Cisco.com

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**
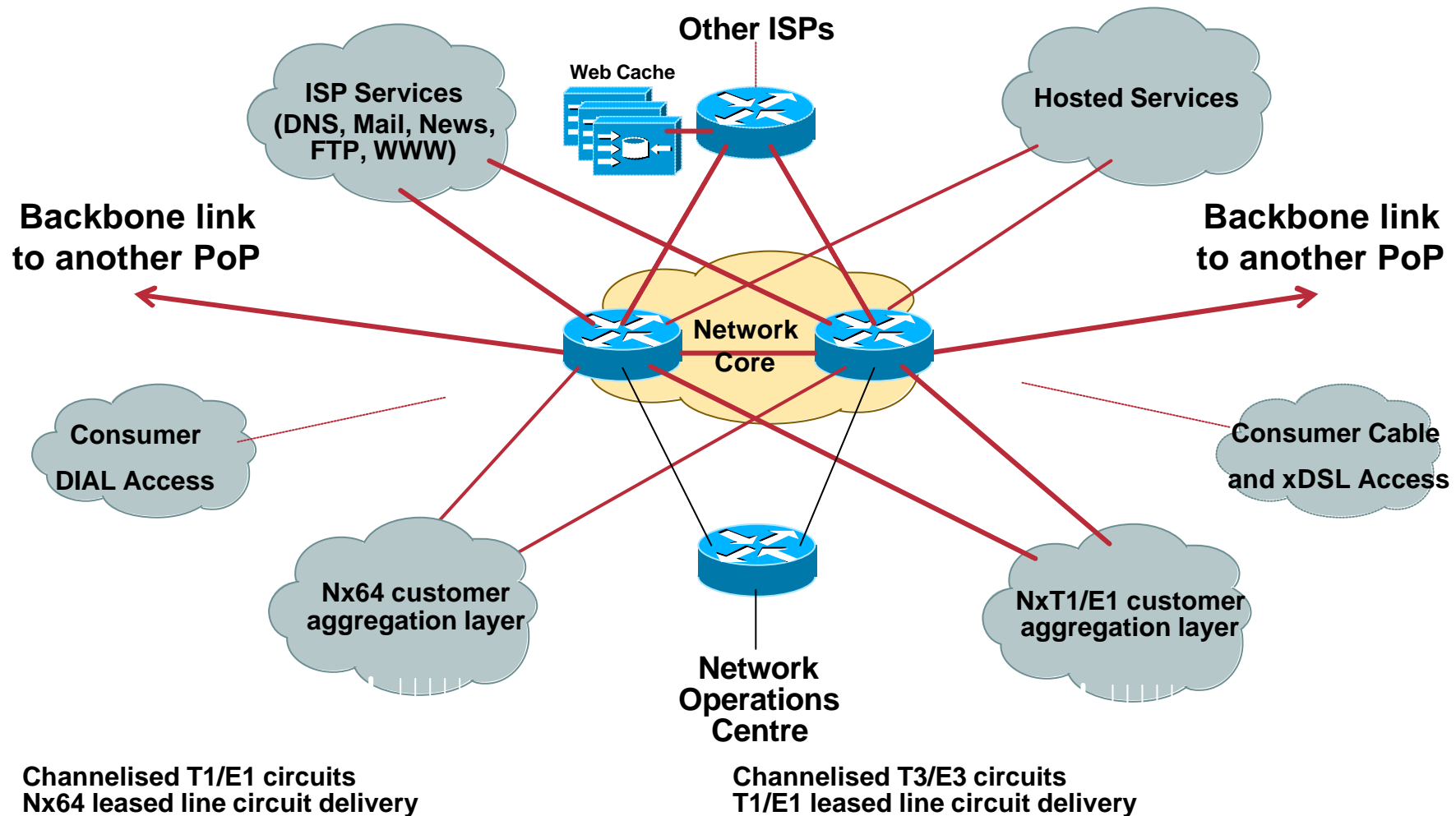
Point of Presence Topologies

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

Other ISPs

Web Cache

ISP Services (DNS, Mail, News, FTP, WWW)

Hosted Services

Backbone link to another PoP

Network Core

Backbone link to another PoP

Consumer DIAL Access

Consumer Cable and xDSL Access

Nx64 customer aggregation layer

NxT1/E1 customer aggregation layer

Network Operations Centre

Channelised T1/E1 circuits
Nx64 leased line circuit delivery

Channelised T3/E3 circuits
T1/E1 leased line circuit delivery

Modular Routing Protocol Design

- **Modular IGP implementation**

  IGP "area" per module

  aggregation/summarisation 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 and Cable

  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)

