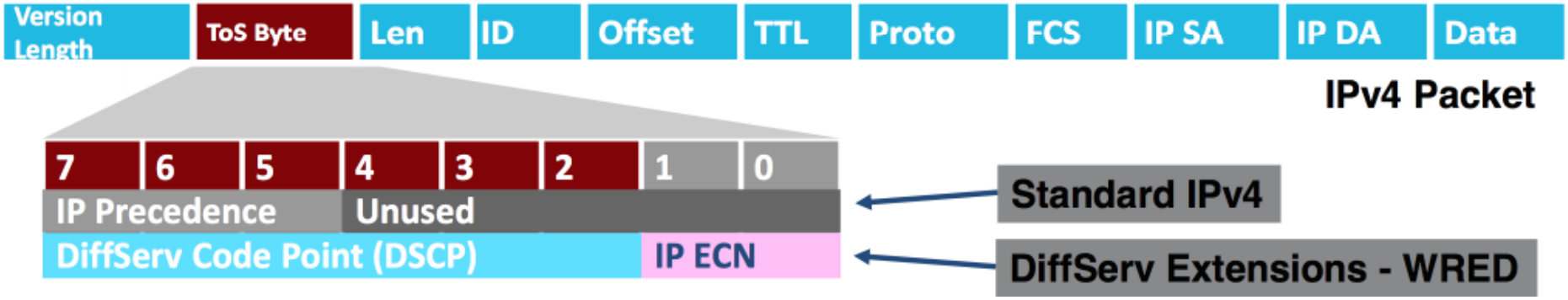
(*) It is general info , however specification is deferent per each model

Layer 2- Ethernet 802.1Q Class of Service

DSCP is backward-compatible with IP precedence



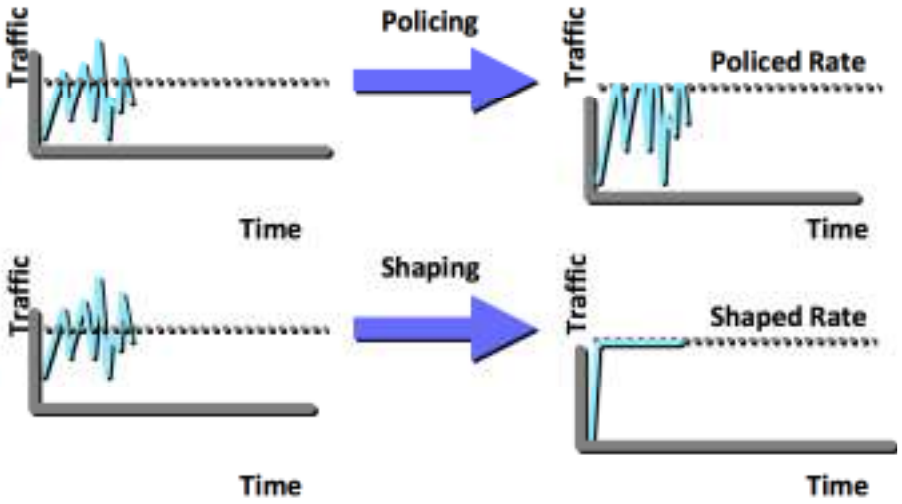
Layer 3- IP Precedence and DiffServ Code Points



