SANOG

South Asian Network Operators Group

Network Operation
Tips and Tricks

Simon Sohel Baroi
Fiber@Home Ltd

Md. Zobair Khan
Fiber@Home Ltd

Case Studies :

1. TCP MSS Tweaks
2. MPLS L2 VPN Tweaks
3. IPv6 Subnetting
4. Prefix Announcement : Use of Community
5. Route Redistribution
6. Router Security (IPv4/IPv6)
7. Route Optimization

TCP MSS Tweaks

TCP MSS Tweaks

Assumption :

- ISP Infrastructure has MPLS Network.
- Upstream Provider has MPLS Network in between some hops.

TCP MSS Tweaks

Problem :

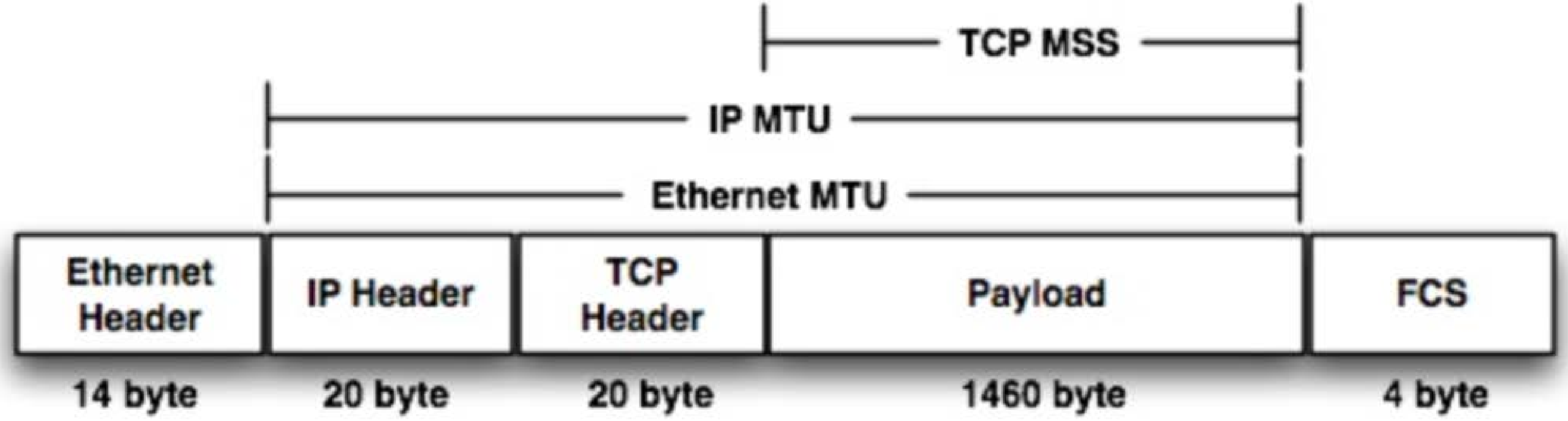
- Users were not able to access most of the WWW contents
- Users were not able to perform e-Mail transactions with or without attachments
- Both IP VPN and MPLS L3 VPN users faced similar problems with site-to-site data traffic

TCP MSS Tweaks

Why :

- Maximum Transmission Unit (MTU) is 1500 by default for Ethernet excluding ethernet headers & trailers

MSS adjustment process:



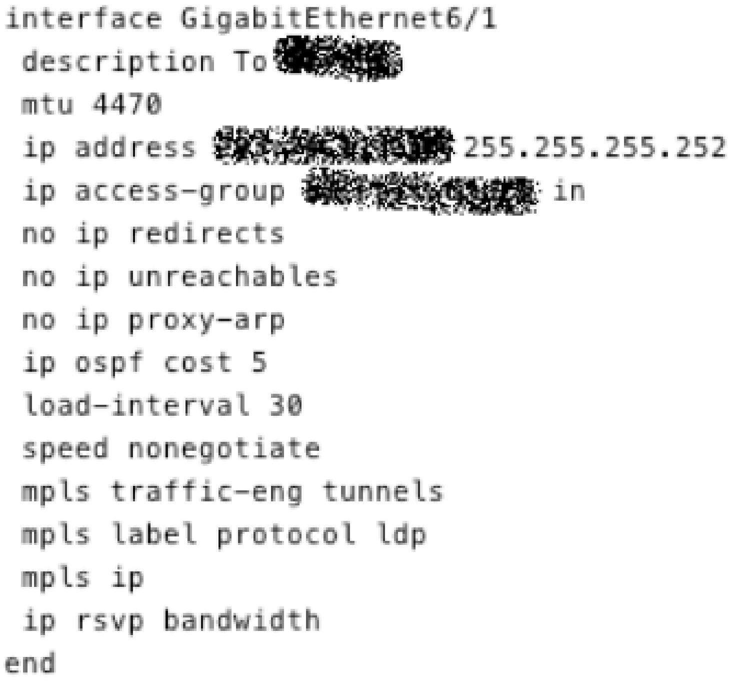
TCP MSS Tweaks

Also :

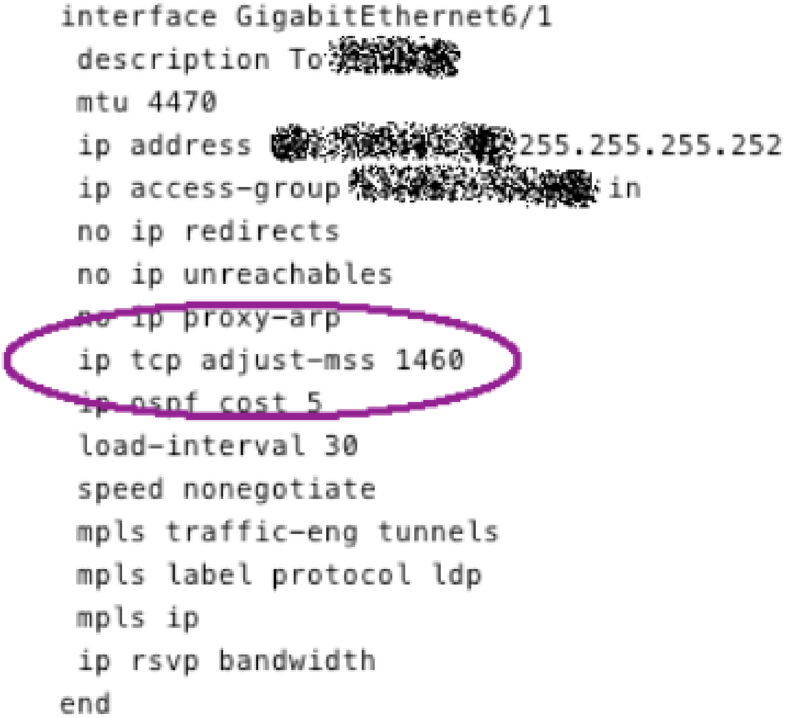
- We can’t increase IP MTU of ethernet interface because
  - if a node construct a full size packet and then with MPLS encapsulation the maximum frame size exceed the 1500 bytes.
- By using **TCP MSS** adjustment, nodes can be signaled to reduce the payload size.

TCP MSS Tweaks

Peering Interface CFG ***before/after MSS*** tweaking:



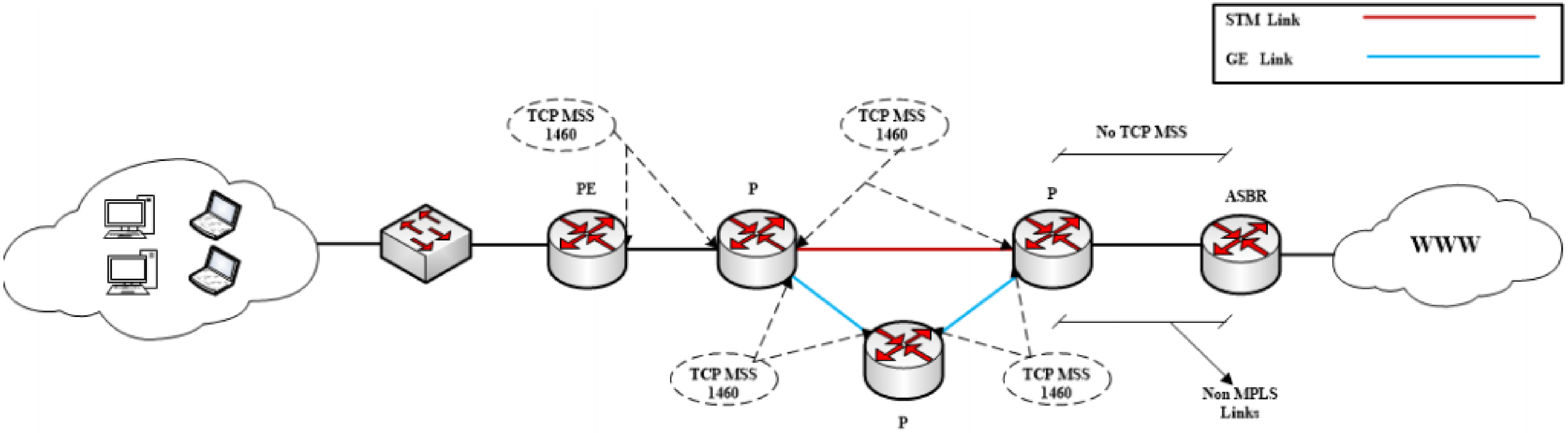
```
interface GigabitEthernet6/1
 description To [redacted]
 mtu 4470
 ip address [redacted] 255.255.255.252
 ip access-group [redacted] in
 no ip redirects
 no ip unreachables
 no ip proxy-arp
 ip ospf cost 5
 load-interval 30
 speed nonegotiate
 mpls traffic-eng tunnels
 mpls label protocol ldp
 mpls ip
 ip rsvp bandwidth
end
```



```
interface GigabitEthernet6/1
 description To [redacted]
 mtu 4470
 ip address [redacted] 255.255.255.252
 ip access-group [redacted] in
 no ip redirects
 no ip unreachables
 no ip proxy-arp
 ip tcp adjust-mss 1460
 ip ospf cost 5
 load-interval 30
 speed nonegotiate
 mpls traffic-eng tunnels
 mpls label protocol ldp
 mpls ip
 ip rsvp bandwidth
end
```

TCP MSS Tweaks

Where to Implement :



TCP MSS Tweaks

Router CPU Problem :

- Packet Per Second will Increase.

Solution :

- Monitoring CPU Load.
  - Observium.

MPLS L2 VPN Tweak

MPLS L2 VPN Tweaks

MPLS L2 VPN : Requirements

- End-to-End **Jumbo Frame** support across the ISP backbone.
- End-to-End **Error free Full Duplex** Links

MPLS L2 VPN Tweaks

MTU:

Maximum Transmission Unit: default 1500 bytes

Jumbo Frames: Frames which are larger than standard 1500 bytes

A simple peak at what goes through the wire:

- 14 bytes: Ethernet Header
- 20 bytes: IP Header
- 20 bytes: Transport Header
- 1500 bytes: Max. Data Payload
- 4 bytes: FCS (or in other words 32 bit CRC - Ethernet Trailer)

MPLS L2 VPN Tweaks

<u>The simple math:</u>

Total Header Size: 58 bytes max.

Payload Size: 1500 bytes max.

Hence in full load a frame may hit 1558 bytes.

So we already have exceeded MTU by 58 bytes. And this is just traditional frame without MPLS.

Activating MPLS adds more header bytes.

MPLS L2 VPN Tweaks

MPLS Headers:

- **4 bytes:** MPLS LDP Header

- **4 bytes:** MPLS L3/L2 VPN Header

- **4 bytes:** MPLS TE Header (only if MPLS TE is active)

Therefore, we end up with a Frame size of:

1558 + 4 + 4 + 4 = 1570 bytes at least.

MPLS L2 VPN Tweaks – Solution

We increase MTU size of the transmission channel by either of the two following means:

- Increase Peering Interface MTU with “mtu xxxx” command
- Increase Peering Interface MPLS MTU with “mpls mtu xxxx” command

Also, we need to increase Switch system MTU with “system mtu ZZZZ”
- setting the switch to its highest supported MTU settings will be

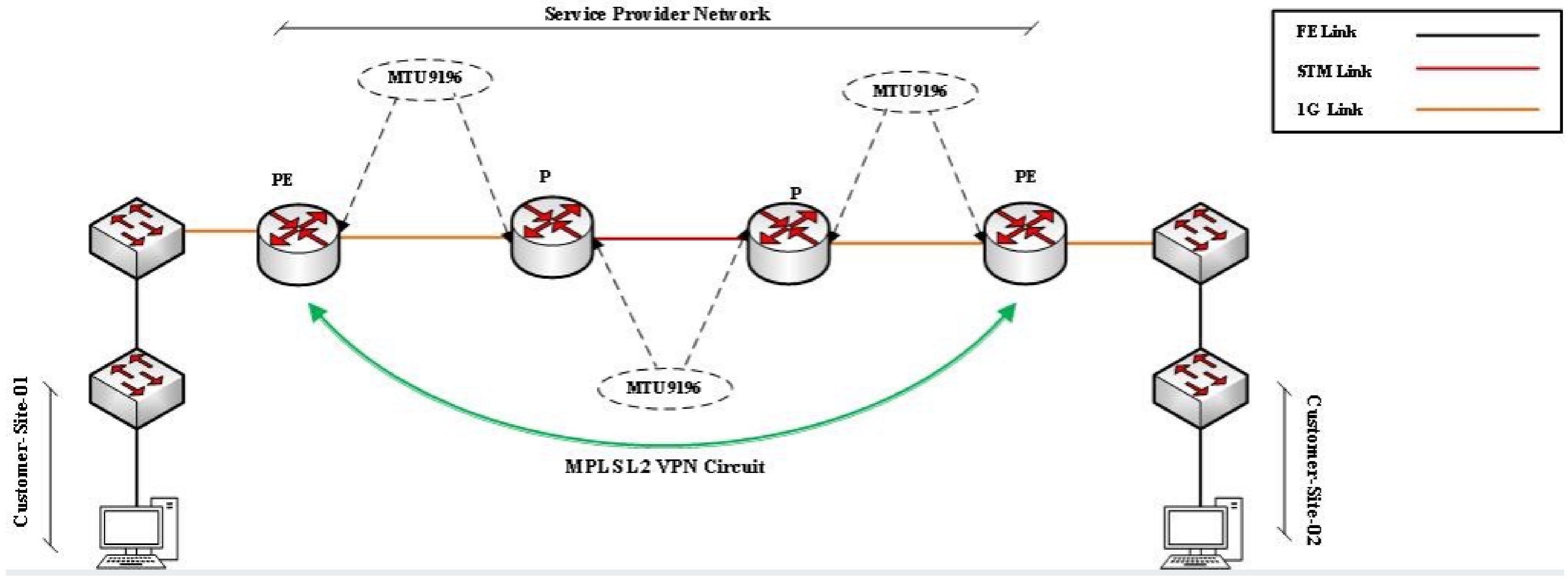
Next concern: what to set for “XXXX” ?

MPLS L2 VPN Tweaks – Solution

- MTU value of 9196 is minimum as per our experience operating with multiple transmission technologies [TDM/SDH/Ethernet].
- In case of POS/SONET only we have tested down to 4470 with successful results.
- But with TDM/SDH transmission channel 9196 is mandatory according to our experience for MPLS L2 VPN service to work properly.

Note: This may not be same for all. Things may differ from one network to another. But this can be considered as a head start.