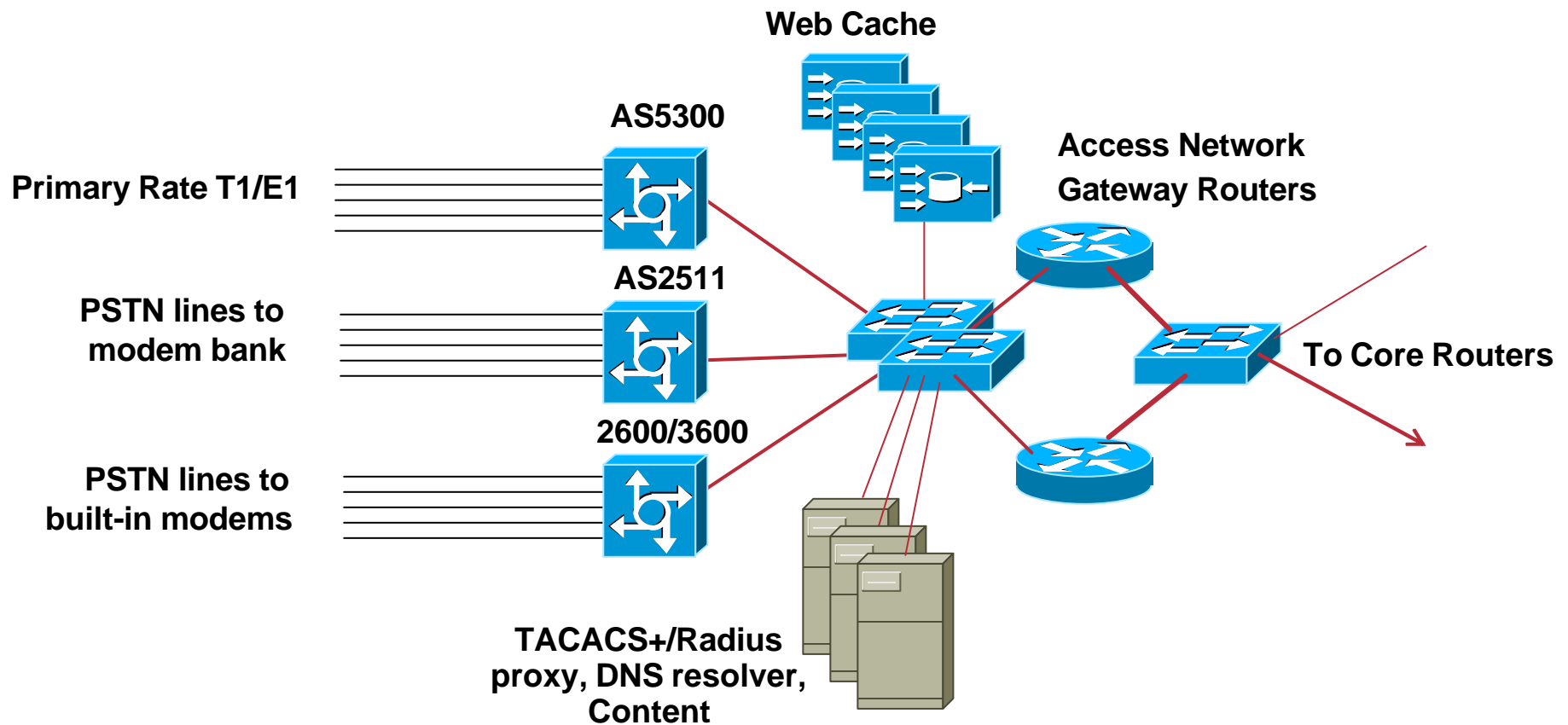- ## Hosted Services

    Virtual Web, WWW (server, proxy, cache)
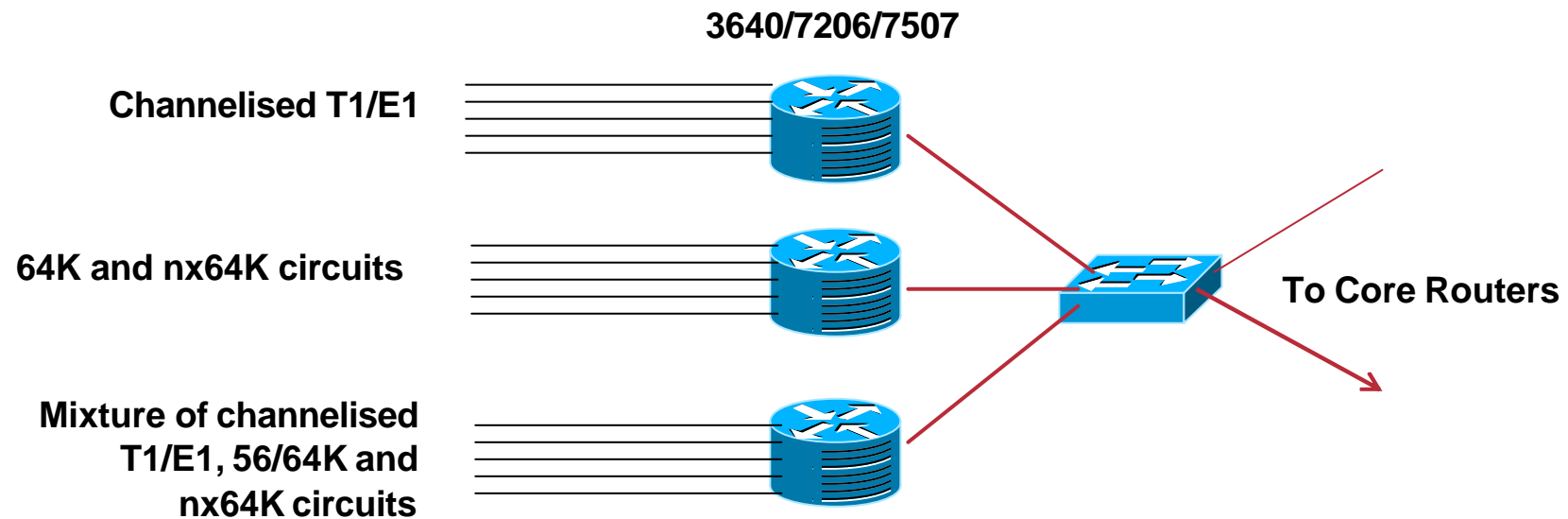
    Information/Content Services

