



ISP Network Design

ISP/IXP Workshops

ISP Network Design

- **PoP Topologies and Design**
- **Backbone Design**
- **ISP Systems Design**
- **Addressing**
- **Routing Protocols**
- **Security**
- **Out of Band Management**
- **Operational Considerations**



Point of Presence Topologies

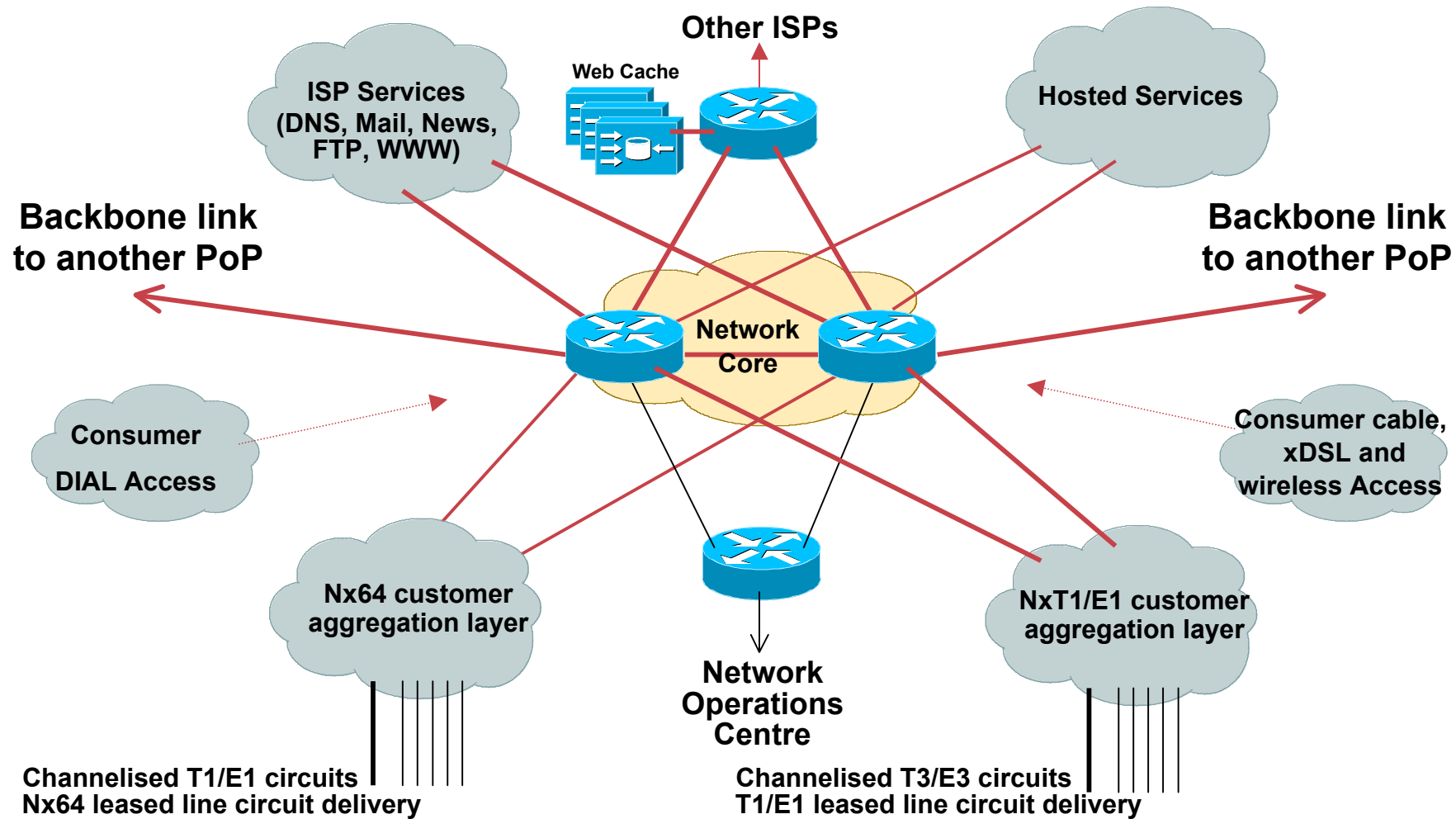
PoP Topologies

- **Core** routers – high speed trunk connections
- **Distribution** routers and **Access** routers – high port density
- **Border** routers – connections to other providers
- **Service** routers – hosting and servers
- Some functions might be handled by a single router

PoP Design

- **Modular Design**
- **Aggregation Services separated according to**
 - connection speed**
 - customer service**
 - contention ratio**
 - security considerations**

Modular PoP Design



Modular Routing Protocol Design

- **Modular IGP implementation**

- IGP “area” per module**

- aggregation/summarisation where possible into the core**

- **Modular iBGP implementation**

- BGP route reflector cluster per module**

- core routers are route-reflectors**

- clients peer with core only**



Point of Presence Design

PoP Modules

- **Low Speed customer connections**
 - PSTN/ISDN dialup**
 - low bandwidth needs**
 - low revenue, large numbers**
- **Medium Speed customer connections**
 - 56/64K to sub-T1/E1 speeds**
 - low bandwidth needs**
 - medium revenue, medium numbers**

PoP Modules

- **High Speed customer connections**
 - E1++ speeds**
 - medium bandwidth needs**
 - high revenue, low numbers**
- **Broad Band customer connections**
 - xDSL, Cable and Wireless**
 - high bandwidth needs**
 - low revenue, large numbers**

PoP Modules

- **PoP Core**

Two dedicated routers

High Speed interconnect

Backbone Links **ONLY**

Do not touch them!

- **Border Network**

dedicated border router to other ISPs

the ISP's "front" door

transparent web caching

PoP Modules

- **ISP Services**

- DNS (cache, secondary)**

- News, Mail (POP3, Relay)**

- WWW (server, proxy, cache)**

- **Hosted Services**

- Virtual Web, WWW (server, proxy, cache)**

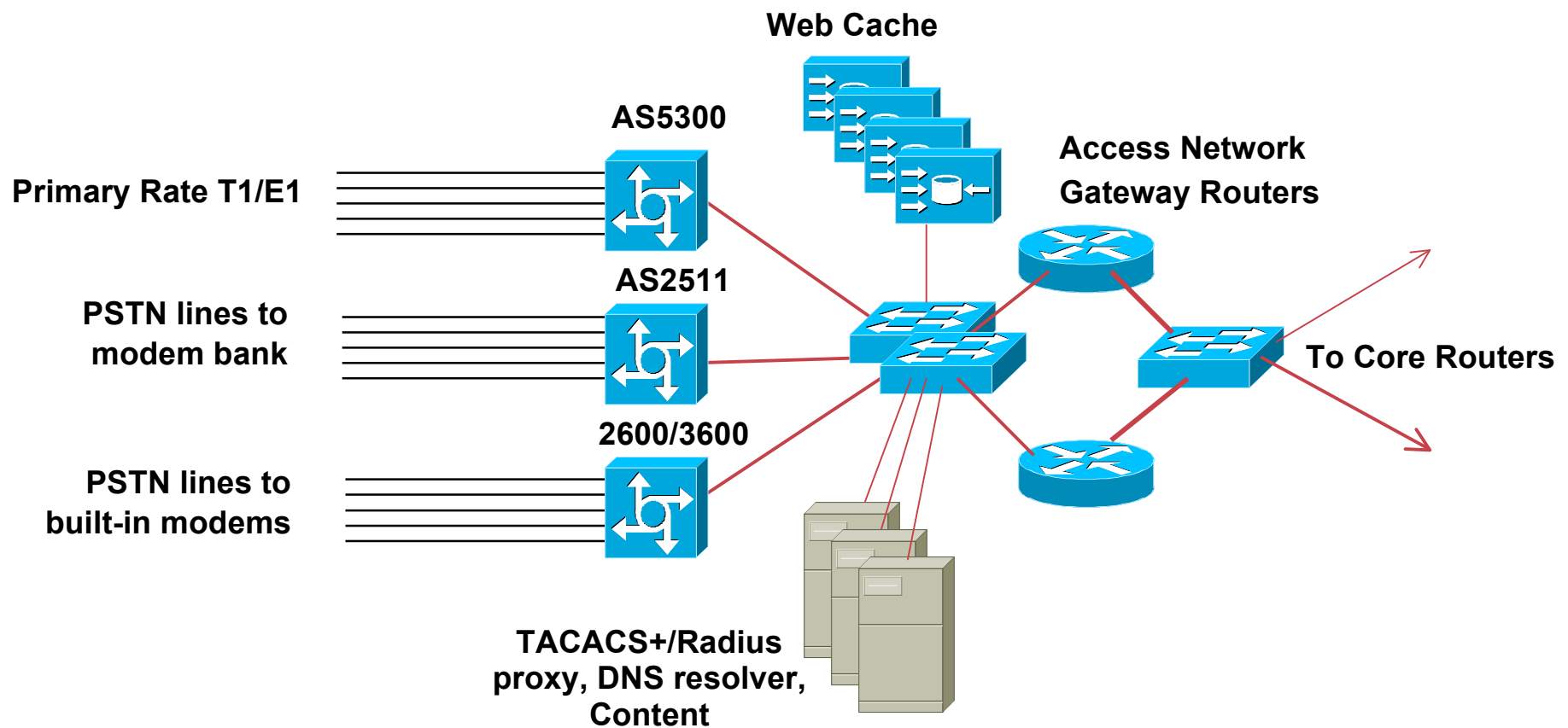
- Information/Content Services**

- Electronic Commerce**

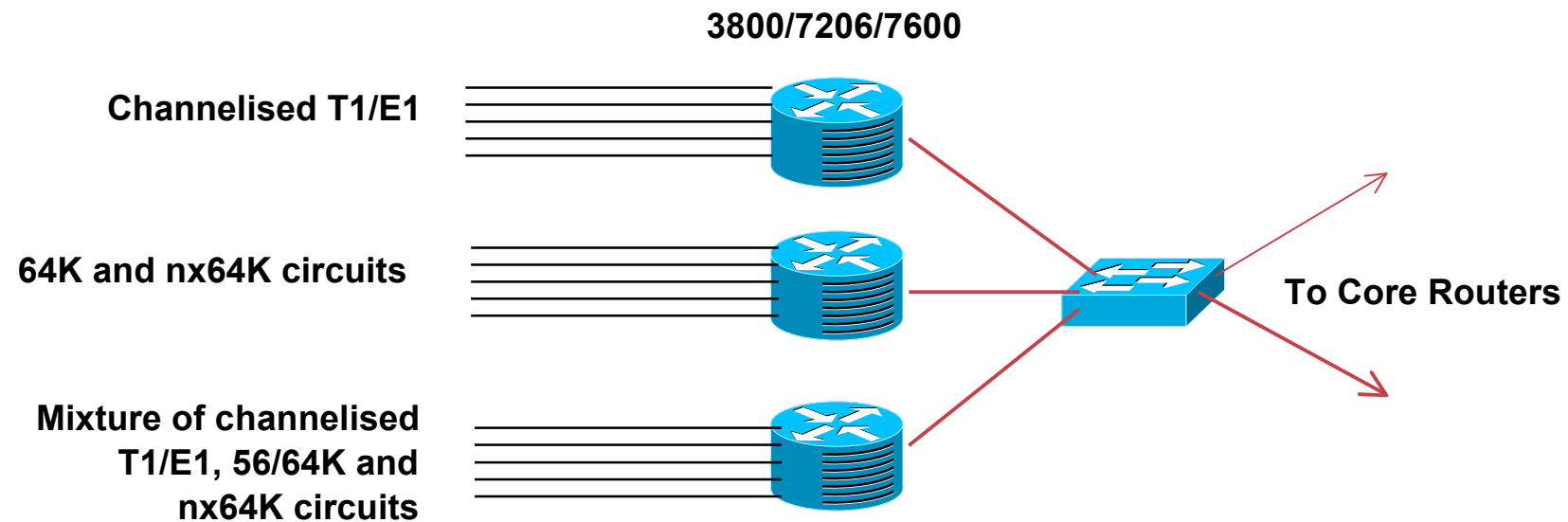
PoP Modules

- **Network Operations Centre**
primary and backup locations
network monitoring
statistics and log gathering
direct but secure access
- **Out of Band Management Network**
The ISP Network “Safety Belt”