MPLS L2 VPN Tweaks – Solution



IPv6 Subnetting

IPv6 Deployment

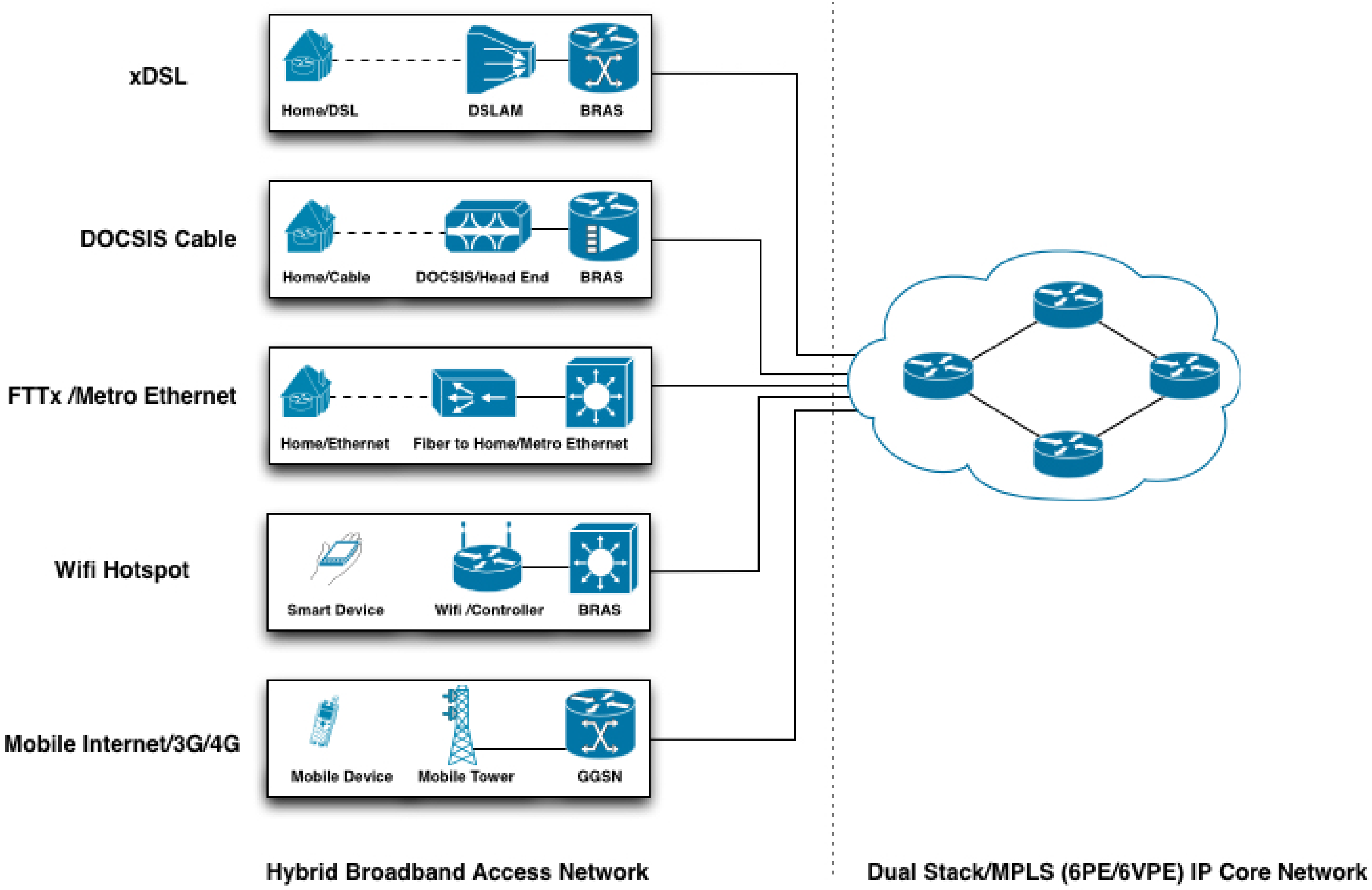
IPv4 BGP Reports

APNIC R&D	**5,61,890**
Route-Views.Oregon-ix.net	**5,87,977**

IPv6 BGP Reports

APNIC R&D	**23,766**
Route-Views.Oregon-ix.net	**24,855**

Access network :

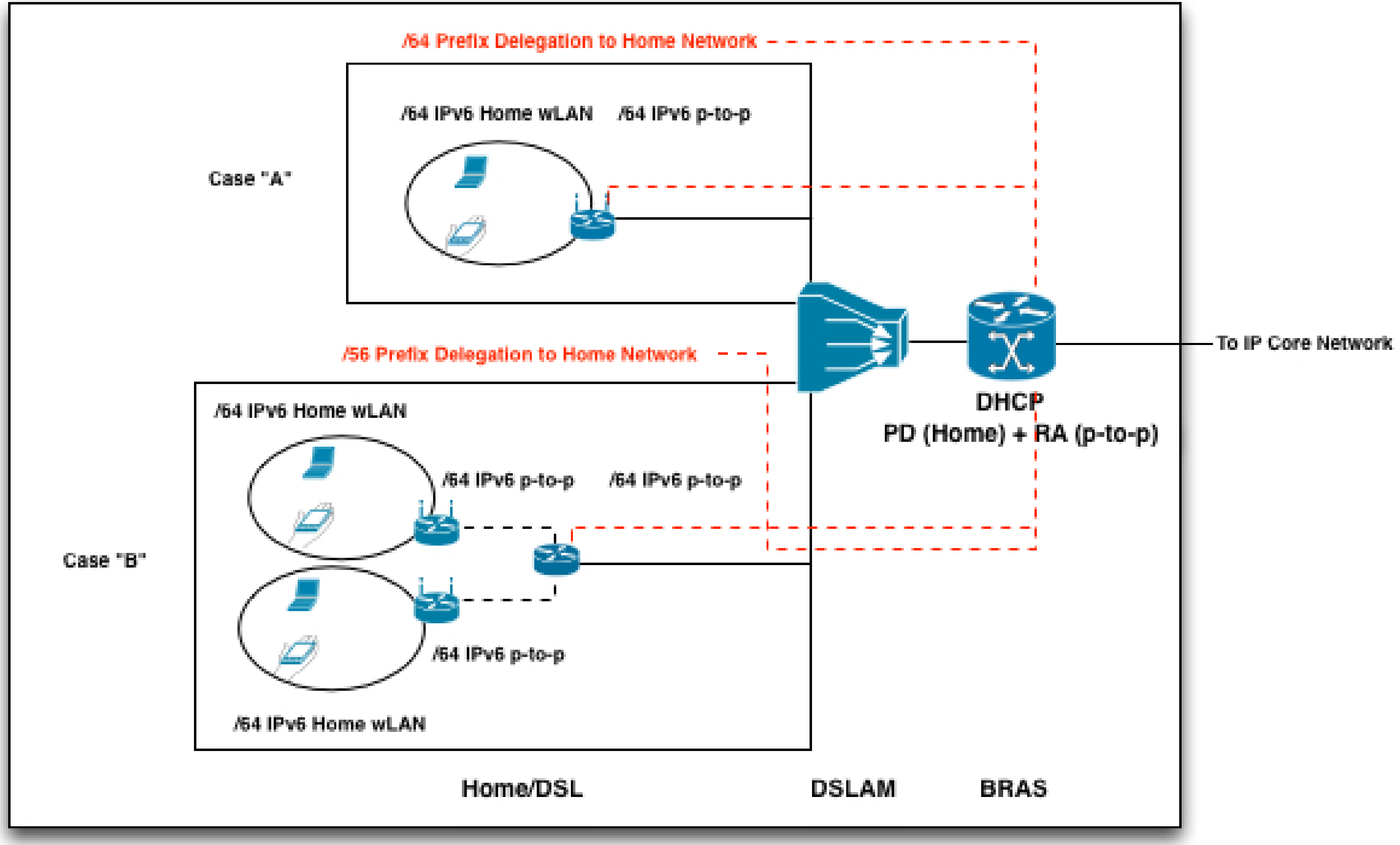


Deployment Cases : Broadband Network

Case A – A single network link where all end user devices will be connected.

Case B – Multiple network links at end user segment.

Deployment Cases :



Best Practice :

Case A :

- /64 where it is known that only one subnet is required.

Case B :