    Electronic Commerce

PoP Modules

- ## Network Operations Centre

    primary and backup locations

    network monitoring

    statistics and log gathering

    direct but secure access

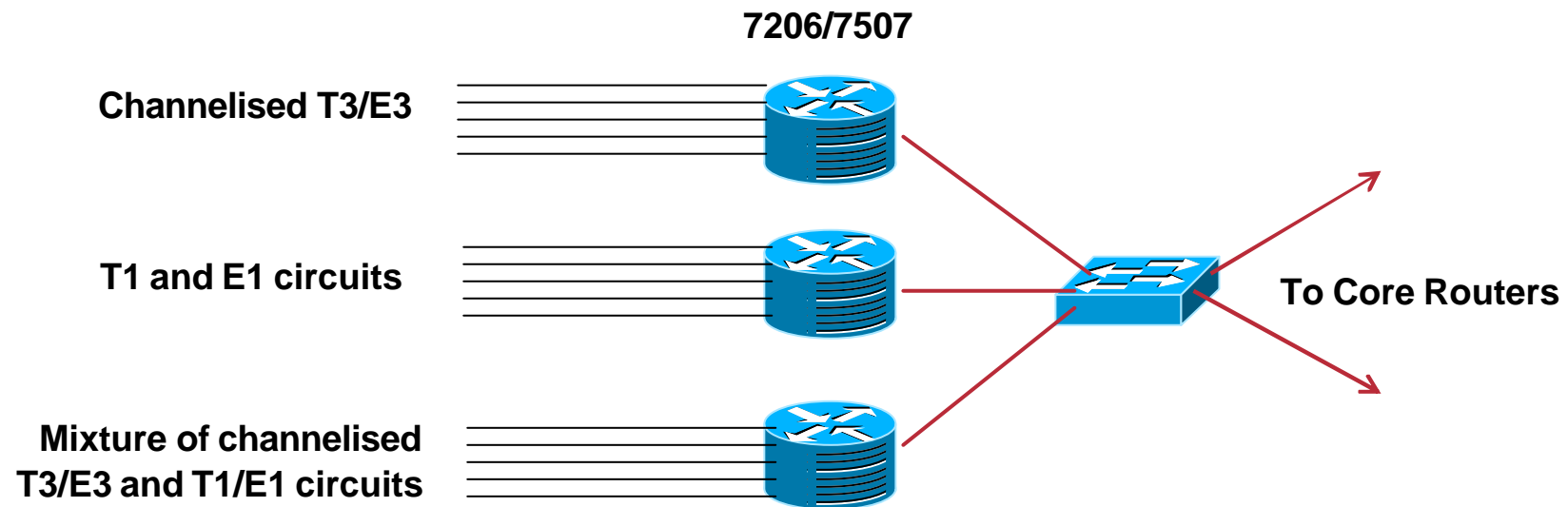- ## Out of Band Management Network

    The ISP Network "Safety Belt"

Low Speed Access Module

Web Cache

AS5300

Primary Rate T1/E1

**Access Network
Gateway Routers**

AS2511

**PSTN lines to
modem bank**

To Core Routers

2600/3600

**PSTN lines to
built-in modems**

**TACACS+/Radius
proxy, DNS resolver,
Content**

Medium Speed Access Module

3640/7206/7507

Channelised T1/E1

64K and nx64K circuits

Mixture of channelised
T1/E1, 56/64K and
nx64K circuits

To Core Routers

High Speed Access Module

7206/7507

Channelised T3/E3

T1 and E1 circuits

**Mixture of channelised
T3/E3 and T1/E1 circuits**

To Core Routers

Broad Band Access Module

Web Cache

Telephone Network

61xx

6400

IP, ATM

**Access Network
Gateway Routers**

uBR7246

The cable system

To Core Routers

**SSG, DHCP, TACACS+
or Radius Servers/Proxies,
DNS resolver, Content**

ISP Services Module

To core routers

Service Network Gateway Routers

WWW cache

DNS secondary

POP3

Mail Relay

NEWS

DNS cache

Hosted Services Module

To core routers

Hosted Network
Gateway Routers

Customer 1

Customer 2

Customer 3

Customer 4

Customer 5

Customer 6

Customer 7

Border Module

**To local IXP -
NB - no default route +
local AS routing table only**

ISP1   **ISP2**

**Network
Border Routers**

To core routers

NOC Module

Critical Services Module

To core routers

Out of Band

Management Network

Hosted Network Gateway Routers

Corporate LAN

Firewall

2620/32async

NetFlow Analyser

TACACS+ server

SYSLOG server

Primary DNS

Billing, Database and Accounting Systems

Network Operations Centre Staff

Out of Band Network

Out of Band

Management Network

Router consoles

2620/32async

To the NOC

NetFlow Collector

Out of Band Ethernet

NetFlow enabled routers
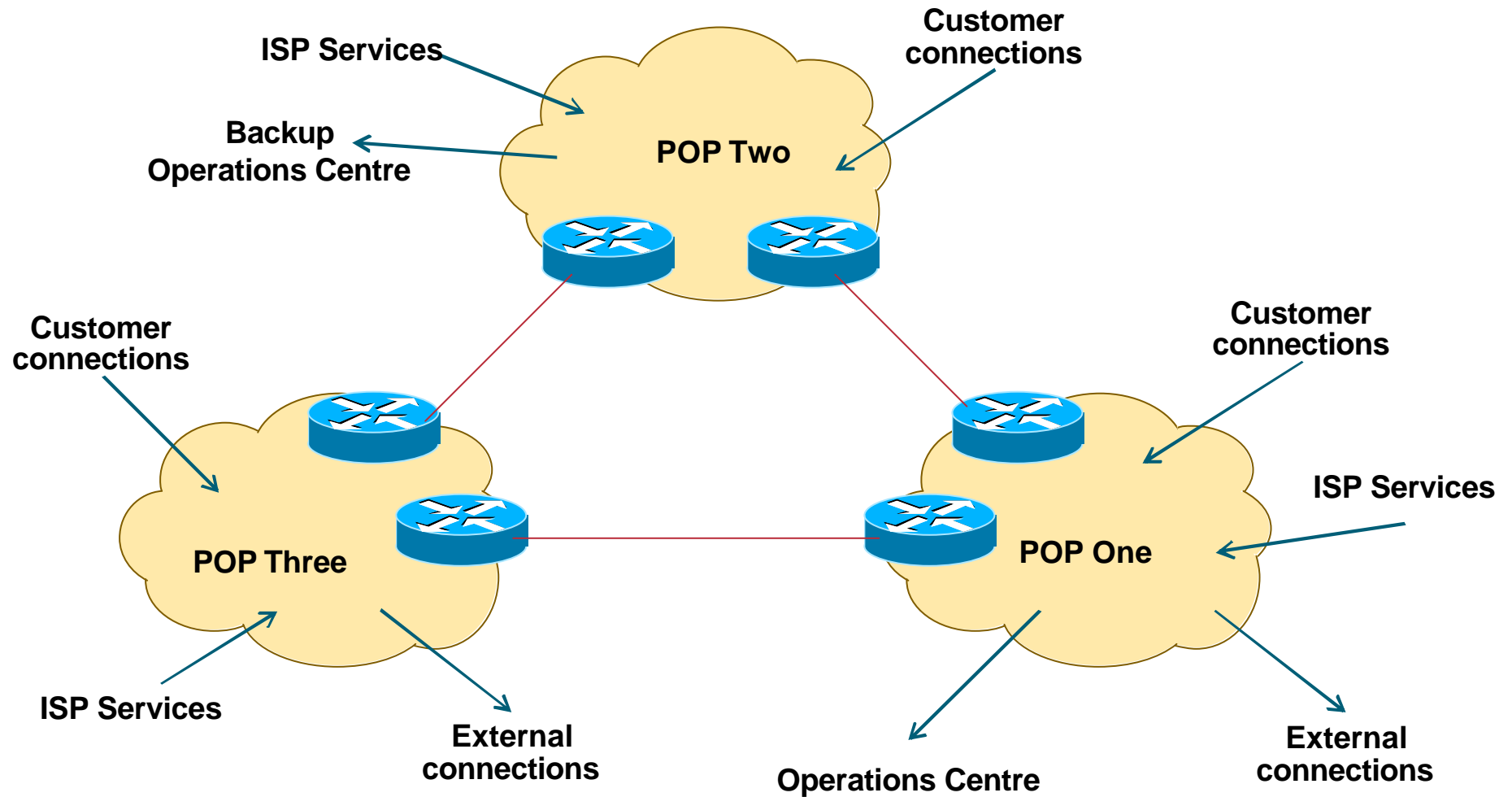
Backbone Network Design