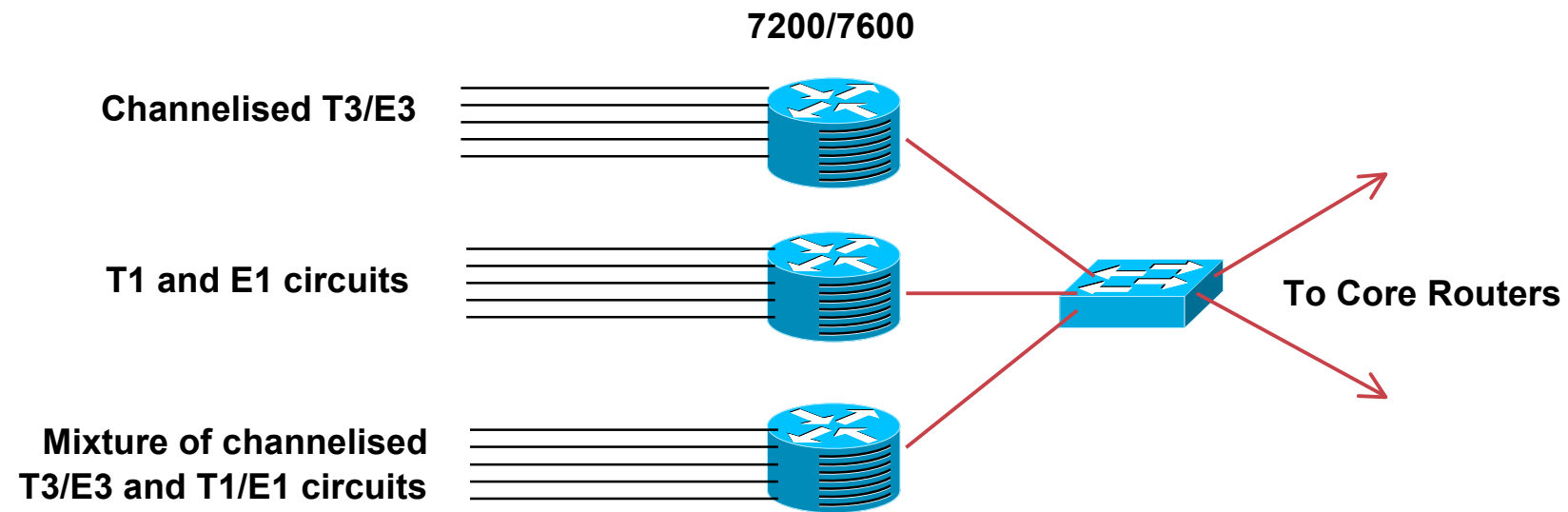
Low Speed Access Module



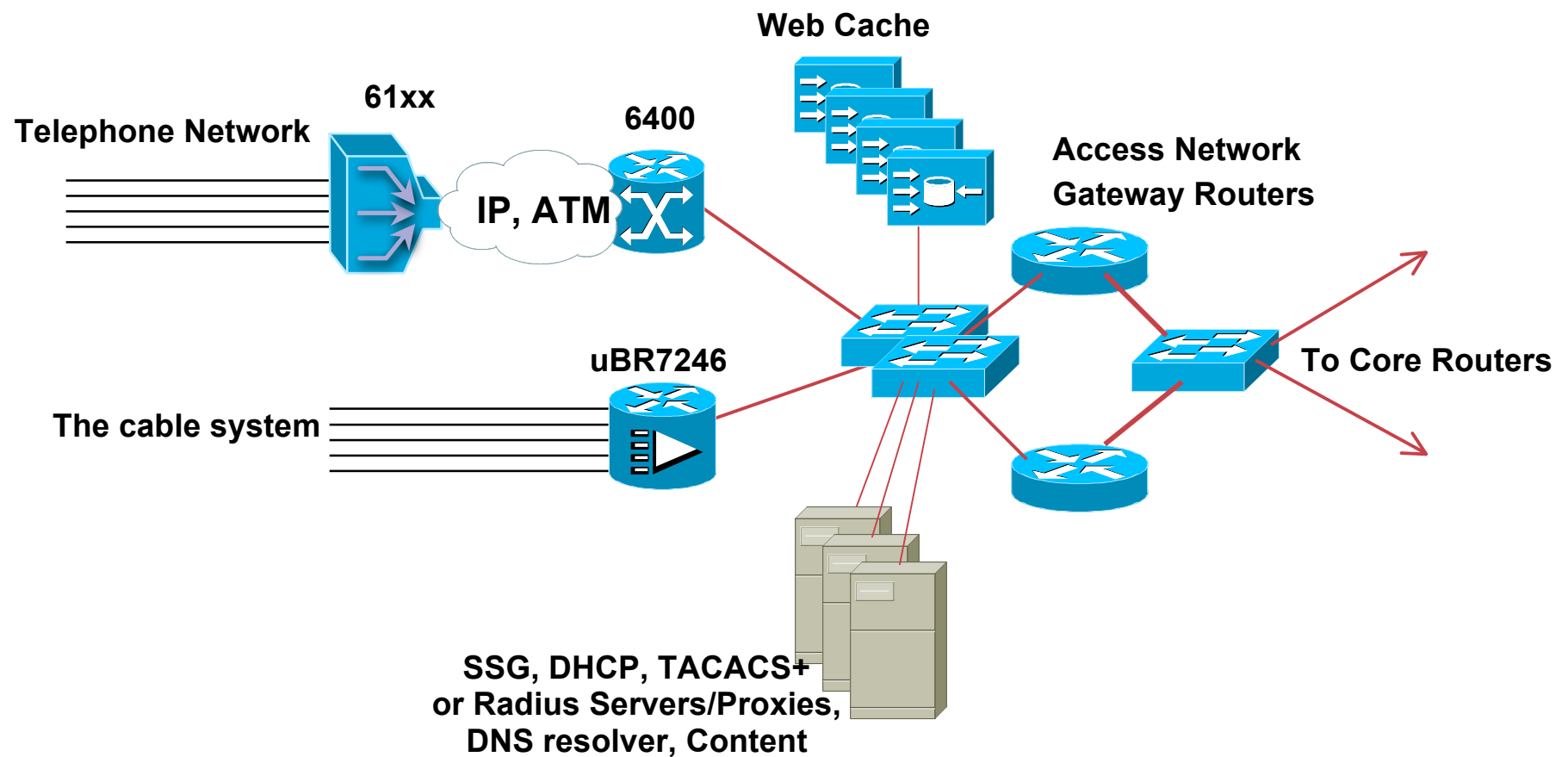
Medium Speed Access Module



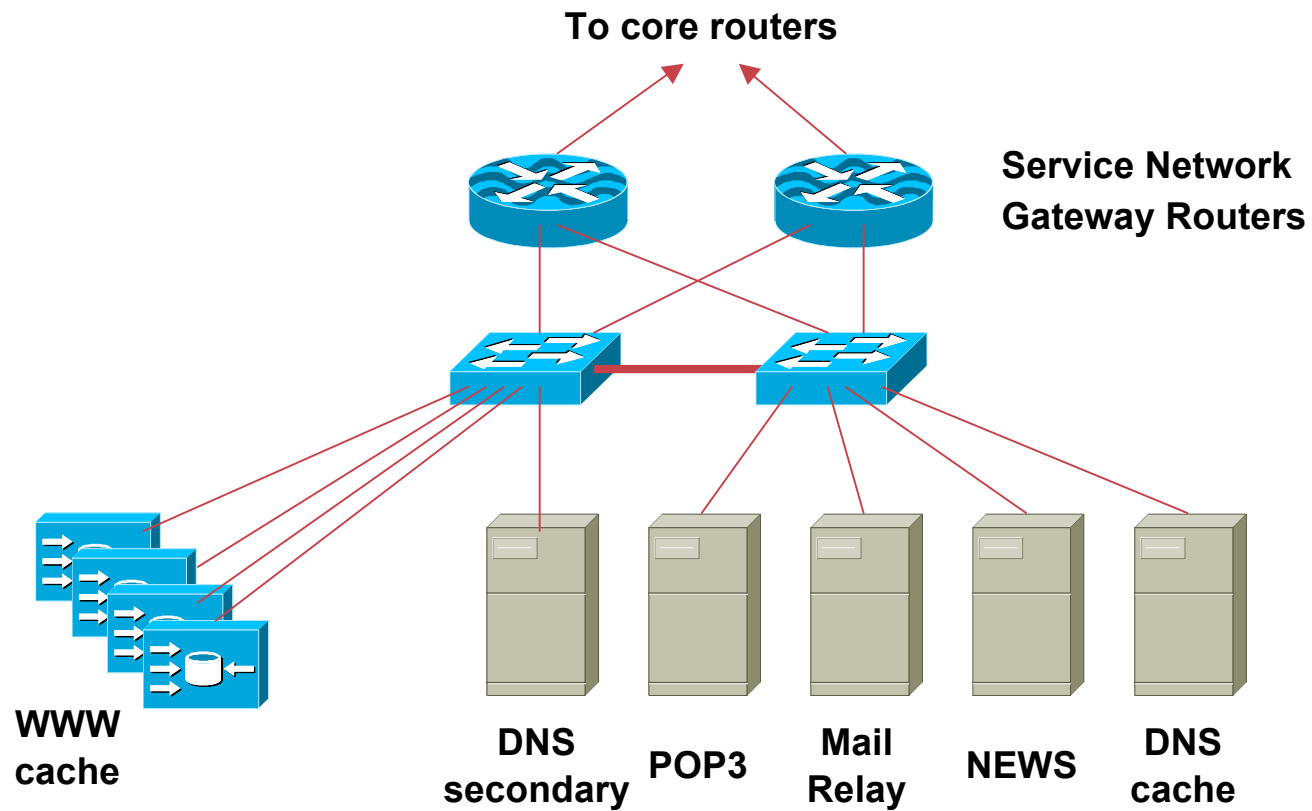
High Speed Access Module



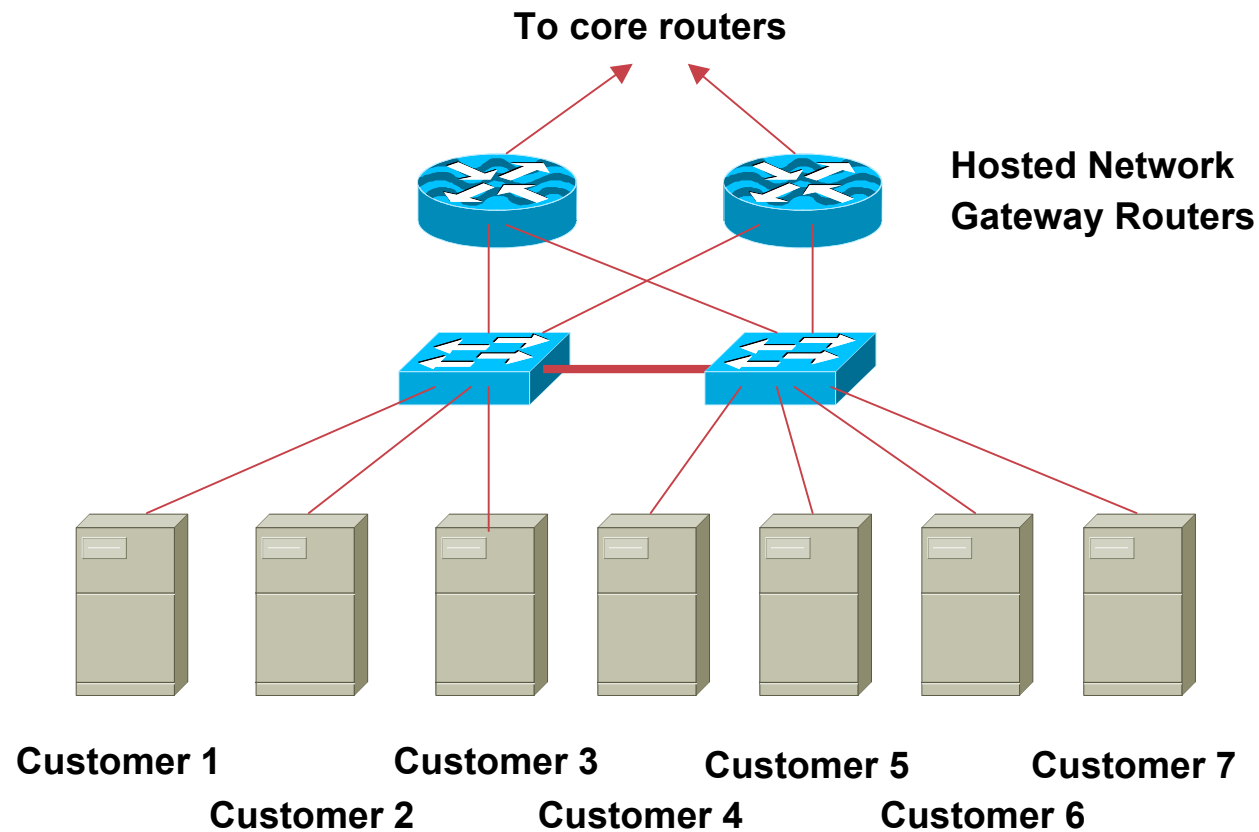
Broad Band Access Module



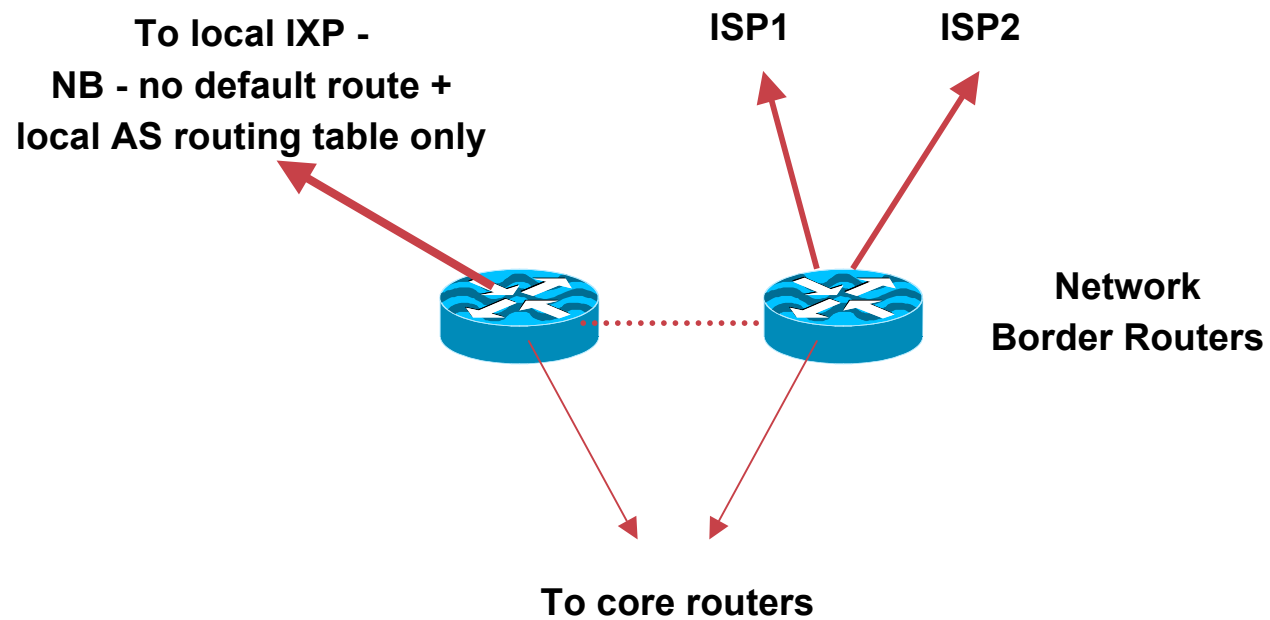
ISP Services Module



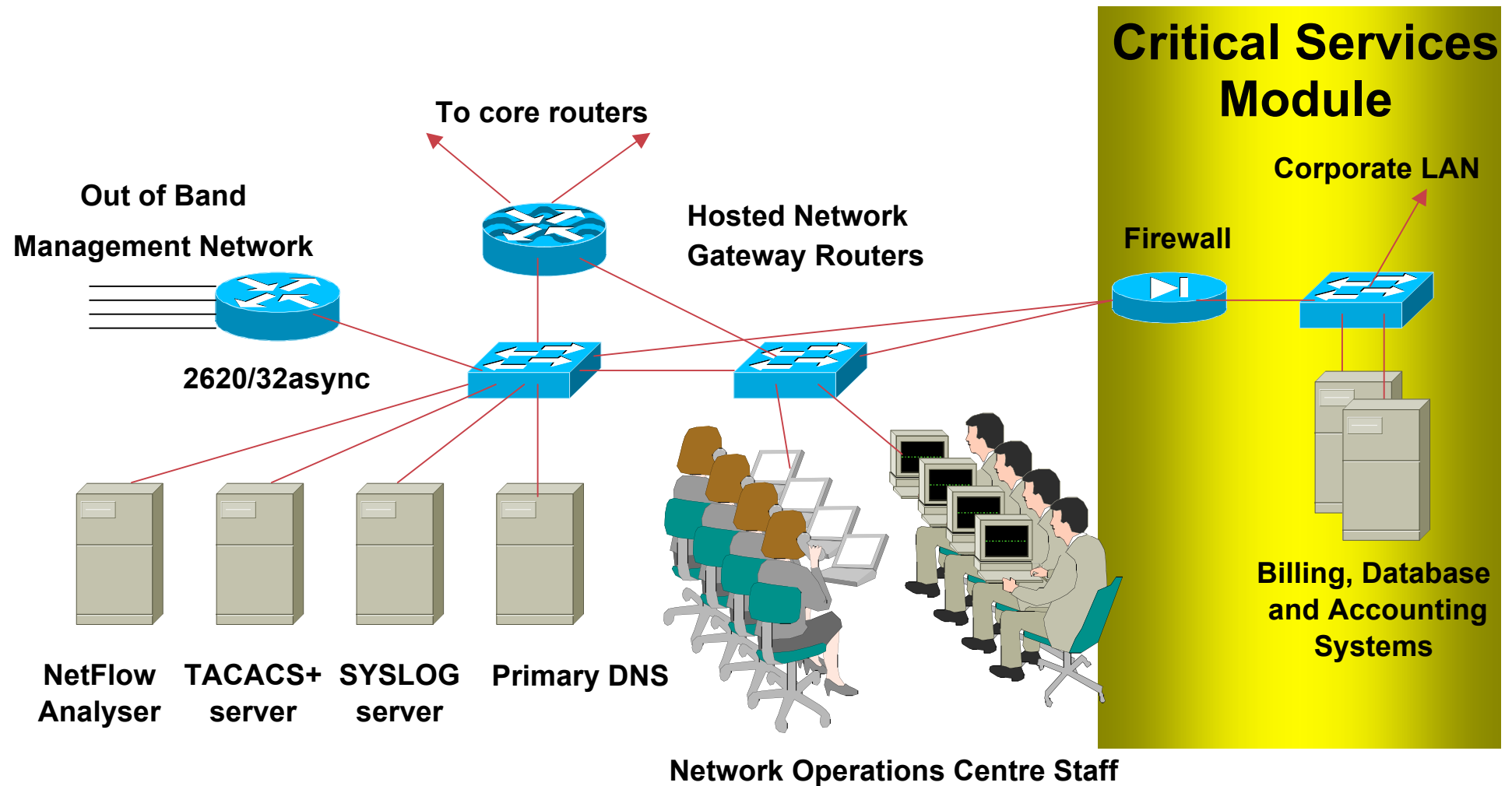
Hosted Services Module



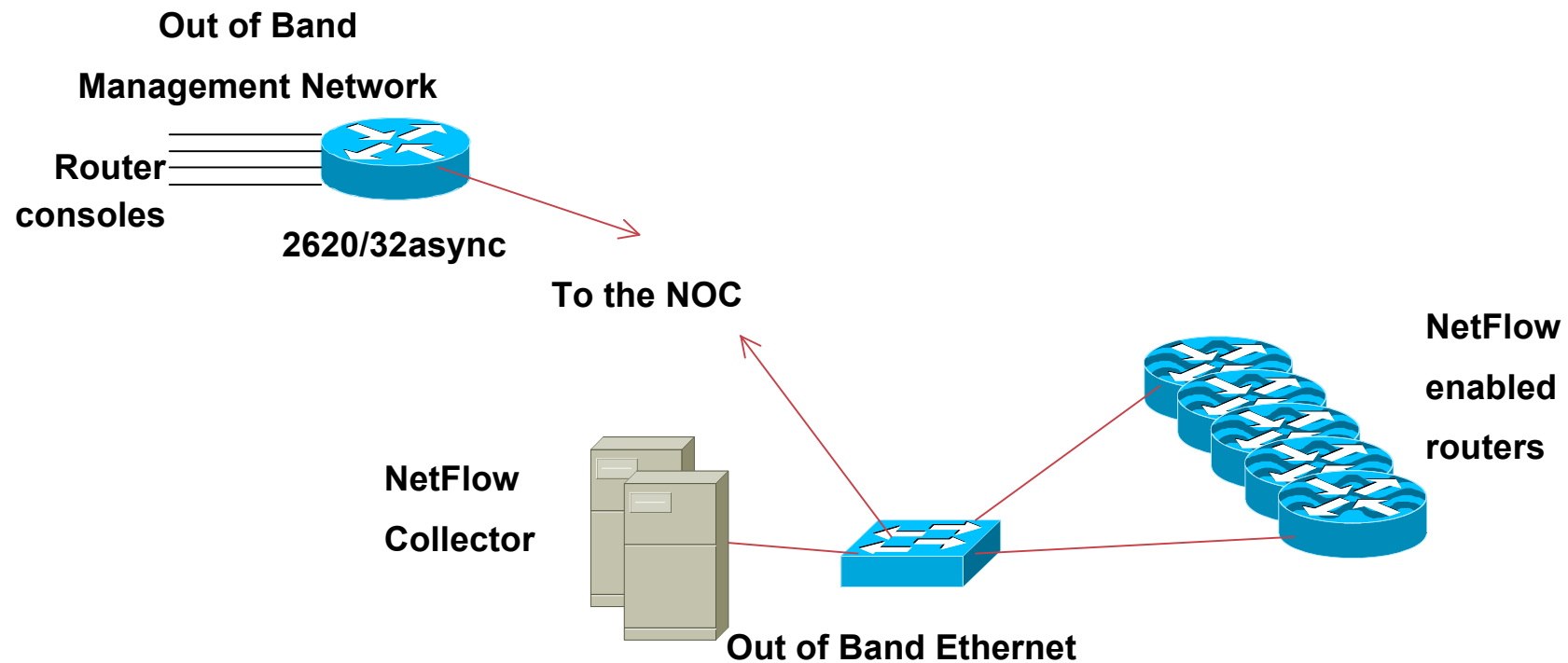
Border Module



NOC Module



Out of Band Network





Backbone Network Design