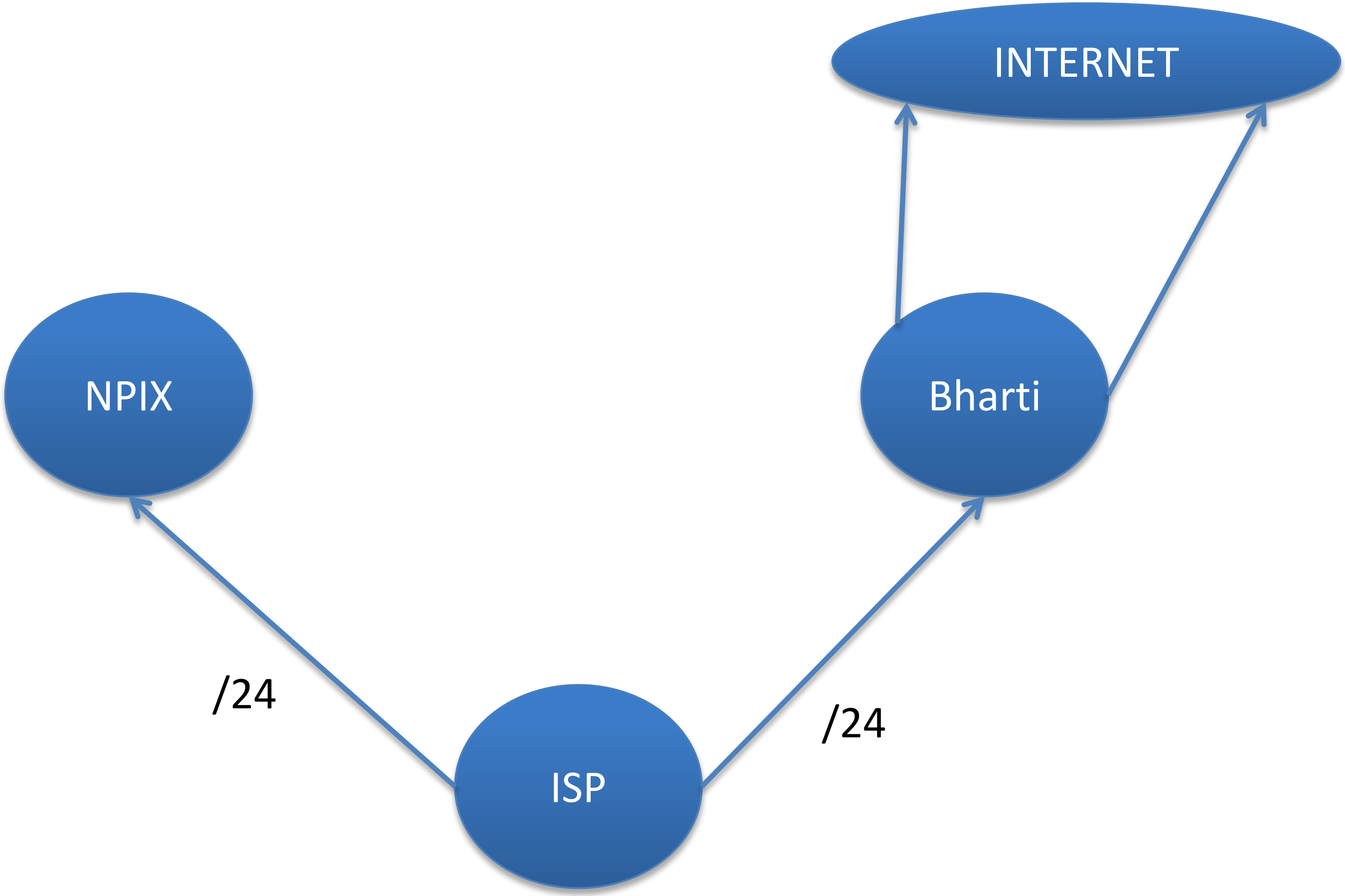
- /56 for small sites where it is expected only a few subnets will be required. Subscribers can receive a /56 when connecting through on-demand or always-on connections such as small office and home office enterprises.

- /48 for larger sites, or if an end site is expected to grow into a large network and multihome.

Prefix Announcement :

Use of Community

Prefix Announcement :



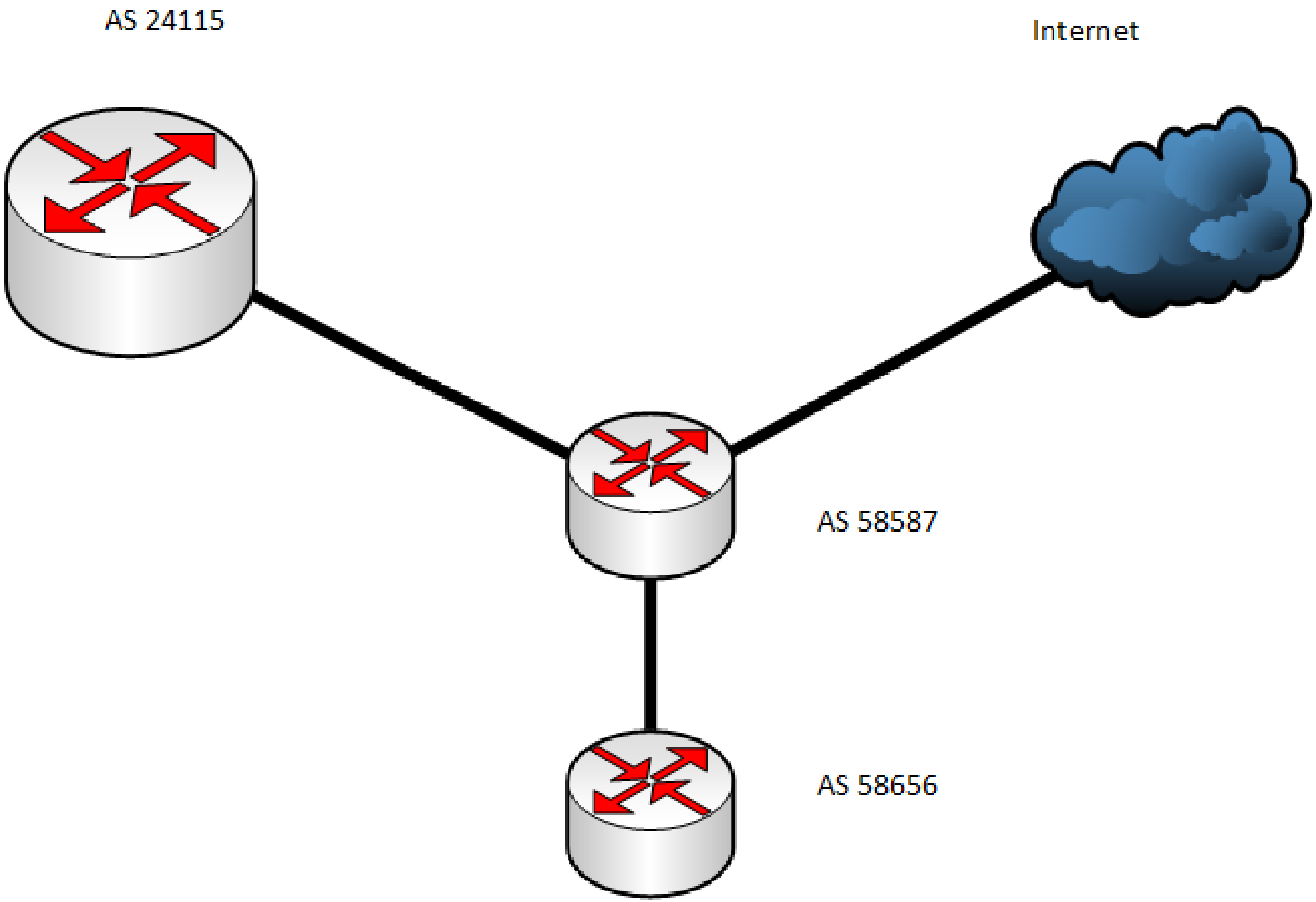
Community :

- Community information is included as a path attribute in BGP update messages.
- A community value is a 32-bit field that is divided into two main sections as a total of 4 octets.
- The first 16 bits of the value encode the AS number of the network that originated the community, while the last 16 bits carry a unique number assigned by the AS.
- Community Notation -> as-number:community-value

Some Predefine Community :

- no-export—Do not advertise to eBGP peers. Keep this route within an AS.
- no-advertise—Do not advertise this route to any peer, internal or external.
- internet—Advertise this route to the Internet community. Any router belongs to this community.
- local-as—Use in confederation scenarios to prevent the transmit of packets outside the local AS.

Use of Community : Towards Commercial IX.



Use of Community :