(*) It is general info , however specification is deferent per each model

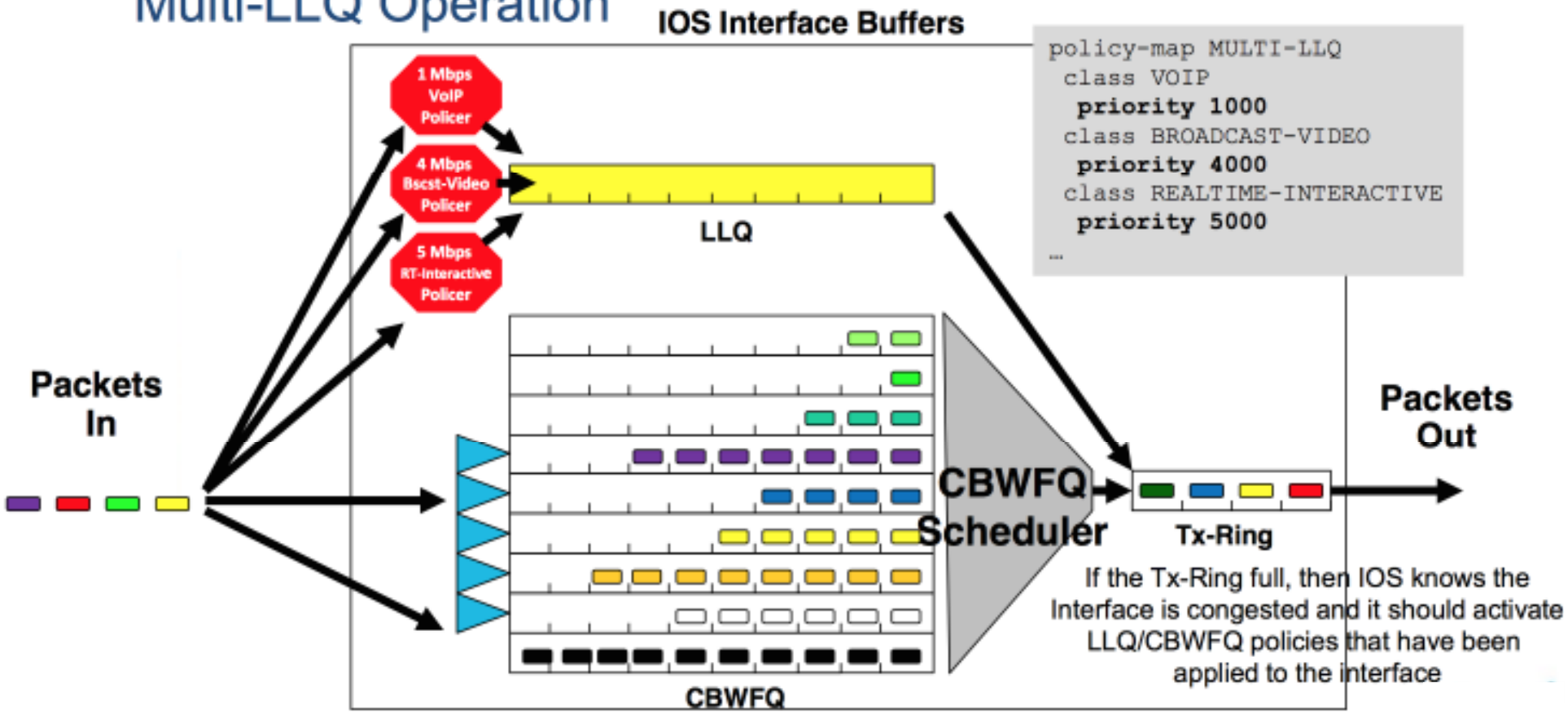
Policing vs. Shaping

- Policing typically drops out-of-contract traffic
- Effectively policing acts to cut the peaks off bursty traffic
- Shaping typically delays out of contract traffic
- Shaping acts to smooth the traffic profile by delaying the peaks
  - Resulting packet stream is “smoothed” and net throughput for TCP traffic is higher with shaping
  - Shaping delay may have an impact on some services such as voip and video