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

ISP Services

Customer connections

Backup Operations Centre

POP Two

Customer connections

Customer connections

ISP Services

POP Three

POP One

ISP Services

External connections

Operations Centre

External connections

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

POP Two

Long distance link

POP Three

POP One

Alternative/Backup Path
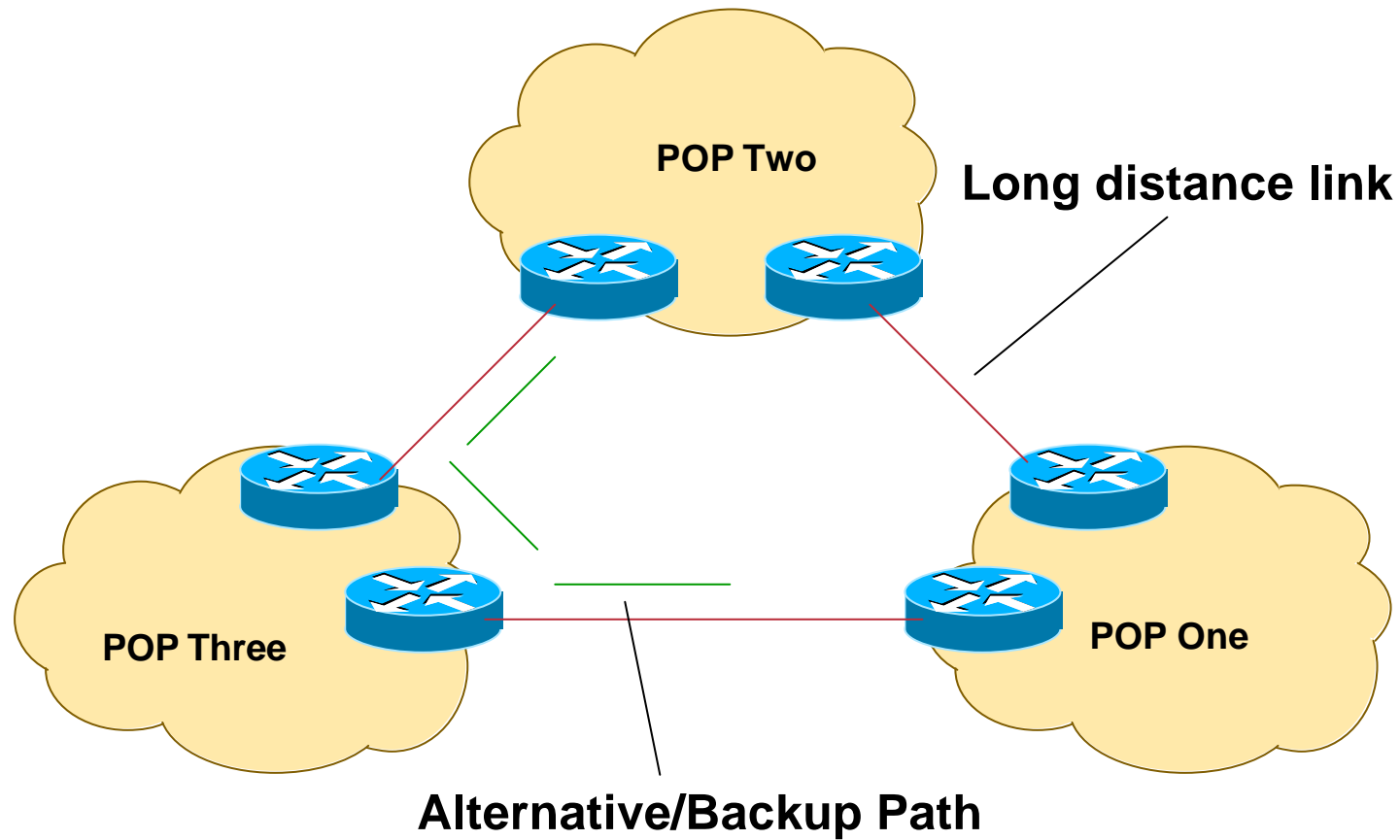
Metropolitan Area Backbone Links