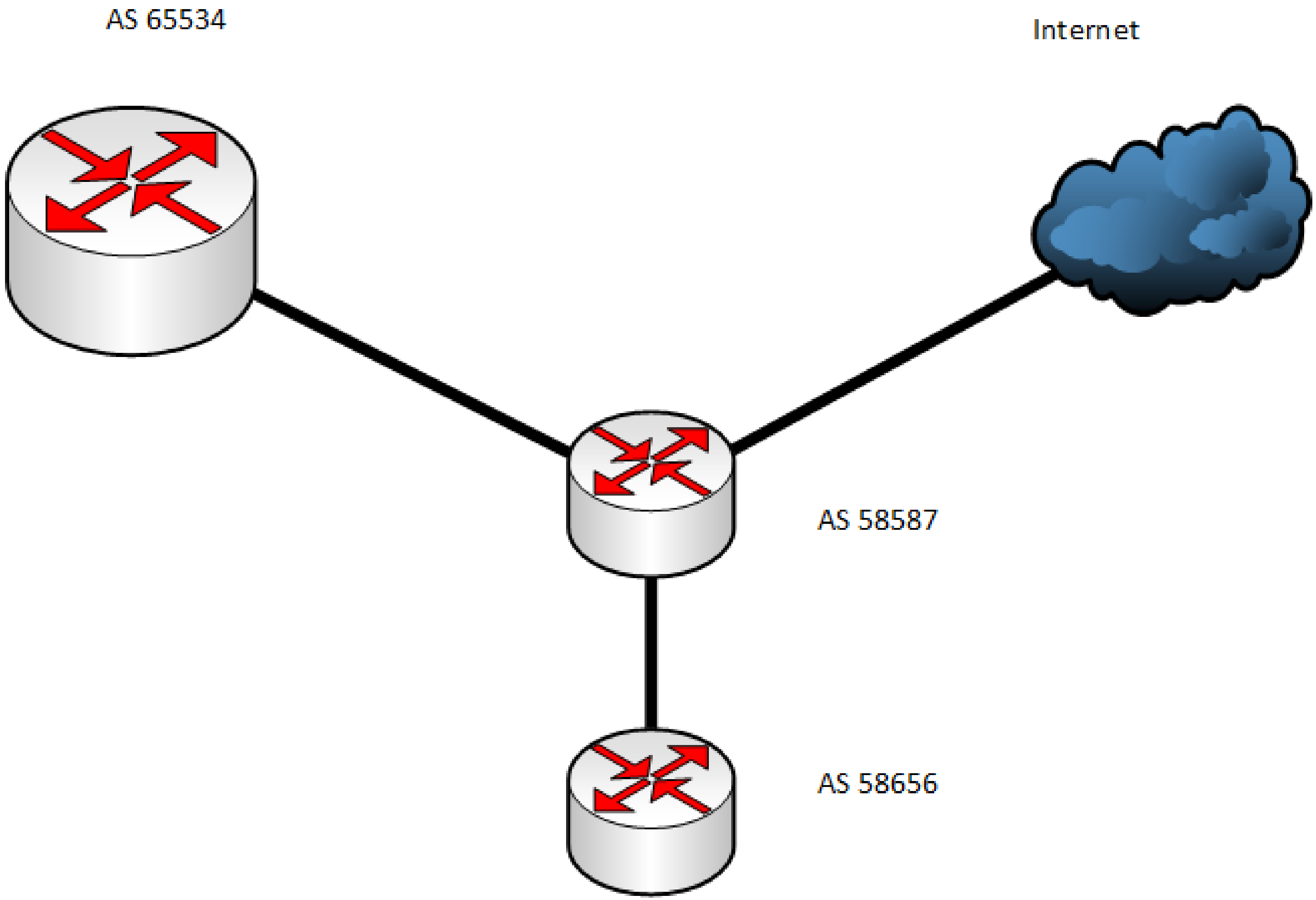
For Receiving from Equnix-IX :

```
set policy-options policy-statement equinix-import term 2 from route-filter 0.0.0.0/0 prefix-length-range /8-/24 community add 58587:24115

set policy-options policy-statement equinix-import term 2 from route-filter 0.0.0.0/0 prefix-length-range /8-/24 accept
```

For Advertisement to Client:

```
set policy-options policy-statement bdhub-equinix-only term advertise from community 58587:24115

set policy-options policy-statement bdhub-equinix-only term advertise then accept
```

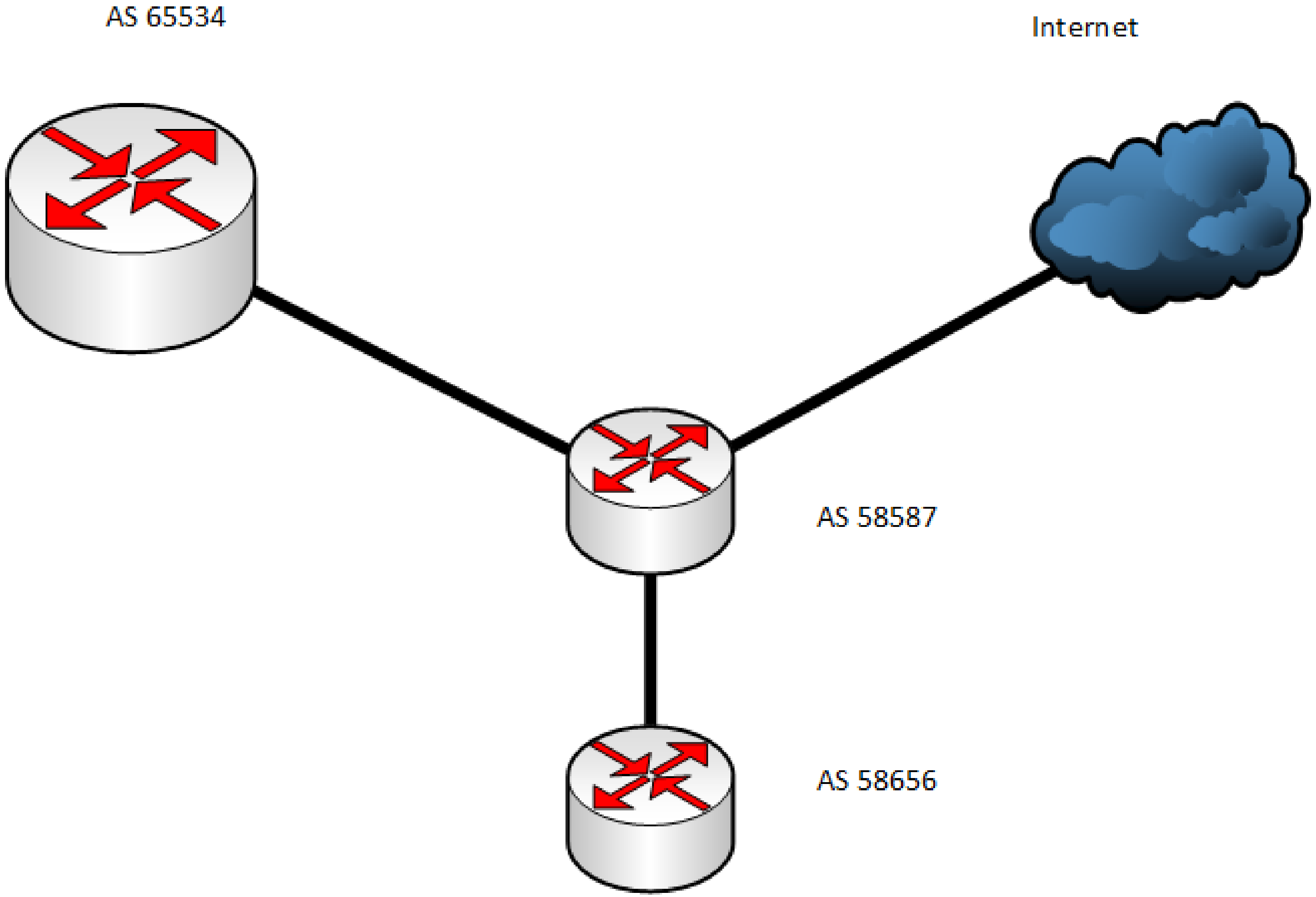
Use of Community : RTBH



Use of Community :

```
ip route 192.0.2.1 255.255.255.255 Null0

ip community-list 30 permit 58587:777

route-map RTBH permit 10
match community 30
set ip next-hop 192.0.2.1

 address-family ipv4
  network 103.7.248.70 mask 255.255.255.255 route-map RTBH
```

Community : Local IX (NPIX)



Community : Towards Client

For Receiving from Upstream:

```
ip as-path access-list 100 deny _58587_
ip as-path access-list 100 permit .*

route-map bdix-in permit 10
 match as-path 100
 set community 58587:65534
```

For Advertisement to Downstream:

```
ip community-list 10 permit 58587:65534

route-map bdix-prefix permit 10
 match community 10
 set community no-export additive
```

Community : Towards BDIX

For Receiving from Downstream:

```
route-map teletalk-v4-map permit 100
 match ip address prefix-list teletalk-v4-in
 set community 58587:700
```

For Advertisement to Upstream:

```
ip community-list 20 permit 58587:700  # only for local IX
ip community-list 30 permit 58587:800  # only for internet
ip community-list 40 permit 58587:900  # for Both

route-map bdix-out permit 100
 match community 20
route-map bdix-out permit 110
 match community 40
```

Route Redistribution

(If you really need to do)