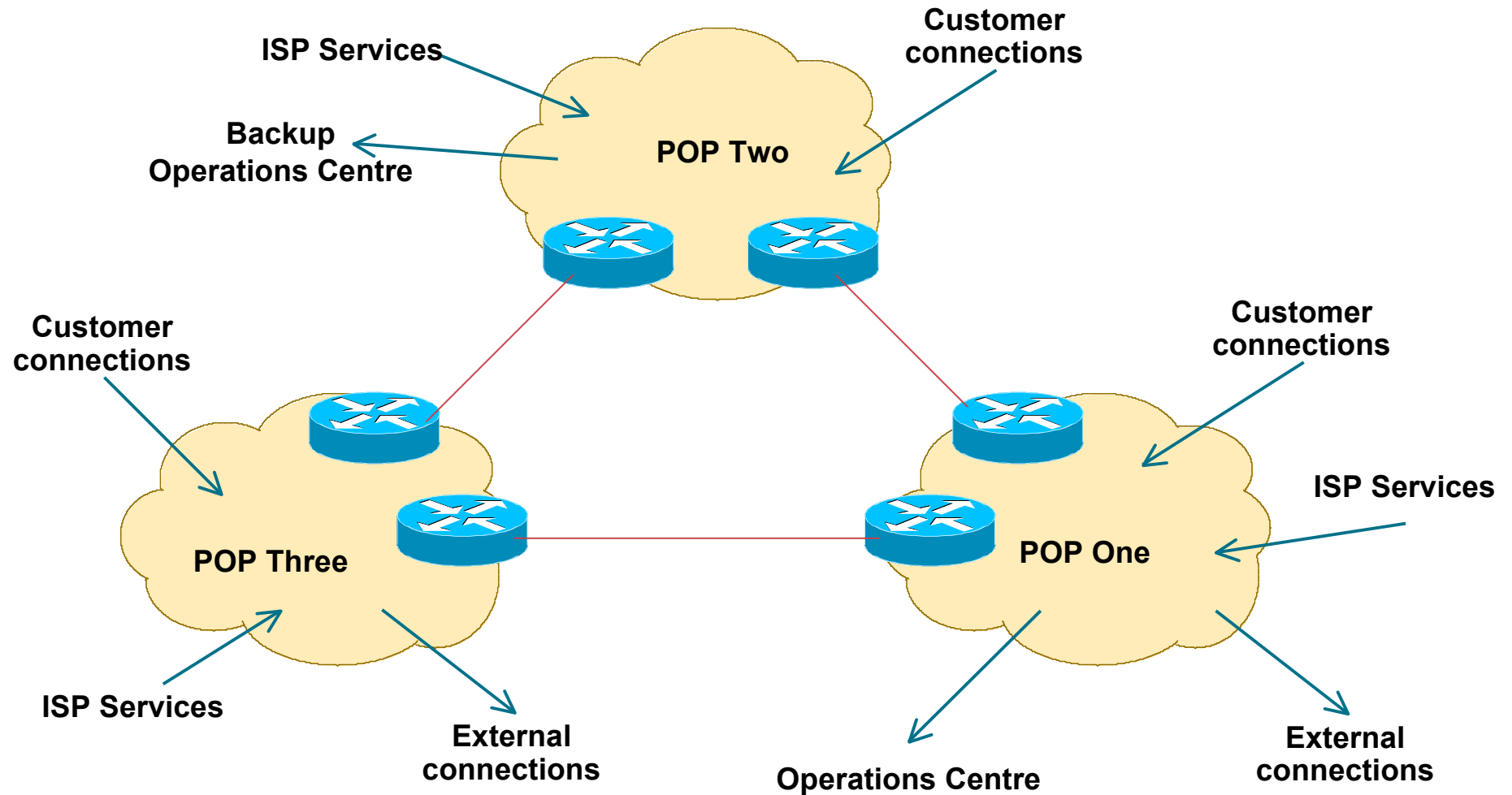
Backbone Design

- **Routed Backbone**
- **Switched Backbone**
- **Leased point-to-point circuits**
nx64K, T1/E1, T3/E3, OC3, OC12,...
- **ATM/Frame Relay service from telco**
T3, OC3, OC12,... delivery
easily upgradeable bandwidth (CIR)

Distributed Network Design

- **PoP design “standardised”**
operational scalability and simplicity
- **ISP essential services distributed around backbone**
- **NOC and “backup” NOC**
- **Redundant backbone links**

Distributed Network Design



Backbone Links

- **ATM/Frame Relay**