IOS QoS Mechanisms and Operation

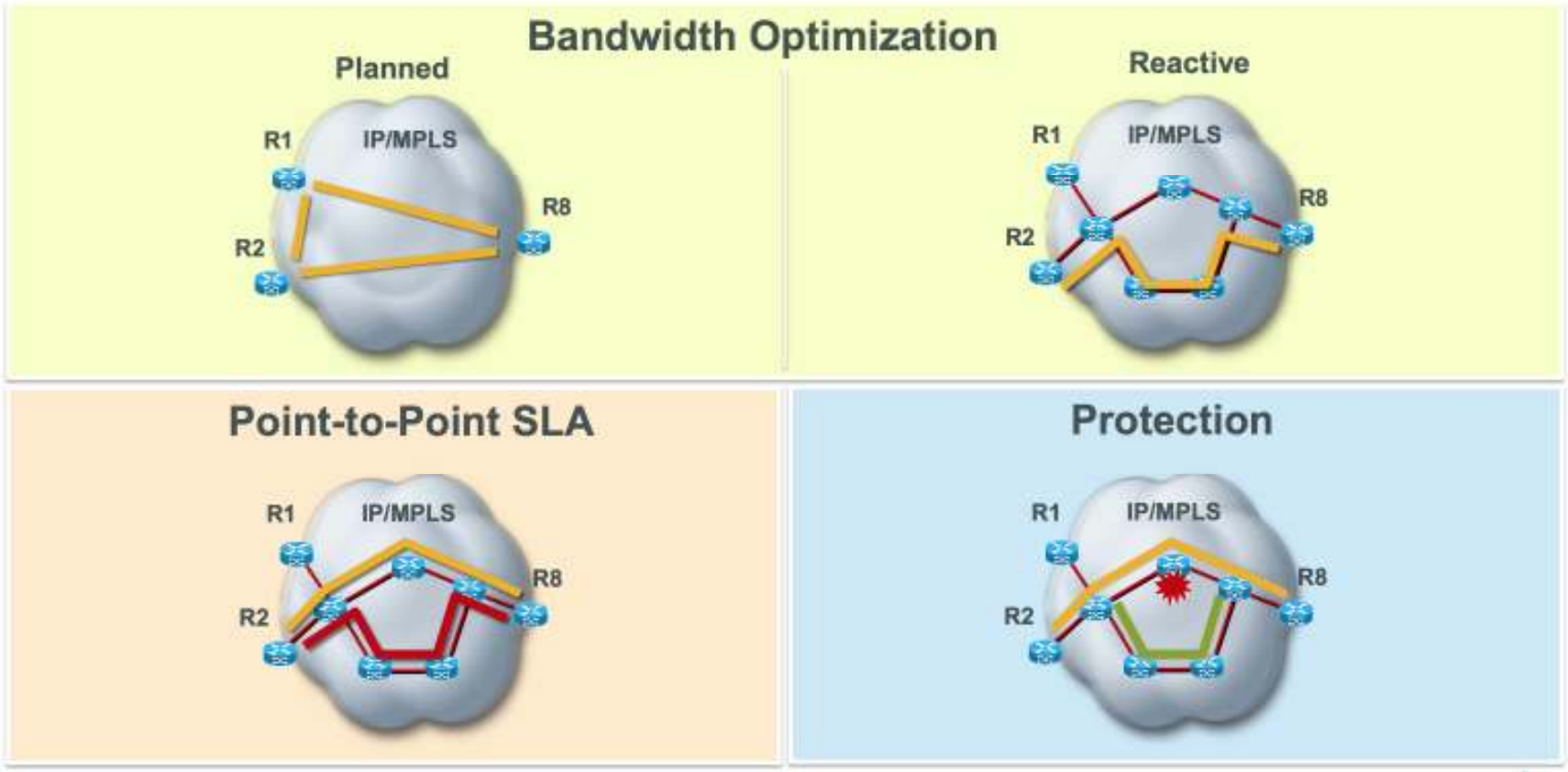
Multi-LLQ Operation



(*) It is general info , however specification is deferent per each model

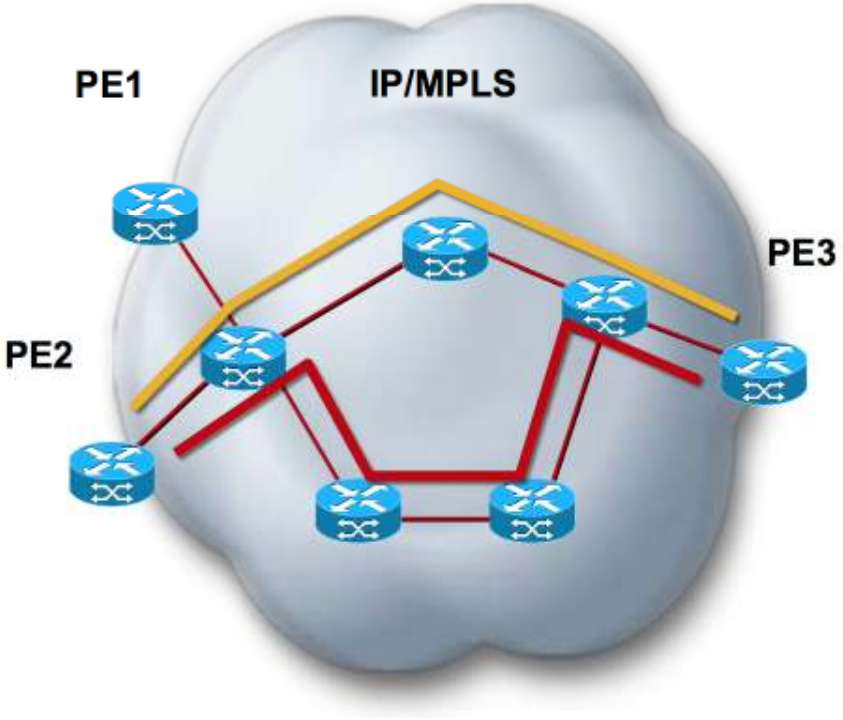
KDDI Myanmar