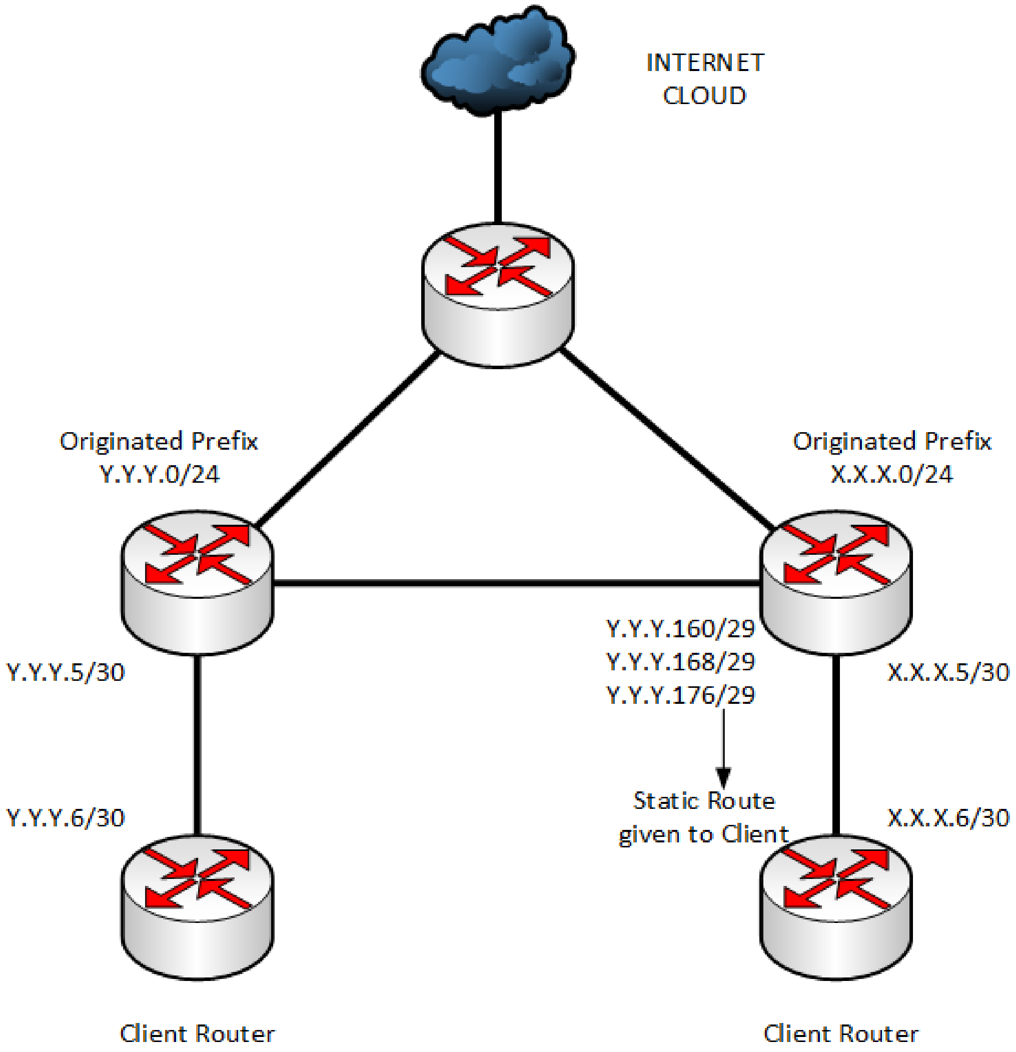
Route Redistribution

People may need to redistribute routes from different protocols to different protocols.

Assumption :

- A static route of a IP Block is given towards client which is originated in different distant router.

Route Redistribution



Route Redistribution

How :

- Identify the subnets to be redistributed.
- Make an ACL for those subnets.
- Make a Route-Map and match that ACL.
- While redistribute, make sure that you are using that route-map.

Route Redistribution

Example :

- Identify the subnets to be redistributed.
  - Y.Y.Y.160/29
  - Y.Y.Y.168/29
  - Y.Y.Y.176/29

Route Redistribution

Example :

- Make an ACL for those subnets.
  - access-list 10 permit Y.Y.Y.160 0.0.0.7
  - access-list 10 permit Y.Y.Y.168 0.0.0.7
  - access-list 10 permit Y.Y.Y.176 0.0.0.7

Route Redistribution

Example :

- Make a Route-Map and match that ACL.
  - route-map static-red-ospf permit 10
  - match ip address 10

Route Redistribution

Example :

- While redistribute, make sure that you are using that route-map.
  - router ospf X
  - redistribute static subnets route-map static-red-ospf

Route Redistribution

Caution :

- Don’t redistribute IGP into BGP
- Don’t redistribute BGP into IGP

Router Security (IPv4 & IPv6)

Router Security (IPv4 & IPv6)

- Control Plane
- Management Plane
- Data Plane

Management Plane Filters

- Authenticate Access
- Define Explicit Access To/From Management Stations
  - SNMP
  - Syslog
  - TFTP
  - NTP
  - SSH, Telnet, etc.

Securing SNMP

```
access-list 99 permit 192.168.1.250
access-list 99 permit 192.168.1.240

snmp-server community N3T-manag3m3nt ro 99
```

Securing SSH

```
ipv6 access-list AUTHORIZED_IPV6_HOST
 permit ipv6 host 2405:7600:0:6::250 any
 deny ipv6 any any log
!
ip access-list extended AUTHORIZED_IPV4_HOST
 permit tcp host 103.21.75.5 any eq 22
 deny   tcp any any log
!
line vty 0 4
 access-class AUTHORIZED_IPV4_HOST in
 ipv6 access-class AUTHORIZED_IPV6_HOST in
```

Secure Access with Passwords and Logout Timers

```
line console 0
    login
    password console-pw
    exec-timeout 1 30
!
line vty 0 4
    login
    password vty-pw
    exec-timeout 5 00
!
enable secret enable-secret
username bob secret bob-secret
```

[redacted] dual Users

Restrict Access To Trusted Hosts

- Use filters to specifically permit hosts to access an infrastructure device
- Example

```
access-list 103 permit tcp host 192.168.200.7 192.168.1.0
   0.0.0.255 eq 22 log-input
access-list 103 permit tcp host 192.168.200.8 192.168.1.0
   0.0.0.255 eq 22 log-input
access-list 103 permit tcp host 192.168.100.6 192.168.1.0
   0.0.0.255 eq 23 log-input
access-list 103 deny ip any any log-input
!
line vty 0 4
access-class 103 in
transport input ssh
```

Banner – What Is Wrong ?

```
banner login ^C

     You should not be on this device.

     Please Get Off My Router!!
^C
```

More Appropriate Banner

```
!!!!  WARNING !!!!
You have accessed a restricted device.
All access is being logged and any
unauthorized access will be prosecuted to
the full extent of the law.
```

Centralized Log (syslog)

```
Router(config)# logging 192.168.0.30
Router(config)# logging trap 3
Router(config)# logging facility local3
```

Trap:
Emergency: 0
Alert: 1
Critical: 2
Error: 3
Warning: 4
Notice: 5
Informational: 6
Debug: 7

Facility Type Keyword	Description
auth	Authorization system
cron	Cron facility
daemon	System daemon
kern	Kernel
local0-7	Locally defined messages
lpr	Line printer system
mail	Mail system
news	USENET news
sys9	System use
sys10	System use
sys11	System use
sys12	System use
sys13	System use
sys14	System use
syslog	System log
user	User process
uucp	UNIX-to-UNIX copy system

Configuration change logging

```
Router# configure terminal
Router(config)# archive
Router(config-archive)# log config
Router(config-archive-log-config)# logging enable
Router(config-archive-log-config)# logging size 200
Router(config-archive-log-config)# hidekeys
Router(config-archive-log-config)# notify syslog
```

```
768962: Feb  1 20:59:45.081 UTC: %PARSER-5-CFGLOG_LOGGEDCMD: User:fakrul  logged
command:!exec: enable
768963: Feb  1 21:03:17.160 UTC: %PARSER-5-CFGLOG_LOGGEDCMD: User:fakrul  logged command:no
ipv6 prefix-list dhakacom_AS23956_IN_IPv6 description
768965: Feb  1 21:03:19.182 UTC: %SYS-5-CONFIG_I: Configured from console by fakrul on vty0
(2405:7600:0:6::250)
```

Turn Off Unused Services

Feature	Description	Default	Recommendation	Command
CDP	Proprietary layer 2 protocol between Cisco devices	Enabled		`no cdp run`
TCP small servers	Standard TCP network services: echo, chargen, etc	11.3: disabled 11.2: enabled	This is a legacy feature, disable it explicitly	`no service tcp-small-servers`
UDP small servers	Standard UDP network services: echo, discard, etc	11.3: disabled 11.2: enabled	This is a legacy feature, disable it explicitly	`no service udp-small-servers`
Finger	Unix user lookup service, allows remote listing of logged in users.	Enabled	Unauthorized persons don't need to know this, disable it.	`no service finger`
HTTP server	Some Cisco IOS devices offer web-based configuration	Varies by device	If not in use, explicitly disable, otherwise restrict access	`no ip http server`
Bootp server	Service to allow other routers to boot from this one	Enabled	This is rarely needed and may open a security hole, disable it	`no ip bootp server`