- **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

POP Two

Metropolitan Links

POP Three

POP One

Metropolitan Links
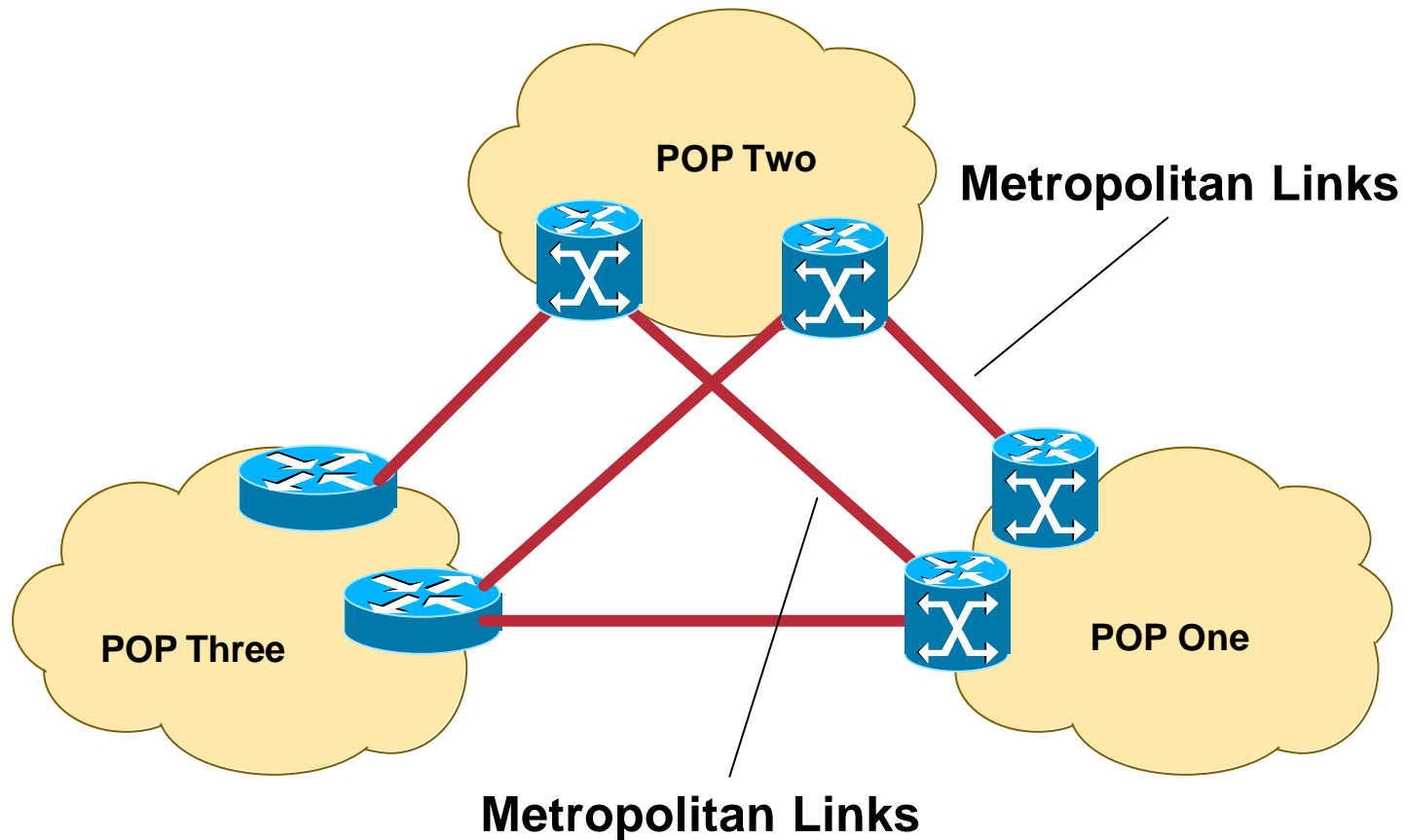
Traditional Point to Point Links

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 (128Mbytes), 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 (128Mbytes), 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

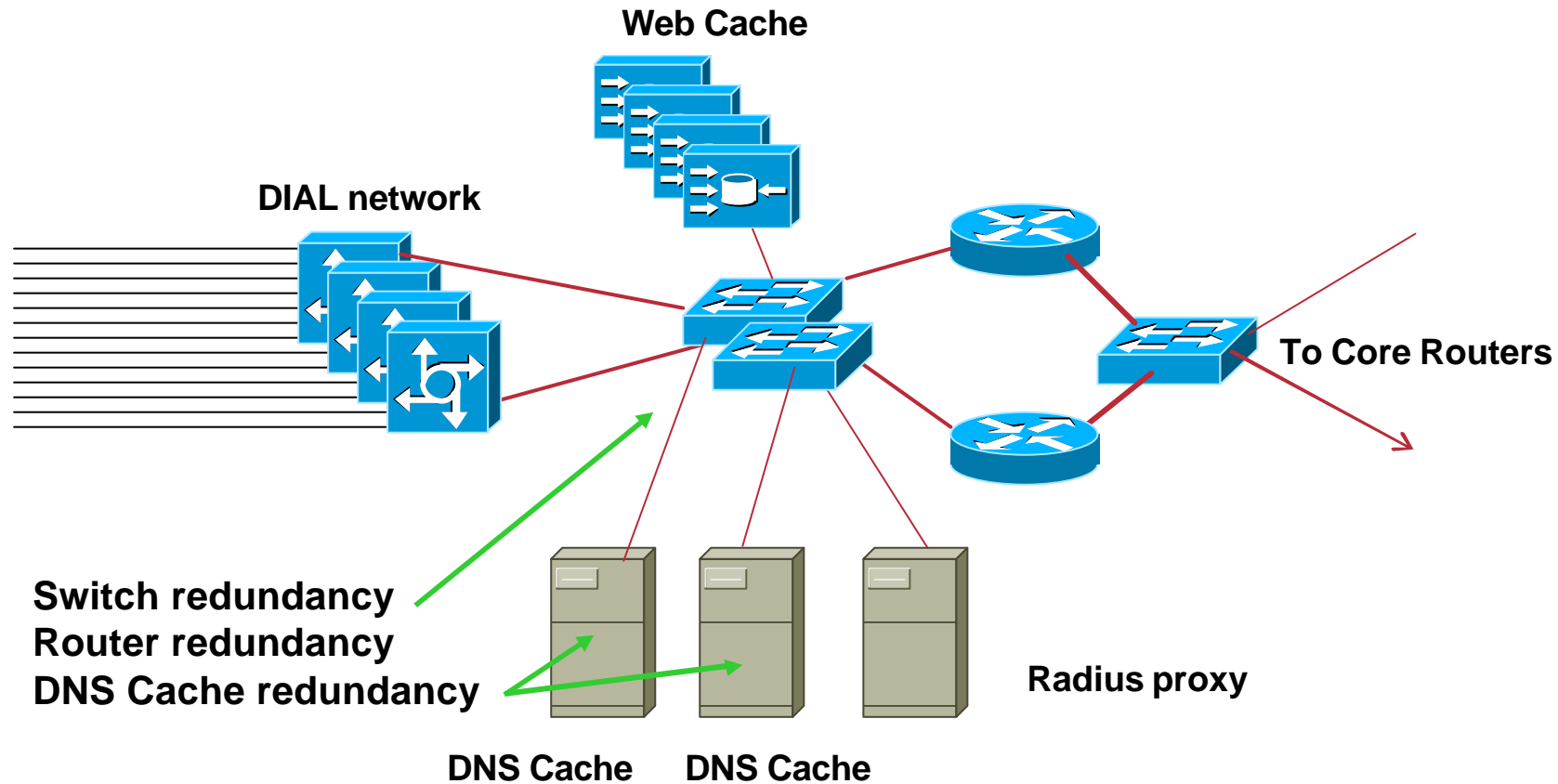   **Your customers use this as resolver, NOT your primary or secondary DNS**