now less popular due to overhead, extra equipment, and shared with other customers of the telco

- **Leased Line**

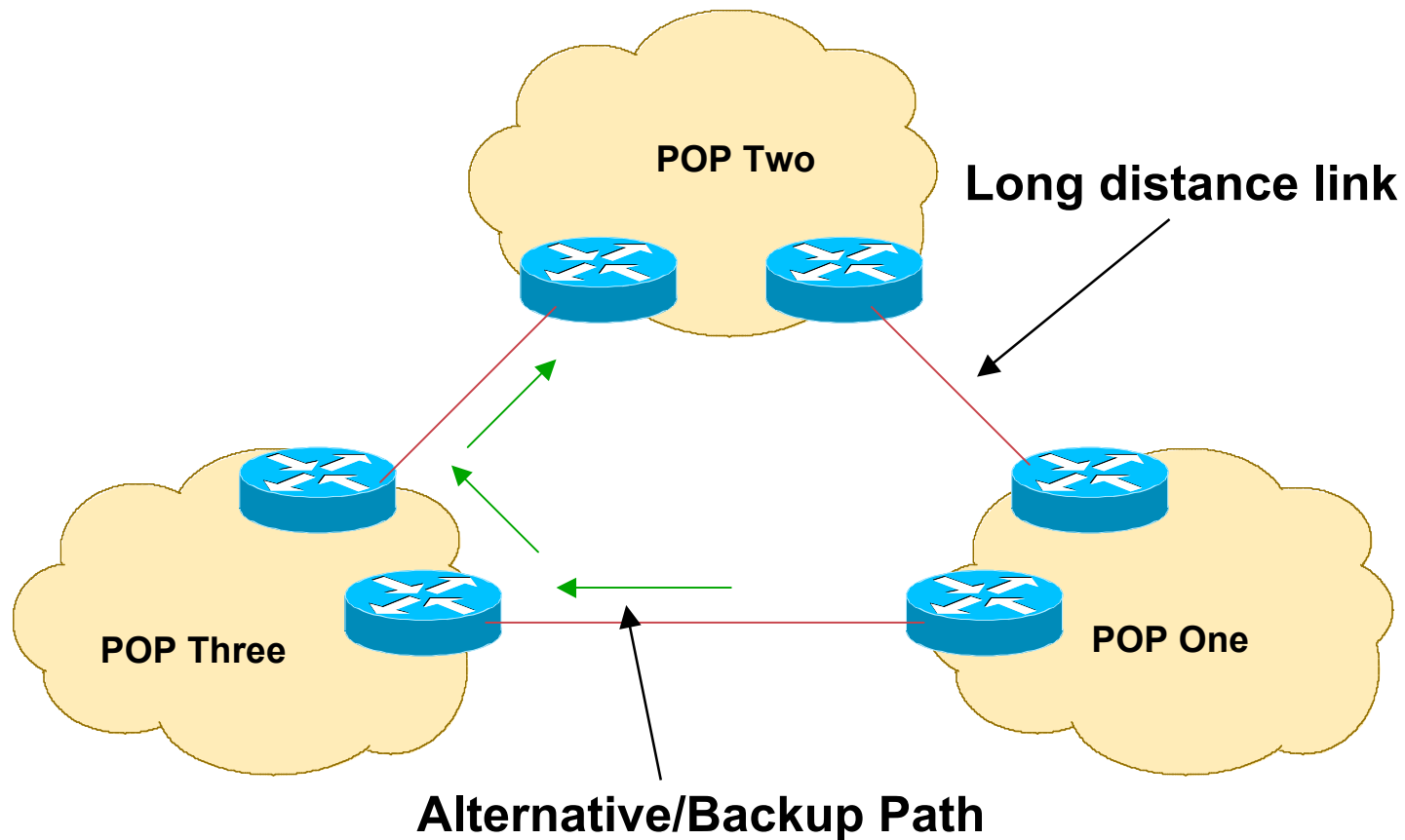
more popular with backbone providers

IP over Optics and MPLS coming into the mainstream

Long Distance Backbone Links

- **Tend to cost more**
- **Plan for the future (at least two years ahead) but stay in budget**
 - Unplanned “emergency” upgrades can be disruptive without redundancy**
- **Allow sufficient capacity on alternative paths for failure situations**
 - sufficient can be 20% to 50%**