MPLS TE Deployment Models



KDDI Myanmar

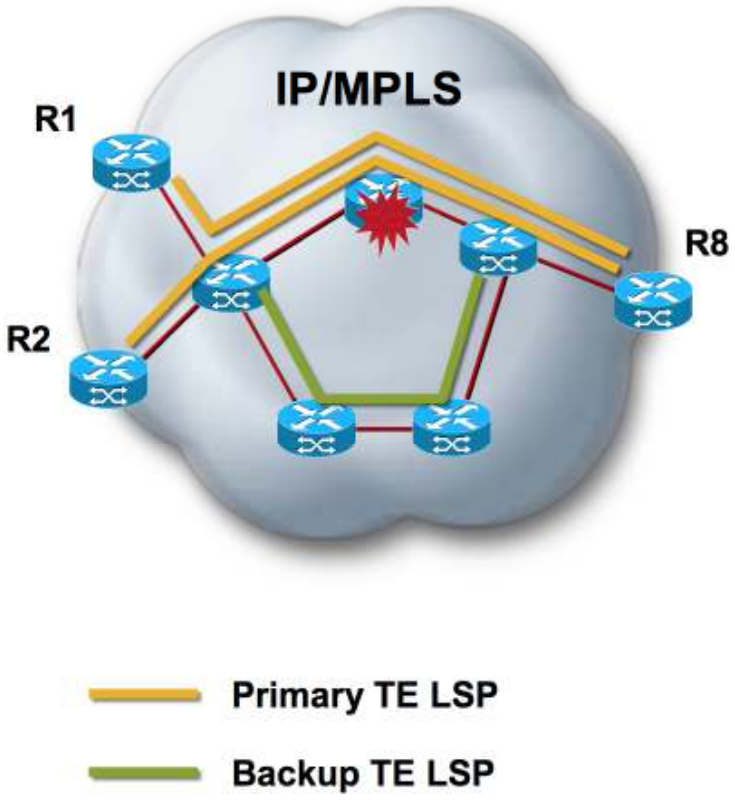
Motivations



- Point-to-point SLAs
- Admission control
- Integration with DiffServ architecture
- Increased routing control to improve network performance