Turn Off Unused Services

Feature	Description	Default	Recommendation	Command
PAD Service	Router will support X.25 packet assembler service	Enabled	Disable if not explicitly needed	`no service pad`
IP source routing	Feature that allows a packet to specify its own route	Enabled	Can be helpful in attacks, disable it	`no ip source-route`
Proxy ARP	Router will act as a proxy for layer 2 address resolution	Enabled	Disable this service unless the router is serving as a LAN bridge	`no ip proxy-arp`
IP directed broadcast	Packets can identify a target LAN for broadcasts	Enabled (11.3 & earlier)	Directed broadcast can be used for attacks, disable it	`no ip directed-broadcast`

Configuration (Templates)

```
!configure timezone
service timestamps debug uptime
service timestamps log datetime localtime
service password-encryption
clock timezone UTC +6

! turn off unnecessary services (global)
no ip domain-lookup
no cdp run
no ip http server
no ip source-route
no service finger
no ip bootp server
no service udp-small-servers
no service tcp-small-servers
```

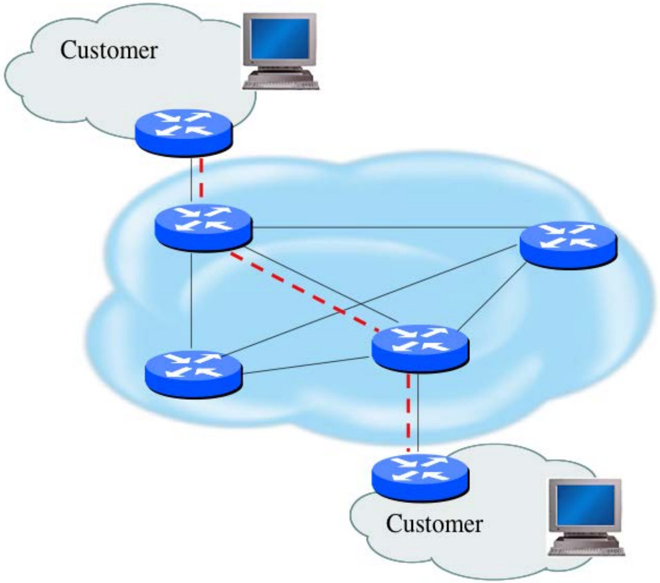
```
! turn off unnecessary services (interface)
Interface GigabitEthernet0/0
no ip redirects
no ip directed-broadcast
no ip proxy arp
no cdp enable

! turn on logging and snmp
logging 192.168.253.56
snmp-server communityTxo~QbW3XM ro 98
!
access-list 99 permit 192.168.253.0 0.0.0.255
access-list 99 deny any log
access-list 98 permit host 192.168.253.51
access-list 98 deny any log
!
```

Configuration (Templates)

```
line vty 0 4
access-class 99 in
exec-timeout 2 0
transport input ssh
!
line con 0
access-class 99 in
exec-timeout 2 0
!
banner motd #
!!!!  WARNING !!!!
You have accessed a restricted device.
All access is being logged and any
unauthorized access will be prosecuted to
the full extent of the law.
#
```

```
!Turn on NTP
ntp authenticate
ntp authentication-key 1 md5 -UN&/6[oh6
ntp trusted-key 1
ntp access-group peer 96
ntp server 192.168.254.57 key 1
access-list 96 permit host 192.168.254.57
access-list 96 deny any log
```

Securing The Data Path



- Filtering and rate limiting are primary mitigation techniques
- Edge filter guidelines for ingress filtering (BCP38/BCP84)
- Null-route and black-hole any detected malicious traffic
- Netflow is primary method used for tracking traffic flows
- Logging of Exceptions

Data Plane (Packet) Filters

- Most common problems
  - Poorly-constructed filters
  - Ordering matters in some devices
- Scaling and maintainability issues with filters are commonplace
- Make your filters as modular and simple as possible
- Take into consideration alternate routes
  - Backdoor paths due to network failures

Filtering Deployment Considerations

- How does the filter load into the router?
- Does it interrupt packet flow?
- How many filters can be supported in hardware?
- How many filters can be supported in software?
- How does filter depth impact performance?
- How do multiple concurrent features affect performance?
- Do I need a standalone firewall?

General Filtering Best Practices

- Explicitly deny all traffic and only allow what you need
- The default policy should be that if the firewall doesn't know what to do with the packet, deny/drop it
- Don't rely only on your firewall for all protection of your network
- Implement multiple layers of network protection
- Make sure all of the network traffic passes through the firewall
- Log all firewall exceptions (if possible)

Filtering Recommendations

- Log filter port messages properly
- Allow only internal addresses to enter the router from the internal interface
- Block packets from outside (untrusted) that are obviously fake or commonly used for attacks
- Block packets that claim to have a source address of any internal (trusted) network.

Filtering Recommendations

- Block incoming loopback packets and RFC 1918 networks
  - 127.0.0.0
  - 10.0.0.0 – 10.255.255.255
  - 172.16.0.0 – 172.31.0.0
  - 192.168.0.0 – 192.168.255.255
- Block multicast packets (if NOT using multicast)
- Block broadcast packets (careful of DHCP & BOOTP users)
- Block incoming packets that claim to have same destination and source address

DoS Filtering

(* these networks were reallocated and are actually used)

Description	**Network**
default	0.0.0.0 /8
loopback	127.0.0.0 /8
RFC 1918	10.0.0.0 /8
RFC 1918	172.16.0.0 /12
RFC 1918	192.168.0.0 /16
Net Test	192.0.2.0 /24
Testing devices *	192.18.0.0 /15
IPv6 to IPv4 relay *	192.88.99.0 /24
RFC 1918 nameservers *	192.175.48.0 /24
End-node auto configuration *	169.254.0.0 /16

Example Incoming IPv4 Bogon Packet Filter

```
ip access-list extended DSL-Incoming
 deny   ip 127.0.0.0 0.255.255.255 any log
 deny   ip 10.0.0.0 0.255.255.255 any log
 deny   ip 169.254.0.0 0.0.255.255 any log
 deny   ip 172.16.0.0 0.15.255.255 any log
 deny   ip 192.168.0.0 0.0.255.255 any log
 deny   ip 224.0.0.0 15.255.255.255 any log
 permit icmp any any ttl-exceeded
 permit icmp any any echo-reply
 permit icmp any any echo
 permit tcp any any eq 22 log
 permit udp host <ip address> eq domain <subnet range>
 permit udp host <ip address> eq domain <subnet range>
 permit udp host <ip address>  <subnet range>  eq ntp
 permit udp host <ip address>  <subnet range>  eq ntp
 permit tcp any <my sybnet>  established
 deny   ip any any log
```

Example Incoming IPv4 Bogon Packet Filter

- Bogon and fullbogon peering use different ASNs
- Advertise all fullbogons (IPv4 and IPv6) over a single BGP peering session
- For details: http://www.team-cymru.org/Services/Bogons/bgp.html

Example Outgoing Packet Filter

```
access-list 121 permit ip 192.168.1.250
0.0.0.255 any
access-list 121 deny ip any any log
!
interface serial 1/1/1.3
    Description Link to XYZ
    ip access-group 121 in
```

Infrastructure Filters

- Permit only required protocols and deny ALL others to infrastructure space
  - Filters now need to be IPv4 and IPv6!
  - Applied inbound on ingress interfaces
- Basic premise: filter traffic destined TO your core routers
- Develop list of required protocols that are sourced from outside your AS and access core routers
  - Example: eBGP peering, GRE, IPSec, etc.
  - Use classification filters as required
- Identify core address block(s)
  - This is the protected address space
  - Summarization is critical for simpler and shorter filters

References

- Articles, documents and templates from Team CYMRU http://www.team-cymru.org/ReadingRoom/
- Google for the information specifics from the vendors you use: “<vendor> security template”