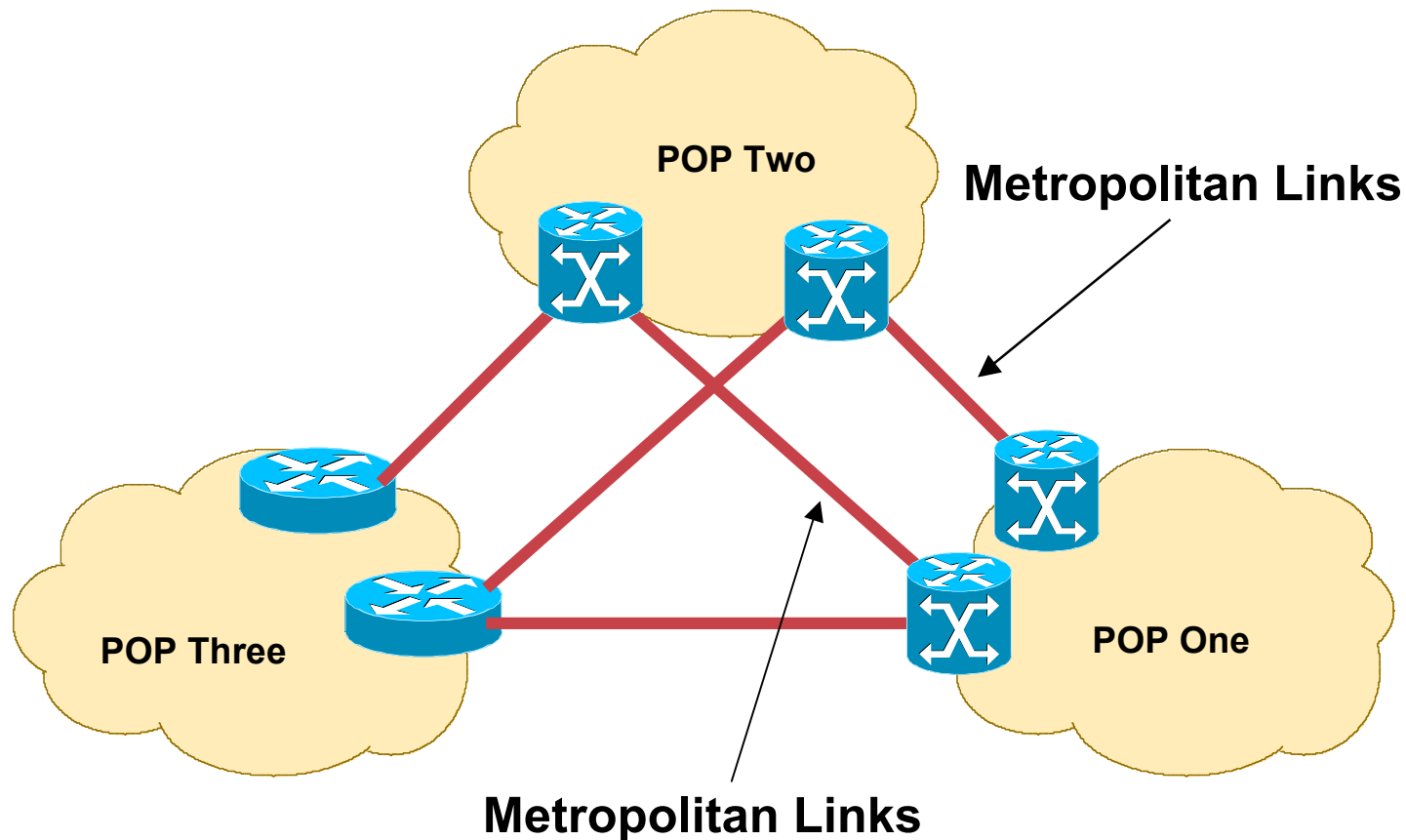
Long Distance Links



Metropolitan Area Backbone Links

- **Tend to be cheaper**
 - Circuit concentration**
 - Choose from multiple suppliers**
- **Think big**
 - More redundancy**
 - Less impact of upgrades**
 - Less impact of failures**

Metropolitan Area Backbone Links



Traditional Point to Point Links



ISP Services

**DNS, Mail, News
design and location**

ISP Services: DNS

- **Domain Name System**

Provides name and address resolution

Servers need to be differentiated, properly located and specified

Primary nameserver

Secondary nameserver

Caching nameserver – resolver

ISP Services: DNS

- **Primary nameserver**

Holds ISP zone files

forward zone (list of name to address mappings) for all ISP's and any customer zones

reverse zone (list of address to name mappings) for all ISP's address space

One Unix server, fast I/O, reasonable amount of memory (512Mbytes), reasonable disk

Located in secure part of net, e.g. NOC LAN

ISP Services: DNS

- **Secondary nameserver**

Holds copies of ISP zone files

At least two are required, more is better

Unix server, fast I/O, reasonable amount of memory (512Mbytes), reasonable disk

Should be geographically separate from each other and the primary DNS

At different PoPs

On a different continent e.g. www.secondary.com

At another ISP

ISP Services: Secondary DNS Example

- **apnic.net zone**
primary DNS in Brisbane
secondary DNS around the world

```
$ dig apnic.net ns
;; ANSWER SECTION:
apnic.net.          50m44s IN NS      svc00.apnic.net.
apnic.net.          50m44s IN NS      ns.ripe.net.
apnic.net.          50m44s IN NS      rs.arin.net.
apnic.net.          50m44s IN NS      ns.apnic.net.

;; ADDITIONAL SECTION:
svc00.apnic.net.    1d23h53m25s IN A      202.12.28.131
ns.ripe.net.        1d23h54m46s IN A      193.0.0.193
rs.arin.net.        1d23h53m25s IN A      192.149.252.21
ns.apnic.net.       1d9h29m16s IN A      203.37.255.97
```

Tokyo

Amsterdam

Washington

Brisbane

ISP Services: Secondary DNS Example

- **apnic.net zone**

primary DNS in Brisbane (ns.apnic.net)

secondary DNS run by APNIC in Tokyo (svc00.apnic.net)

zone secondaried by

RIPE NCC in Amsterdam

ARIN in Washington

Geographical and service provider redundancy – this is the perfect example!

ISP Services: DNS

- **Caching nameserver**

This is the resolver – it is the DNS cache

Your customers use this as resolver, NOT your primary or secondary DNS

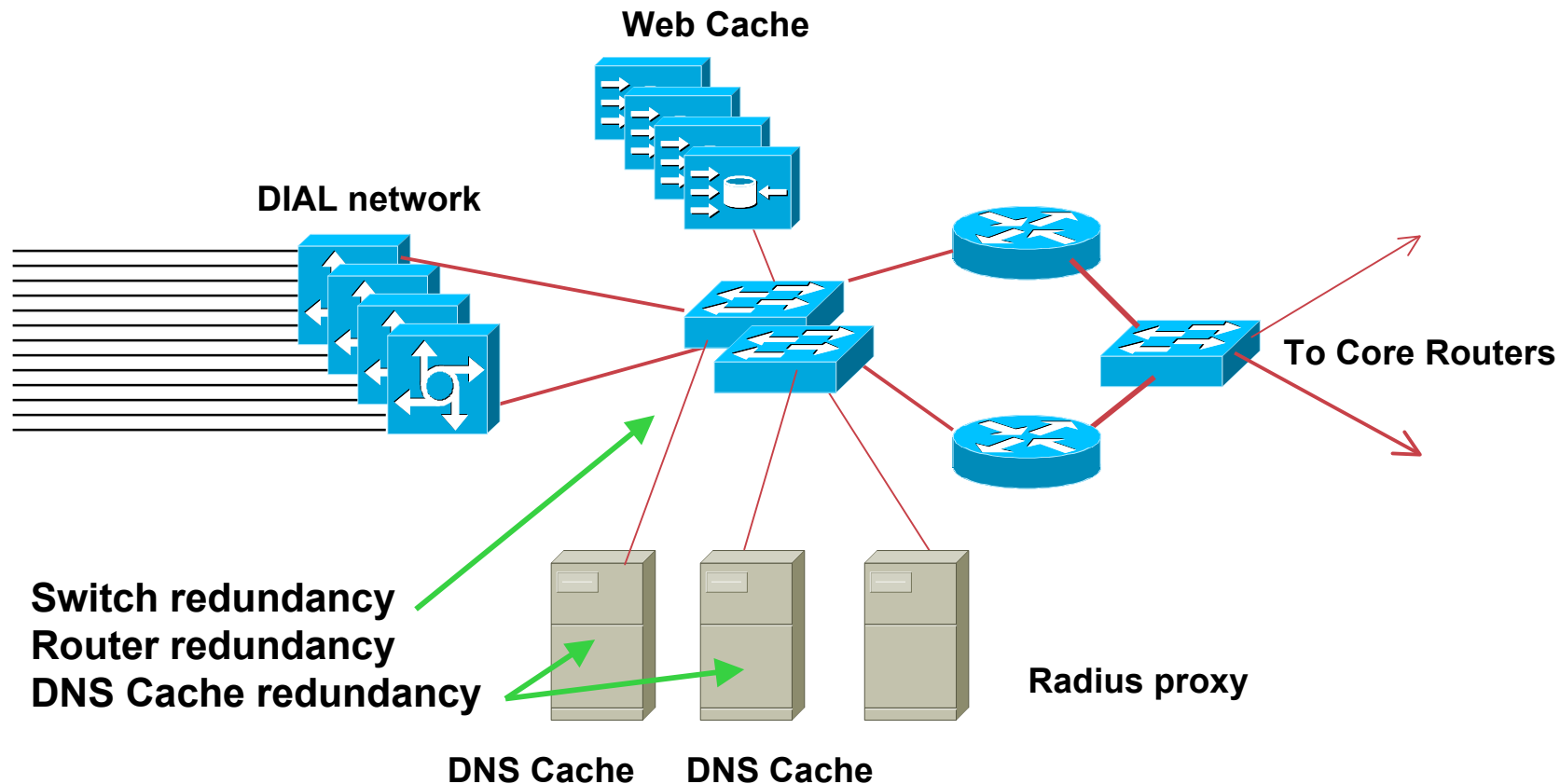
Provides very fast lookups

Does NOT secondary any zones

One, or preferably two per PoP (redundancy)

Unix server, fast I/O, large amount of memory (512Mbytes+ depending on number of zones)

ISP Services: Caching Nameserver



**DIAL users automatically given the IP addresses of
DNS caches when they dial in**

ISP Services:

Anycasting the Caching Nameserver

- **One trick of the trade**

**assign two unique IP addresses to be
for the two DNS resolver systems**

use these two IP addresses in every PoP

route the two /32s across your backbone

**even if the two resolver systems in the local PoP are down,
the IGP will ensure that the next nearest resolvers will be
reachable**

Known as IP Anycast



ISP Services: DNS

- **Efficient and resilient design**

Primary DNS – keep it secure

Secondary DNS – geographical and provider redundancy

Don't ever put them on the same LAN, switched or otherwise

Don't put them in the same PoP

Caching DNS – one or two per PoP

reduces DNS traffic across backbone

more efficient, spreads the load

ISP Services: DNS

- **Software**

Make sure that the BIND distribution on the Unix system is up to date

the vendor's distribution is rarely current

Pay attention to bug reports, security issues

Reboot the DNS cache on a regular (e.g. monthly) basis

clears out the cache

releases any lost RAM

accepted good practice by system administrators

ISP Services: DNS

- **Implementation**

Put all your hosts, point-to-point links and loopbacks into the DNS

under your ISP's domain name

use sensible/meaningful names

Put all your hosts, point-to-point links and loopbacks into the REVERSE DNS also

don't forget about in-addr.arpa – many ISPs do

some systems demand forward/reverse DNS mapping before allowing access

ISP Services: Mail

- **Must have at least two mail hosts (MX records) for all supported domains**
geographical separation helps
- **POP3 server dedicated to that function**
DIAL users get mail from here
- **SMTP gateway dedicated to that function**
DIAL users send mail via here
- **Mail relay open to CUSTOMERS only!**
- **Block port 25 outbound for all customers**
Insist that outbound e-mail goes through SMTP relay

ISP Services: Mail Example

- **telstra.net mail (MX records)**
 - primary MX is mako1**
 - backup MX is postoffice – two addresses**
 - backup MX used if primary unavailable**

```
$ dig telstra.net mx

;; ANSWER SECTION:
telstra.net.          1H IN MX 10 postoffice.telstra.net.
telstra.net.          1H IN MX  5 mako1.telstra.net.

;; ADDITIONAL SECTION:
postoffice.telstra.net. 1H IN A 139.130.4.7
postoffice.telstra.net. 1H IN A 203.50.1.76
mako1.telstra.net.     1H IN A 203.50.0.28
```