    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 (256Mbytes+ depending on number of zones)
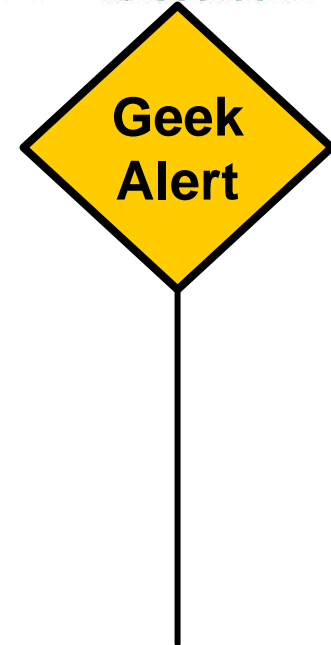
ISP Services
Caching Nameserver

Web Cache

DIAL network

To Core Routers

Switch redundancy
Router redundancy
DNS Cache redundancy

Radius proxy

DNS Cache   **DNS Cache**

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

ISP Services
Caching Nameserver

**Geek
Alert**

- **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!**

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

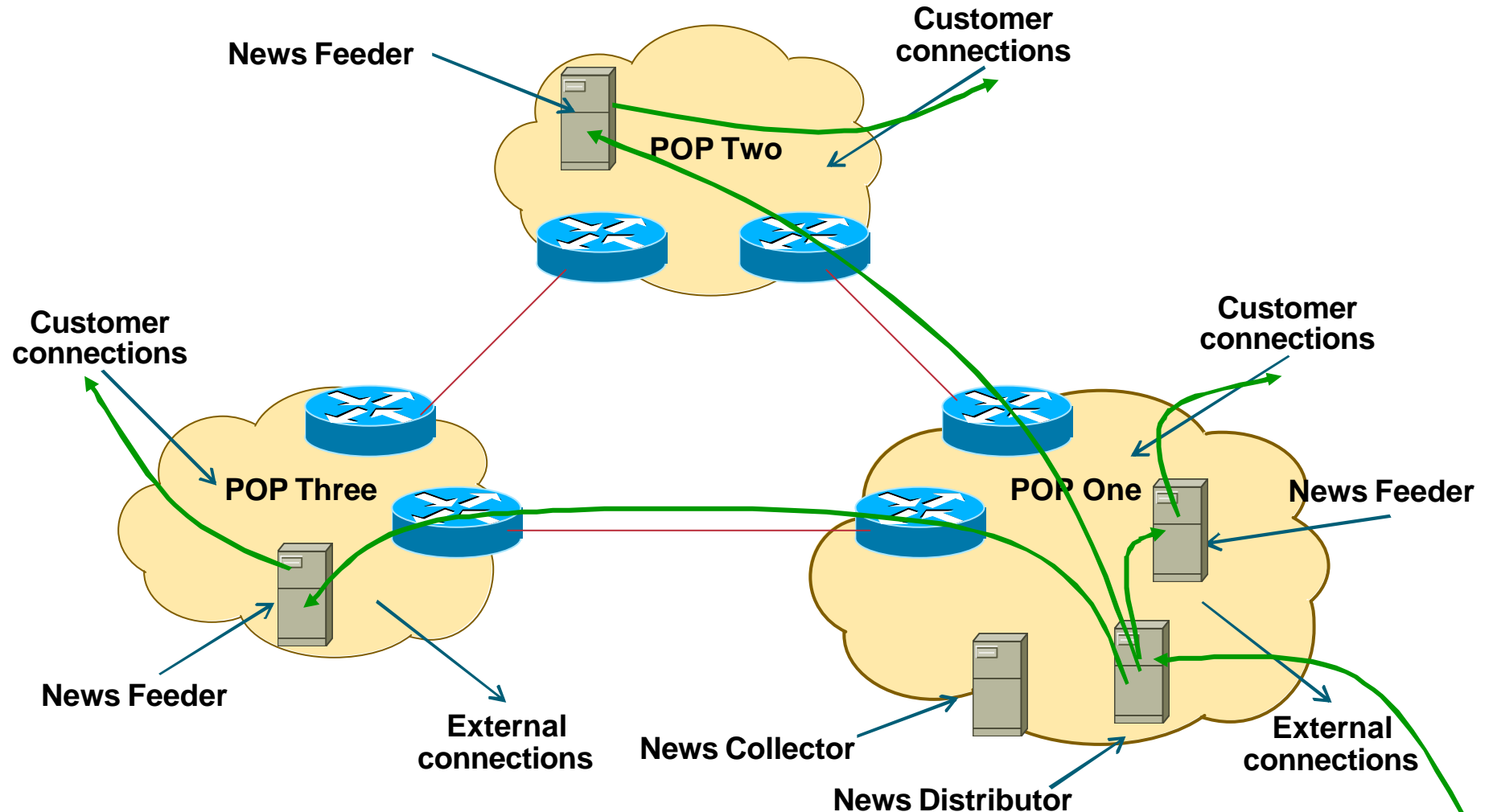 the vendor's distribution are rarely current