KDDI Myanmar

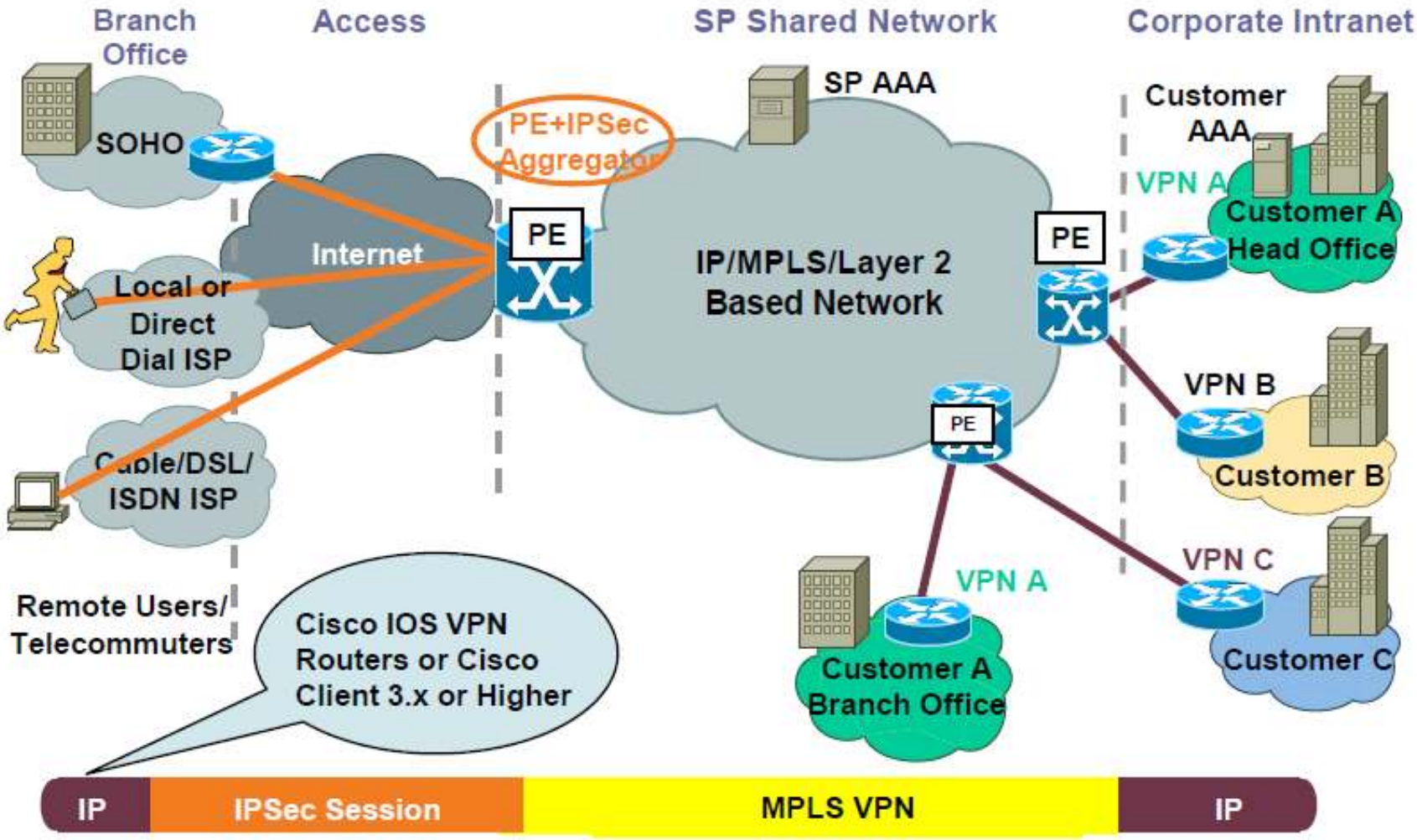
Traffic Protection Using MPLS TE Fast Re-Route (FRR)



- Sub-second recovery against node/link failures
- Scalable 1:N protection
- Greater protection granularity
- Cost-effective alternative to 1:1 protection
- Bandwidth protection

Other (how to connect between Internet VPN and MPLS VPN)

KDDI Myanmar



KDDI Myanmar

Suggestion to MPT

We can also suggest and provide these technology like a MPLS-TE , Internet VPN , VPLS(PtoM L2VPN) ,NNI (for global VPN service) , IPv6 and if you are interested it , please feel free contact to us .

Configuration and troubleshooting

You can refer to OJT documents for understand more detail that we provided you on phase 2 project .

End of all…

KDDI Myanmar

One point advice for you

End of all …

Q : How can be professional ?

A : Practice and experience is most important (certification is not important , but much better)

just try a configuration many time ,then your skill will be improved skill little by little

End of all …

KDDI Myanmar

Q : How can we practice ?

A : We can offer speciall bootcamp for you .
If you are interested it , please feel free contact to us .

Q : What is the bootcamp ?

A : you can use our Lab (we have XE , XR , NE device and we can provide a case study , then you can try step by step . After completed this session , you will be CCIE level ,maybe … ☺

Q : Expensive it ?

A : No , we can offer it by special rate and you can contact to our sales .

End of all …

KDDI Myanmar

Q : How many course do you have ?

A : we can provide all (design , configuration , operation , troubleshooting ,etc..) for L3 network like ISP network (internet service) and SP network (VPN service) technology like IPv4 and IPv6 routing and switching, IGP and EGP protcol detail , internet VPN , IPVPN , MPLS , multicast , IPSec , mGRE , etc …
Meaning we can customize a course for meet your requirement .

Q : where is a Lab location ?

A : Australia(Sydney) or India(Bangalore) or use a simulator at on-site (Myanmar) or Singapore

KDDI Myanmar

Thank you very much for your kind participation today and See you on … !