ISP Services: Mail

- **Software**

Make sure that the MAIL and POP3 distributions on the Unix system are up to date

the vendor's distribution are rarely current

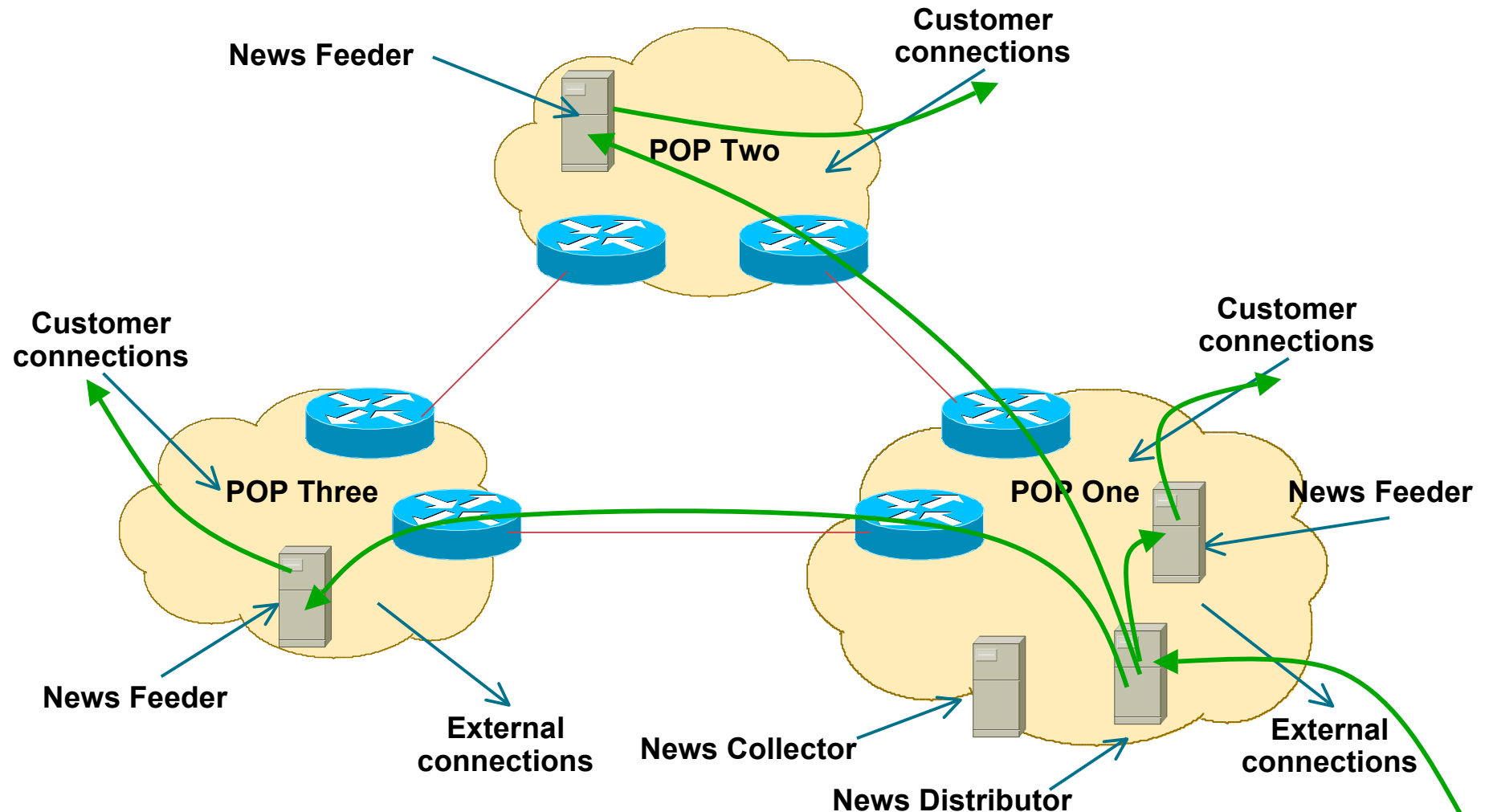
Pay attention to bug reports, security issues, unsolicited junk mail complaints

IMPORTANT: Do NOT allow non-customers to use your mail system as a relay

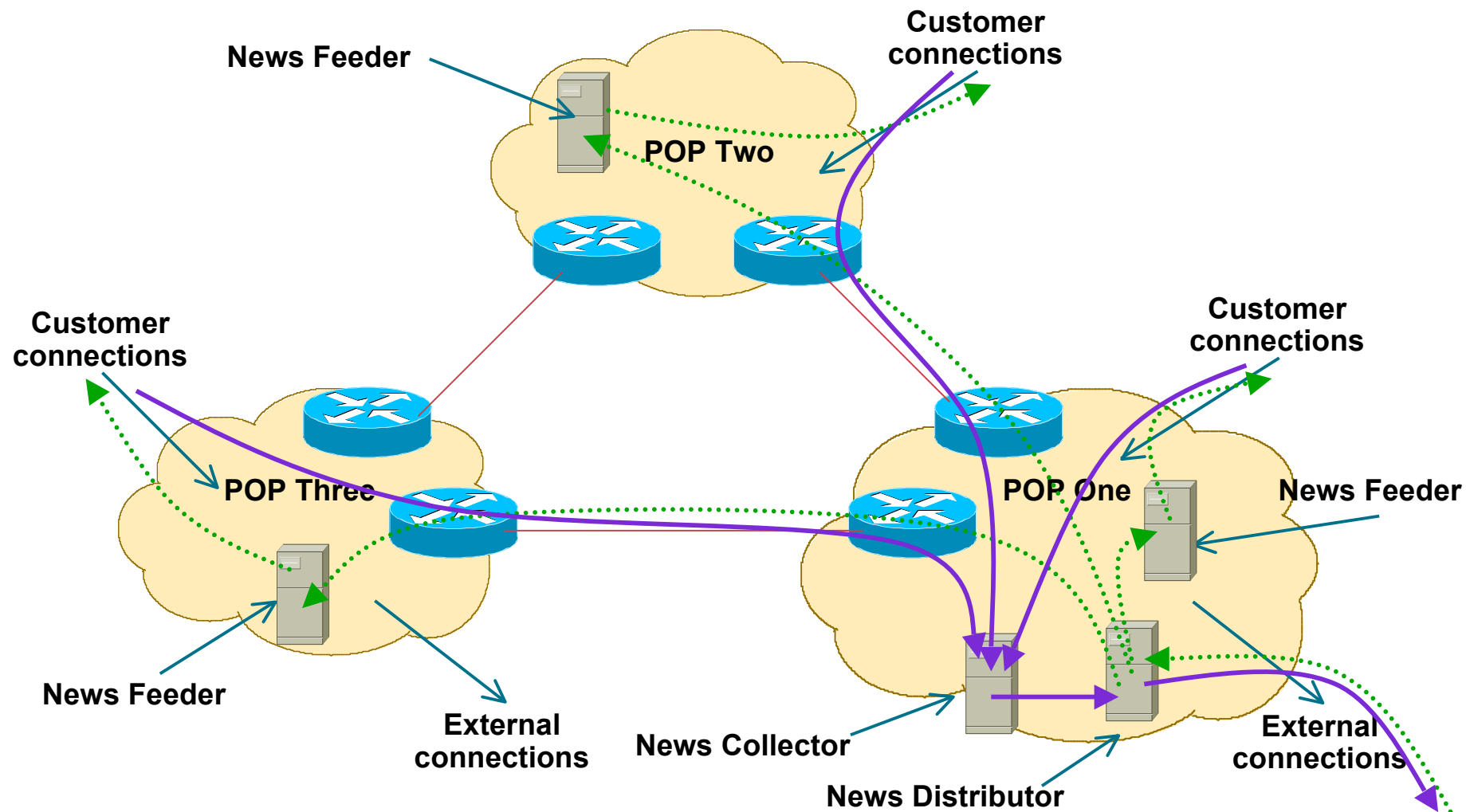
ISP Services: News

- **News servers provide a Usenet news feed to customers**
- **Distributed design required**
 - Incoming newsfeed to one large server**
 - Distributed to feed servers in each PoP**
 - Feed servers provide news feed to customers**
 - Outgoing news goes to another server**
 - Separate reading news system**
 - Separate posting news system**

ISP Services: News System Placement



ISP Services: News System Placement



ISP Services: News

- **Software**

Make sure that the Internet News distribution on the Unix system is up to date

the vendor's distribution is rarely current

Pay attention to bug reports, security issues, unsolicited junk posting complaints

IMPORTANT: Do NOT allow non-customers to use your news system for posting messages



Addressing

Where to get IP addresses and AS numbers

- **Your upstream ISP**
- **Africa**
AfriNIC – <http://www.afrinic.net>
- **Asia and the Pacific**
APNIC – <http://www.apnic.net>
- **North America**
ARIN – <http://www.arin.net>
- **Latin America and the Caribbean**
LACNIC – <http://www.lacnic.net>
- **Europe and Middle East**
RIPE NCC – <http://www.ripe.net>

Internet Registry Regions



Getting IP address space

- Take part of upstream ISP's PA space

or

- Become a member of your Regional Internet Registry and get your own allocation

Require a plan for a year ahead

General policies are outlined in RFC2050, more specific details are on the individual RIR website

- **There is plenty of IPv4 address space**

registries require high quality documentation

Addressing Plans – ISP Infrastructure

- **Address block for router loop-back interfaces**
- **Address block for infrastructure**
 - per PoP or whole backbone**
 - summarise between sites if it makes sense**
 - allocate according to genuine requirements, not historic classful boundaries**

Addressing Plans – Customer

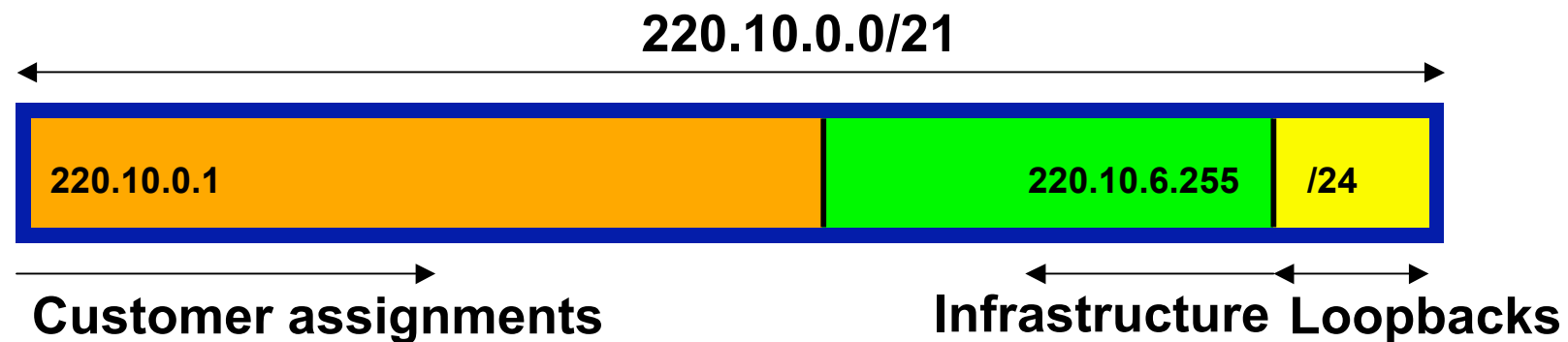
- **Customers assigned address space according to need**
- **Should not be reserved or assigned on a per PoP basis**