  ### 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
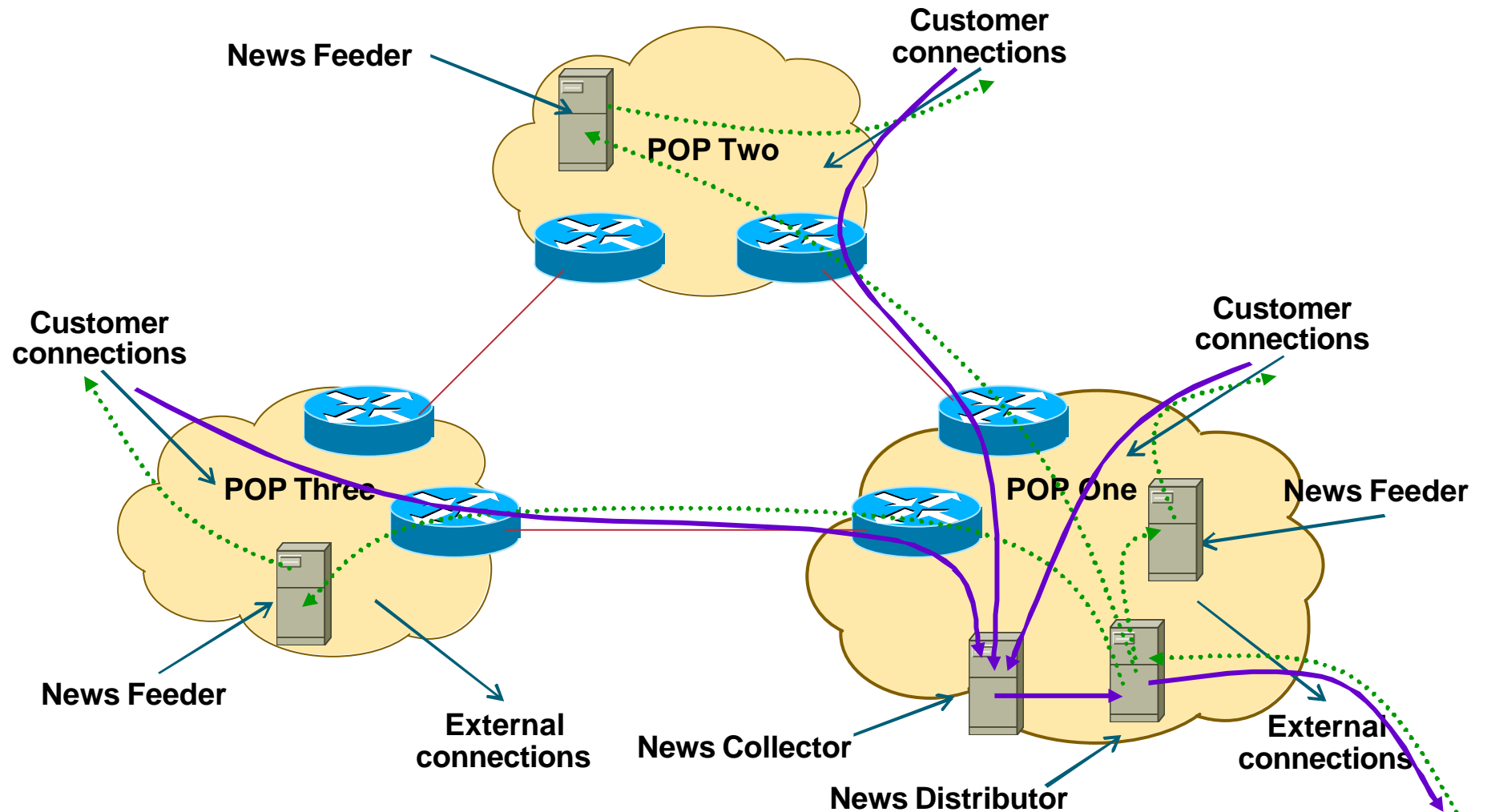
ISP Services
News

- **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

News Feeder

Customer connections

POP Two

Customer connections

Customer connections

POP Three

POP One

News Feeder

News Feeder

External connections

News Collector

News Distributor

External connections

ISP Services
News System Placement

Customer connections

News Feeder

POP Two

POP Three

Customer connections

Customer connections

News Feeder

POP One

News Feeder

News Feeder

External connections

News Collector

News Distributor

External connections

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
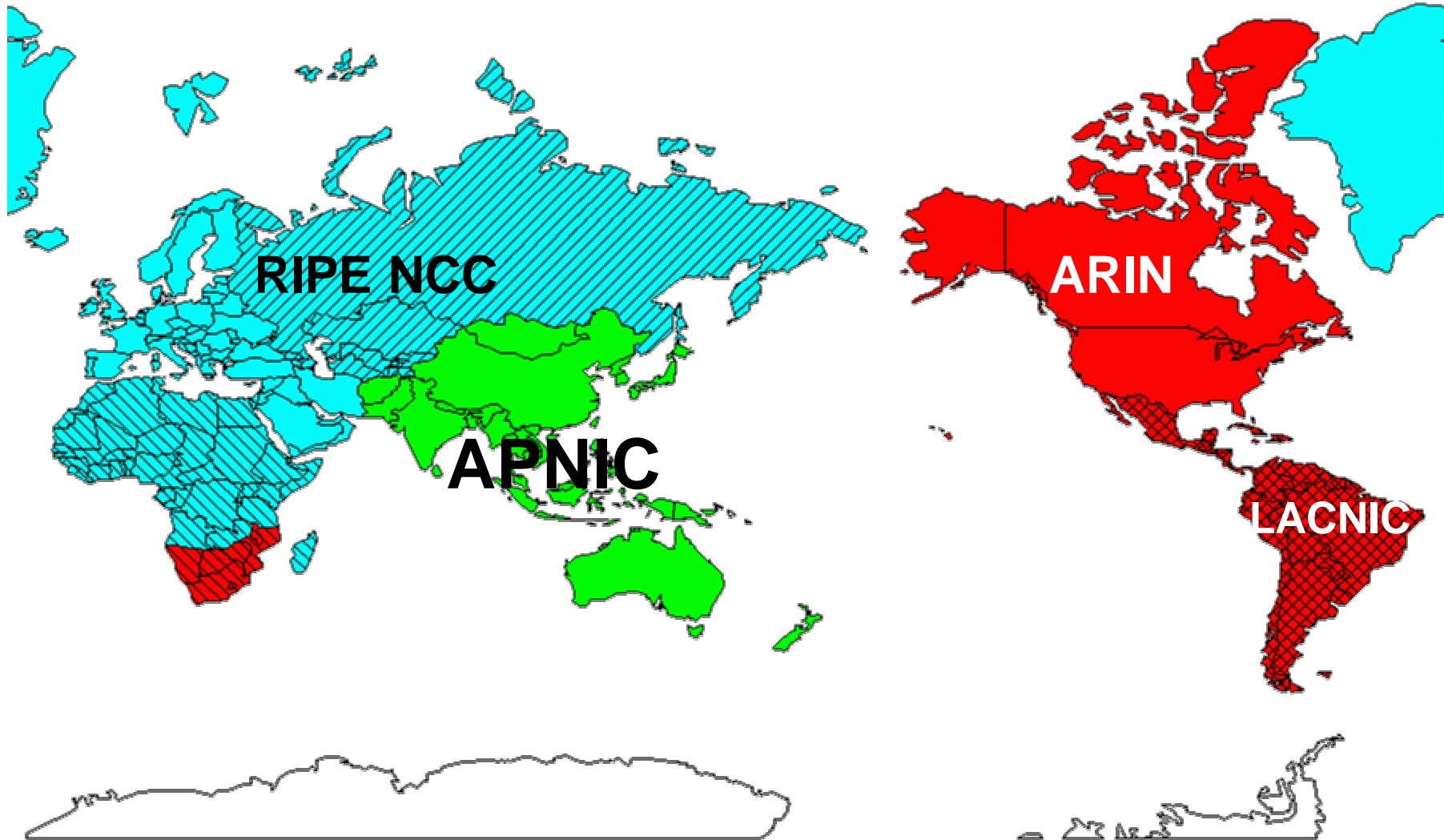
Addressing