Route Optimization

Route Optimization

<u>Routes :</u>

- Default Route Only
- Default + Full Routes
- Full Routes Only
- Partial Routes

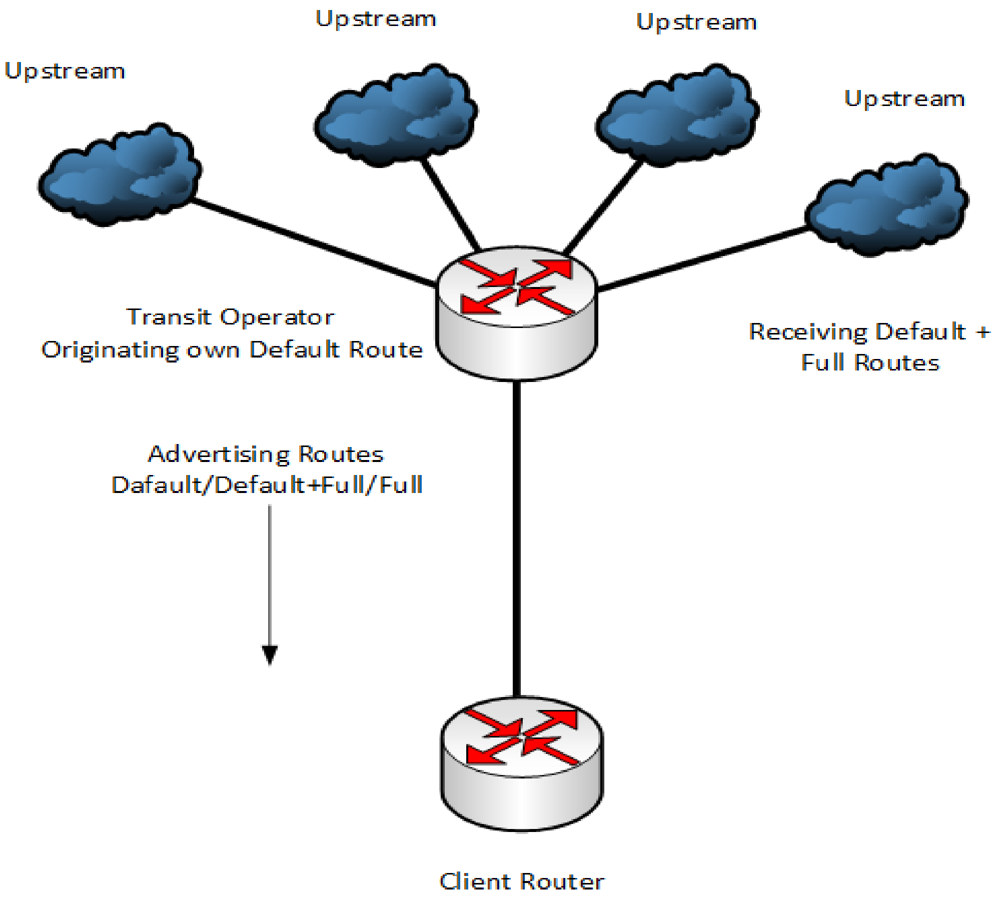
Route Optimization

Default Route Only – Why :

- Routers that are not capable to handle Full Internet Routing Table, receive default route only.
- For advertisement, always prefer to advertise locally originated default route in BGP.

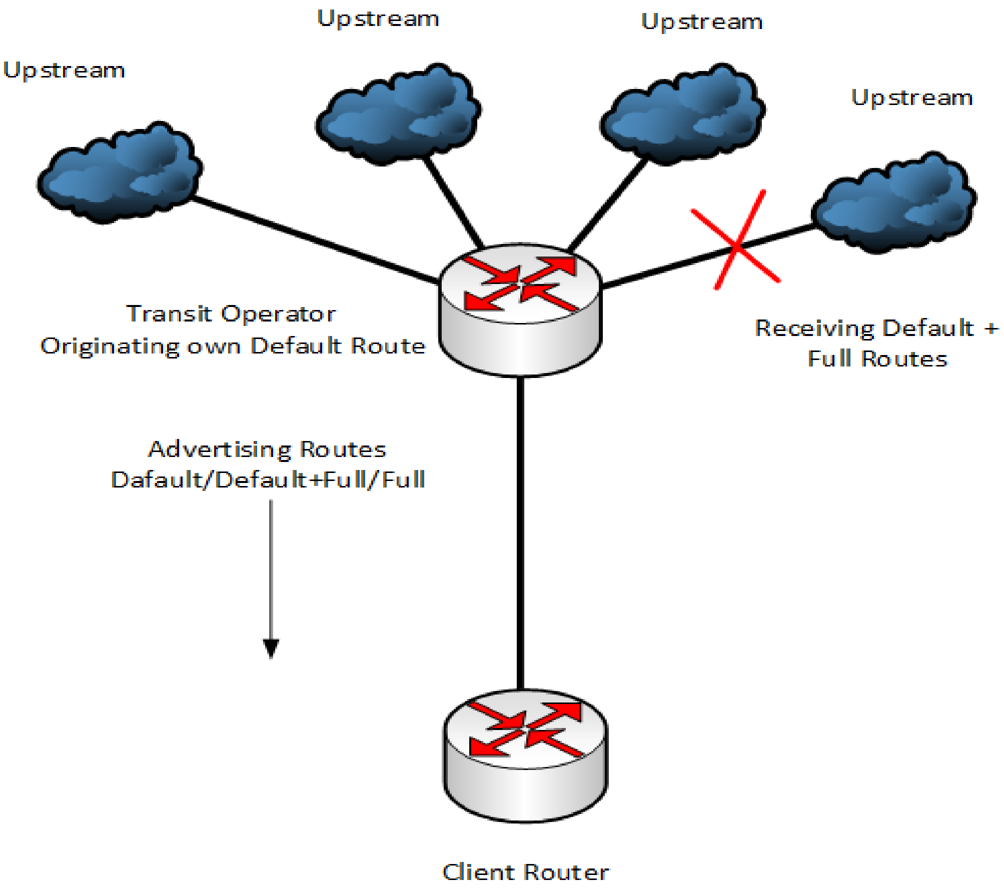
Route Optimization

Default Route Only – Why :



Route Optimization

Default Route Only – Why :



Route Optimization

Default + Full Routes – Why :

- Its better to have both Default and Full Routes from Upstream if your router supports that.
- Full Routes give you the access to all destinations with specific address.
- If your upstream can’t give you any specific route for any destination, Default Route might come handy for that destination.

Route Optimization

Full Routes Only – Why :

- General trend for Tier-1 upstream.
- If you have the whole internet routing table, actually you don’t need Default Route from your upstream.

Route Optimization

Partial Routes – Why :

- You don’t need to make your routing table heavy by taking unnecessary Full Routing Table from multiple upstreams.
- If you have multiple upstream from same region, like east or west, you can take Full Route from one upstream since both of them are likely to have same kind of reachability. For redundancy purpose, you can have Default Route from other.

Route Optimization

Partial Routes – Why :

- If you have multiple upstreams from different regions, like east and west, you might want to take partial routing tables from both of them to make your routing table lite but still efficient.
- You can take 2 or 3 as path distant from both the upstream to have good reachability along with Default Route.

Route Optimization

Partial Routes – Why :

- Default Route is necessary to reach those destinations which are far away from 2 or 3 as path distant.
- You need to use Regular Expression for AS Paths to receive Partial Routes from Upstream.

Route Optimization

Partial Routes – Why :

- Some Regular Expressions
  - ip as-path access-list 65 permit _XXX$
  - ip as-path access-list 65 permit ^[0-9]+$
  - ip as-path access-list 65 permit ^[0-9]+_[0-9]+$
  - ip as-path access-list 65 permit ^[0-9]+_[0-9]+_[0-9]+$

This ACL allows 3 AS Path Distance.

(Regex breakdown: ^ means match, [0-9] indicates any numeral, + means any number of the previous expression, _ is a space, and $ is end-of-line)

Regular Expression

- Like Unix regular expressions
  - . Match one character
  - * Match any number of preceding expression
  - + Match at least one of preceding expression
  - ^ Beginning of line
  - $ End of line
  - \ Escape a regular expression character
  - _ Beginning, end, white-space, brace
  - | Or
  - () brackets to contain expression
  - [] brackets to contain number ranges

Regular Expressions

Reg Expression	Comments
.*	match anything
^$	match routes local to this AS
_1800$	originated by AS1800
^1800_	received from AS1800
_1800_	via AS1800
_790_1800_	via AS1800 and AS790
^1800(_1800)*$	multiple AS1800 in sequence (used to match AS-PATH prepends)
^23956(_23956)*(_55531|_58581)*$	AS 55531 or 58581 via AS 23956 and can do AS-PATH prepends

Regular Expressions

Reg Expression	Comments
^[0-9]+$	Match AS_PATH length of one
^[0-9]+_[0-9]+$	Match AS_PATH length of two
^[0-9]+_[0-9]+_[0-9]+$	Match AS_PATH length of three
^[0-9]+_[0-9]+_[0-9]+_[0-9]+$	Match AS_PATH length of four

Acknowledgement

- Philip Smith, NSRC
- Nurul Islam, APNIC
- Fakrul Alam, APNIC
- Sumon Ahmed Sabir, Fiber@Home

Thank You.

SANOG

South Asian Network Operators Group

Fiber@Home