ISP iBGP carries customer nets

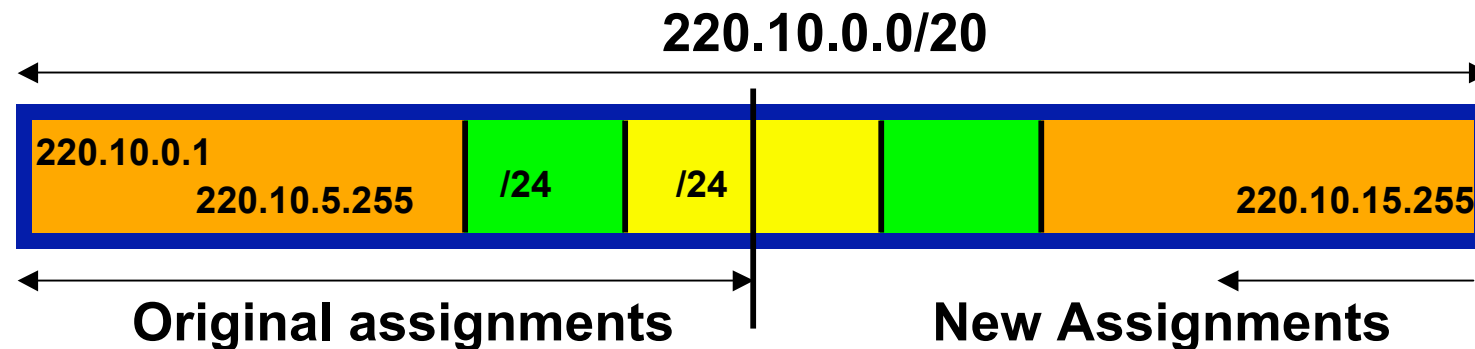
aggregation not required and usually not desirable

Addressing Plans – ISP Infrastructure

Phase One



Phase Two



Addressing Plans Planning

- **Registries will usually allocate the next block to be contiguous with the first allocation**

Minimum allocation is /21

Very likely that subsequent allocation will make this up to a /20

So plan accordingly

Addressing Plans (contd)

- **Document infrastructure allocation**
eases operation, debugging and management
- **Document customer allocation**
contained in iBGP
eases operation, debugging and management
submit network object to RIR Database



Routing Protocols

Routing Protocols

- **IGP – Interior Gateway Protocol**
carries infrastructure addresses, point-to-point links
examples are OSPF, ISIS, EIGRP...
- **EGP – Exterior Gateway Protocol**
carries customer prefixes and Internet routes
current EGP is BGP version 4
- **No link between IGP and EGP**

Why Do We Need an IGP?

- **ISP backbone scaling**

Hierarchy

Modular infrastructure construction

Limiting scope of failure

Healing of infrastructure faults using dynamic routing with fast convergence

Why Do We Need an EGP?

- **Scaling to large network**

Hierarchy

Limit scope of failure

- **Policy**

Control reachability to prefixes

Merge separate organizations

Connect multiple IGPs

Interior versus Exterior Routing Protocols

- **Interior**

- automatic neighbour discovery**
 - generally trust your IGP routers**
 - prefixes go to all IGP routers**
 - binds routers in one AS together**

- **Exterior**

- specifically configured peers**
 - connecting with outside networks**
 - set administrative boundaries**
 - binds AS's together**

Interior versus Exterior Routing Protocols

- **Interior**

Carries ISP infrastructure addresses only

ISPs aim to keep the IGP small for efficiency and scalability

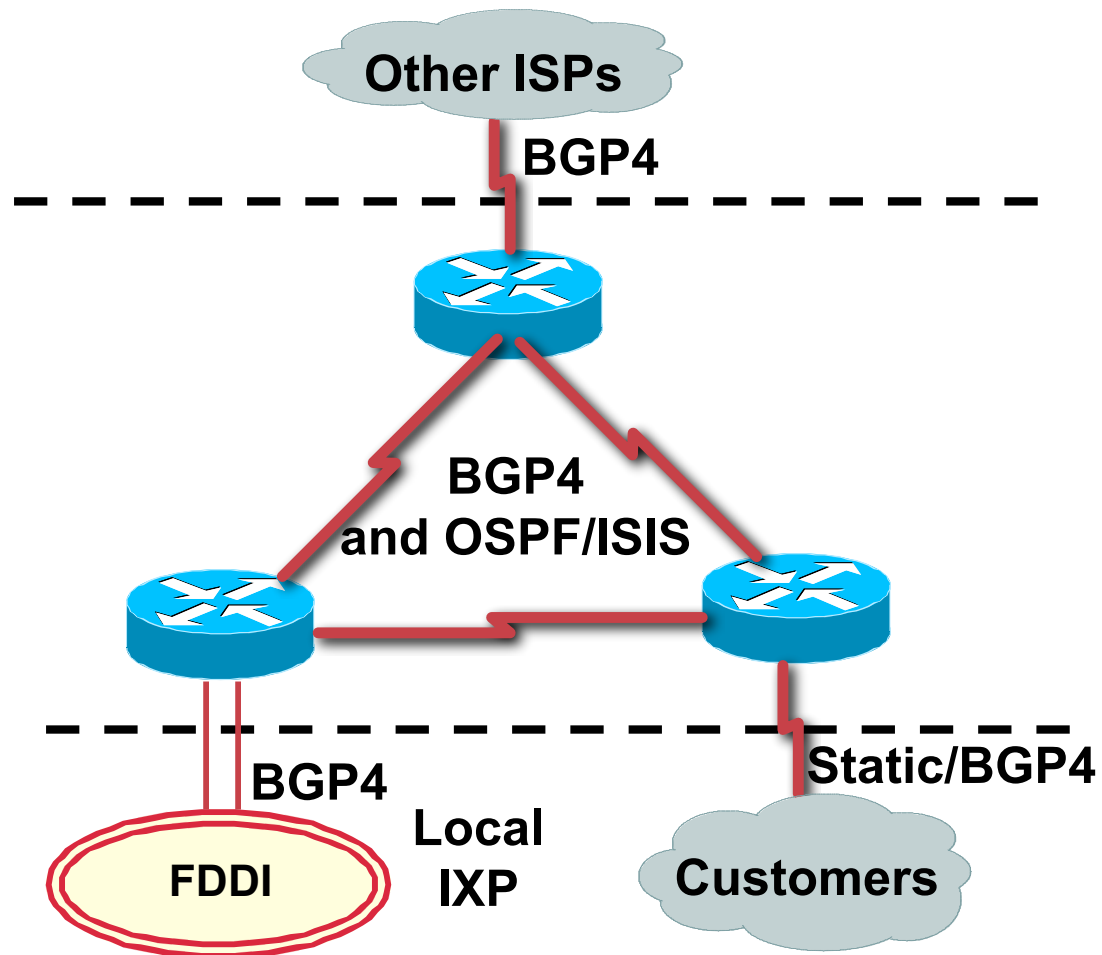
- **Exterior**

Carries customer prefixes

Carries Internet prefixes

EGPs are independent of ISP network topology

Hierarchy of Routing Protocols



Routing Protocols: Choosing an IGP

- **Review the “Introduction to Link State Protocols” presentation**

i.e. – OSPF and ISIS have very similar properties

- **ISP usually chooses between OSPF and ISIS**

Choose which is appropriate for your operators’ experience

In IOS, both OSPF and ISIS have sufficient “nerd knobs” to tweak the IGP’s behaviour

Routing Protocols: IGP Recommendations

- **Keep the IGP routing table as small as possible**

If you can count the routers and the point to point links in the backbone, that total is the number of IGP entries you should see

- **IGP details:**

Should only have router loopbacks, backbone WAN point-to-point link addresses, and network addresses of any LANs having an IGP running on them

Strongly recommended to use inter-router authentication

Use inter-area summarisation if possible

Routing Protocols:

More IGP recommendations

- **To fine tune IGP table size more, consider:**

Using “ip unnumbered” on customer point-to-point links – saves carrying that /30 in IGP

(If customer point-to-point /30 is required for monitoring purposes, then put this in iBGP)

Use contiguous addresses for backbone WAN links in each area – can then summarise into backbone area

Don't summarise router loopback addresses – as iBGP needs those

Use iBGP for carrying anything which does not contribute to the Link State Routing process

Routing Protocols: iBGP Recommendations

- **iBGP should carry everything which doesn't contribute to the IGP routing process**

Internet routing table

Customer assigned addresses

Customer point-to-point links

DIAL network pools, passive LANs, etc

Routing Protocols: More iBGP Recommendations

- **Scalable iBGP features:**

- Use neighbour authentication**

- Use peer-groups to speed update process and for configuration efficiency**

- Use communities for ease of filtering**

- Use route-reflector hierarchy**

- Route reflector pair per PoP (overlaid clusters)**

- Use route flap damping at the network edges**



Security

Security

- **ISP Infrastructure security**
- **ISP Network security**
- **Security is not optional!**
- **ISPs need to:**
 - protect themselves**
 - help protect their customers from the Internet**
 - protect the Internet from their customers**
- **The following slides are general recommendations**
 - do more research on security before deploying any network**

ISP Infrastructure Security

- **router security**

username, passwords, vty filters, TACACS+

Disable telnet on vtys, only use SSH

vtty filters should only allow NOC access, no external access

See IOS Essentials for the recommended practices for ISPs

ISP Infrastructure Security

- **ISP server security**

usernames, passwords, TCP wrappers, IPTABLES

protect *all* servers using routers with strong filters applied

- **Hosted services security**

protect network from hosted servers using routers with strong filters

protect hosted servers from Internet using routers with strong filters

ISP Infrastructure Security

ISP Server Protection

Access-list examples:

Allow tcp/established to all servers

ICMP

DNS 2ary: udp/53 and tcp/53

POP3: tcp/110

Mail Relay: tcp/25 and ISP address range only

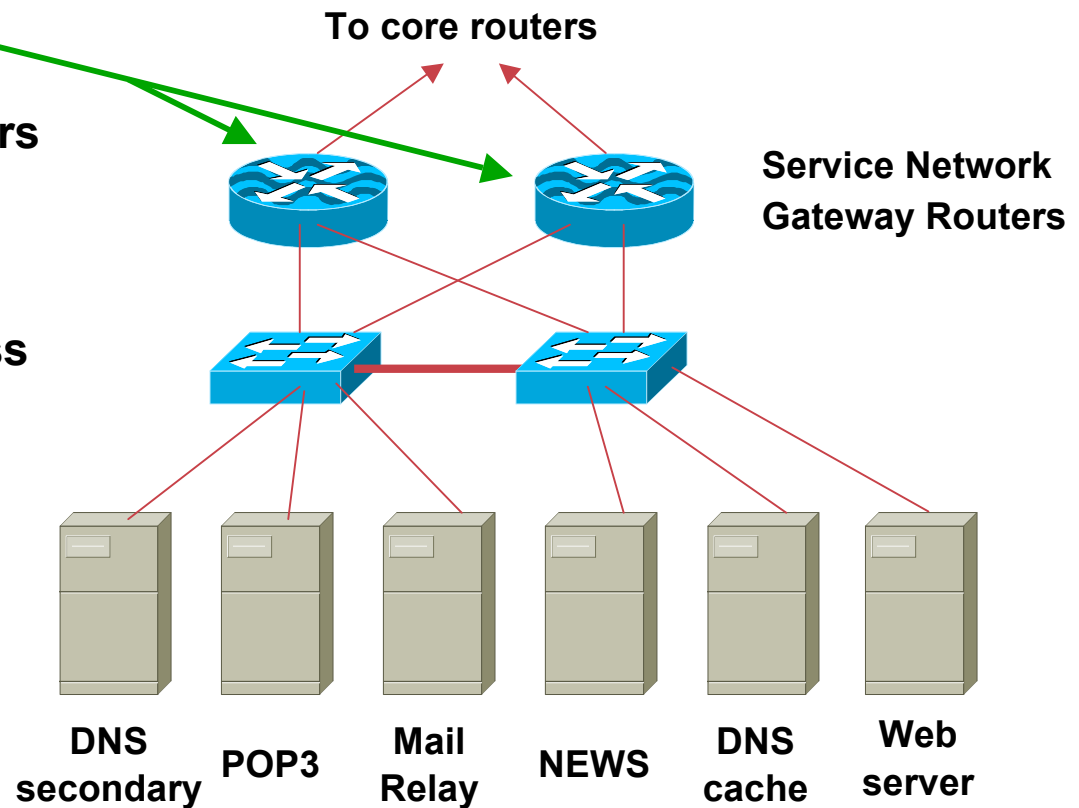
News: tcp/119 and ISP address range only

DNS Cache: udp/53

Web server: tcp/80

Other necessary filters:

All servers: SSH (tcp/22) from NOC LAN only



ISP Infrastructure Security

Hosted Server Protection

Access-list examples:

Inbound

Allow tcp/established to all servers

ICMP

Web server: tcp/80

SSH for customer access

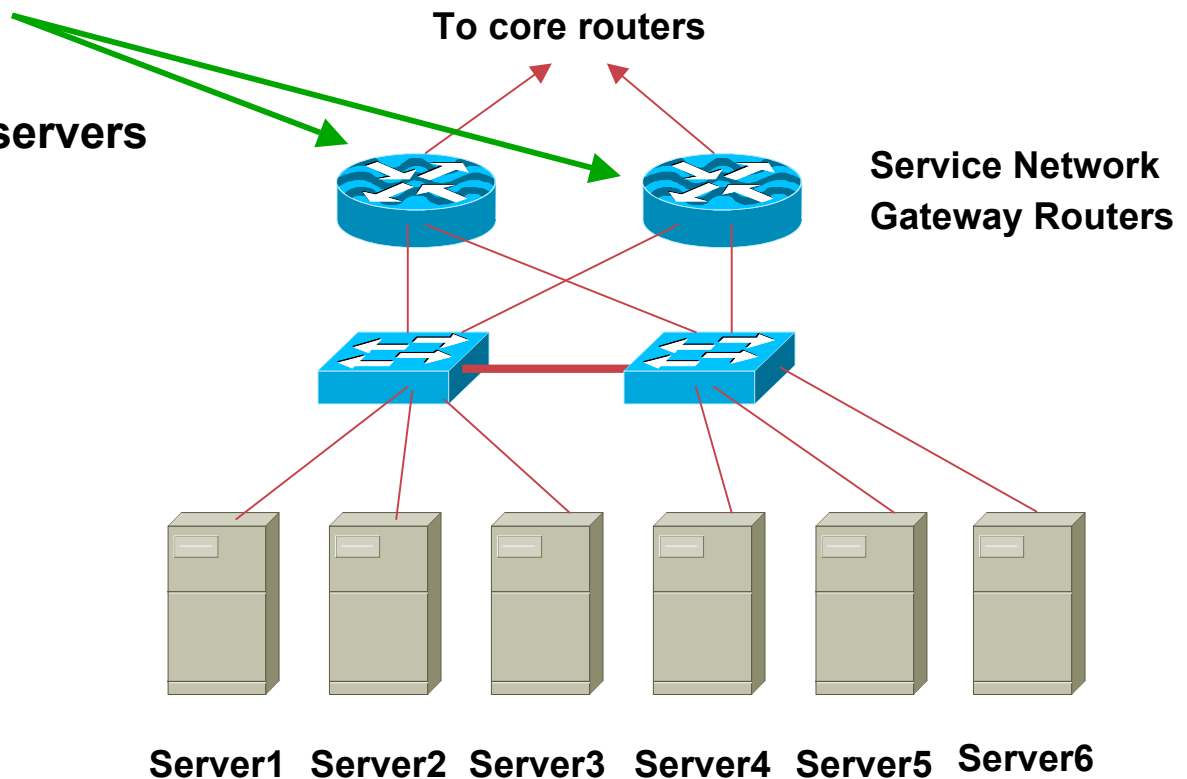
Any other ports for services
sold to customers

Outbound

ICMP

Allow DNS udp/53 and
tcp/53

Block all access to ISP
address range



ISP Infrastructure Security

- **premises security**
 - locks – electronic/card key preferred**
 - secure access – 24x7 security arrangements**
 - environment control – good aircon**
- **staff responsibility**
 - password policy, strangers, temp staff**
 - employee exit procedures**
- **RFC2196**
 - (Site Security Handbook)**
- **RFC3871**
 - (Operational Security Requirements for Large ISP IP Network Infrastructure)**

ISP Network Security

- **Denial of Service Attacks**

eg: “smurfing”

see <http://www.denialinfo.com>

- **Effective filtering**

network borders – see Cisco ISP Essentials

customer connections – unicast RPF

network operation centre

ISP corporate network – behind firewall

ISP Network Security

Secure external access

- **How to provide staff access from outside**
 - set up ssh gateway (Unix system with ssh daemon and nothing else configured)**
 - provide ssh client on all staff laptops**
 - ssh available on Unix and Windows**
 - ssh is Secure Shell – encrypted link**
- **How not to provide access from outside**
 - telnet, rsh, rlogin – these are all insecure**
 - open host – insecure, can be compromised**

Ingress & Egress Route Filtering

Your customers should not be sending *any* IP packets out to the Internet with a source address other than the address you have allocated to them!



Out of Band Management

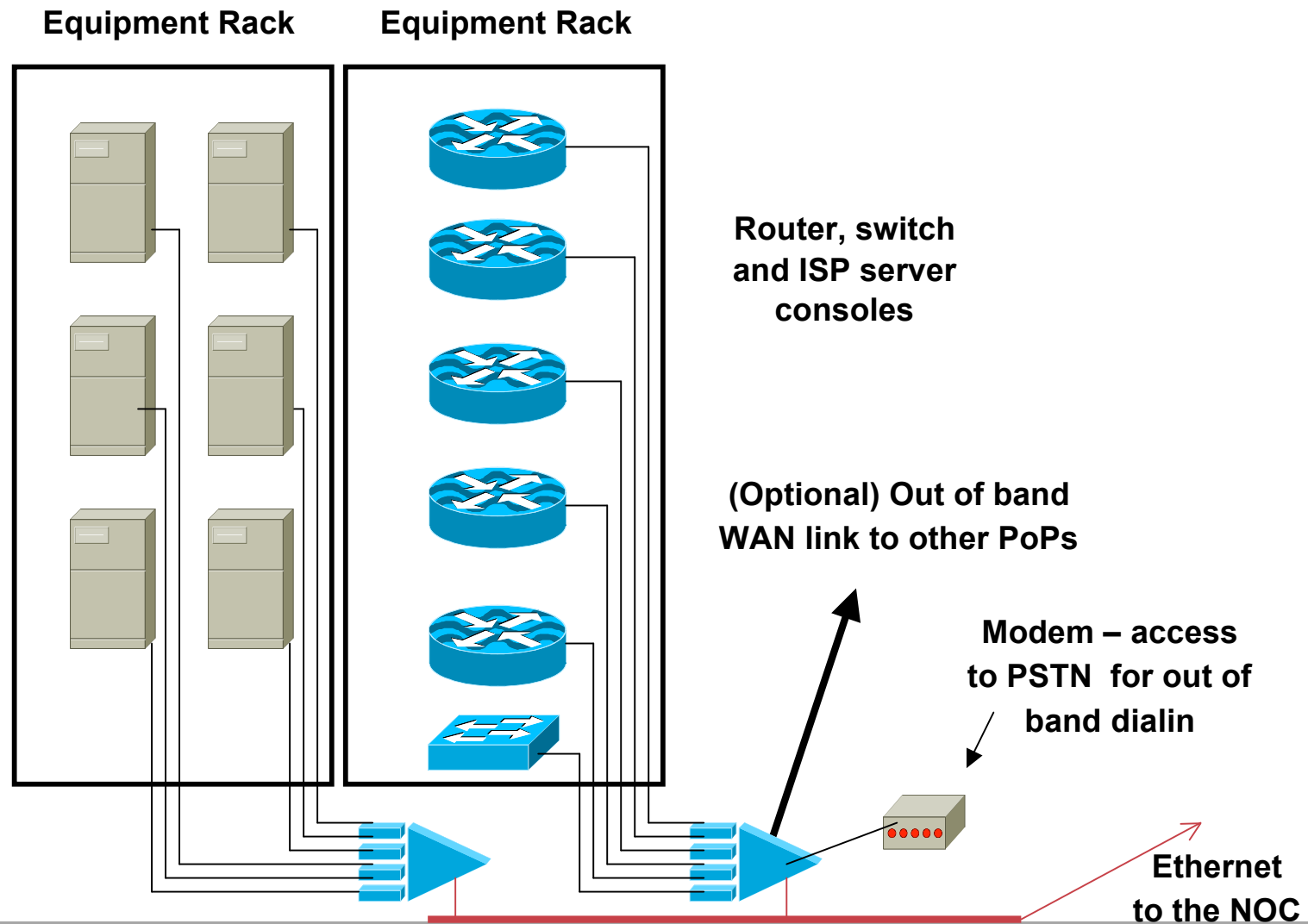
Out of Band Management

- **Not optional!**
- **Allows access to network equipment in times of failure**
- **Ensures quality of service to customers**
 - minimises downtime**
 - minimises repair time**
 - eases diagnostics and debugging**

Out of Band Management

- **OoB Example – Access server:**
 - modem attached to allow NOC dial in**
 - console ports of all network equipment connected to serial ports**
 - LAN and/or WAN link connects to network core, or via separate management link to NOC**
- **Full remote control access under all circumstances**

Out of Band Network



Out of Band Management

- **OoB Example – Statistics gathering:**
 - Routers are NetFlow and syslog enabled**
 - Management data is congestion/failure sensitive**
 - Ensures management data integrity in case of failure**
- **Full remote information under all circumstances**



Test Laboratory

Test Laboratory

- **Designed to look like a typical PoP
operated like a typical PoP**
- **Used to trial new services or new software under realistic conditions**
- **Allows discovery and fixing of potential problems before they are introduced to the network**

Test Laboratory

- **Some ISPs dedicate equipment to the lab**
- **Other ISPs “purchase ahead” so that today’s lab equipment becomes tomorrow’s PoP equipment**
- **Other ISPs use lab equipment for “hot spares” in the event of hardware failure**

Test Laboratory

- **Can't afford a test lab?**

Set aside one spare router and server to trial new services

Never ever try out new hardware, software or services on the live network

- **Every major ISP in the US and Europe has a test lab**

It's a serious consideration



Operational Considerations

Operational Considerations

Why design the world's best network when you have not thought about what operational good practices should be implemented?

Operational Considerations

Maintenance

- **Never work on the live network, no matter how trivial the modification may seem**

Establish maintenance periods which your customers are aware of

e.g. Tuesday 4-7am, Thursday 4-7am

- **Never do maintenance on a Friday**
Unless you want to work all weekend cleaning up
- **Never do maintenance on a Monday**
Unless you want to work all weekend preparing

Operational Considerations Support

- **Differentiate between customer support and the Network Operations Centre**

Customer support fixes customer problems

NOC deals with and fixes backbone and Internet related problems

- **Network Engineering team is last resort**

they design the next generation network, improve the routing design, implement new services, etc

they do not and should not be doing support!

Operational Considerations

NOC Communications

- **NOC should know contact details for equivalent NOCs in upstream providers and peers**
- **Or consider joining the INOC-DBA system**

Voice over IP phone system using SIP

Runs over the Internet

www.pch.net/inoc-dba for more information



ISP Network Design

Summary

ISP Design Summary

- **KEEP IT SIMPLE & STUPID ! (KISS)**
- **Simple is elegant is scalable**
- **Use Redundancy, Security, and Technology to make life easier for yourself**
- **Above all, ensure quality of service for your customers**



ISP Network Design

ISP/IXP Workshops