Where to get IP addresses and AS numbers

- **Your upstream ISP**

- **Asia and Pacific regions**

    APNIC – http://www.apnic.net

- **North America and Southern Africa**

    ARIN – http://www.arin.net

- **Latin America and the Caribbean**

    LACNIC – http://www.lacnic.net

- **Europe, Middle East, Northern Africa**

    RIPE NCC – http://www.ripe.net

Internet Registry Regions

RIPE NCC

ARIN

APNIC

LACNIC

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
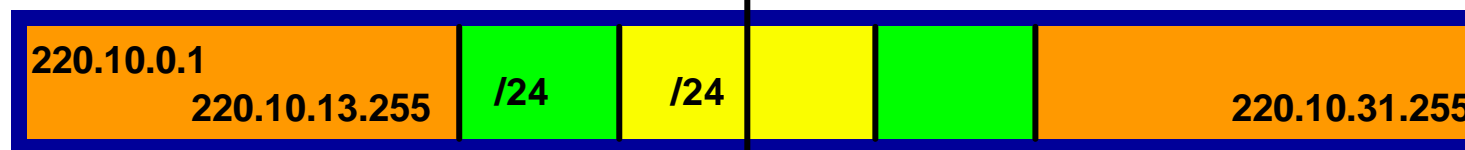
Addressing Plans – ISP Infrastructure

Phase One
220.10.0.0/20

220.10.0.1	220.10.14.255	/24

Customer assignments   **Instrastructure Loopbacks**

Phase Two
220.10.0.0/19

220.10.0.1 220.10.13.255	/24	/24			220.10.31.255

Original assignments   **New Assignments**

Addressing Plans
Planning

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

    Minimum allocation is /20

    Very likely that subsequent allocation will make this up to a /19

    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

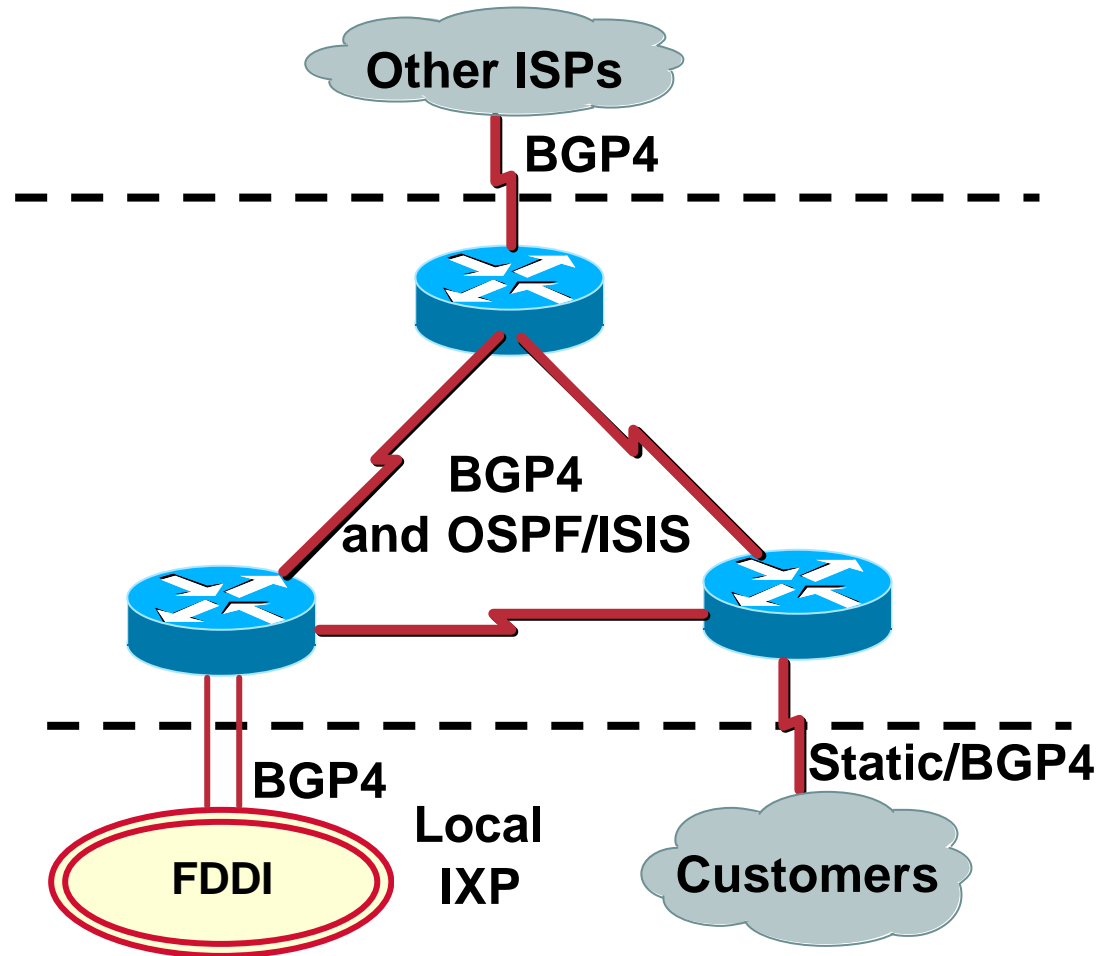 **ISPs aim to keep the IGP small for efficiency and scalability**

- **Exterior**

  Carries customer prefixes

  Carries Internet prefixes

  EGPs are independent of ISP network topology

Hierarchy of Routing Protocols

Other ISPs

BGP4

**BGP4
and OSPF/ISIS**

BGP4

**Local
IXP**

FDDI

Static/BGP4

Customers

Routing Protocols

- **IGP recommendations**

    keep the IGP routing table as small as possible

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

    use inter-router authentication

    use summarisation if possible

Routing Protocols

- **BGP recommendations**

    BGP should carry everything which isn't in the IGP

    - Internet routing table

    - Customer assigned addresses

    - DIAL network pools, passive LANs, etc

    use neighbour authentication

    use peer-groups and route-reflector hierarchy

    use route flap damping at the 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

ISP Infrastructure Security

- **router security**

  - **usernames, passwords, vty filters, TACACS+**

  - **vty 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

    protect *all* servers using routers with strong filters applied

- ## Hosted services security

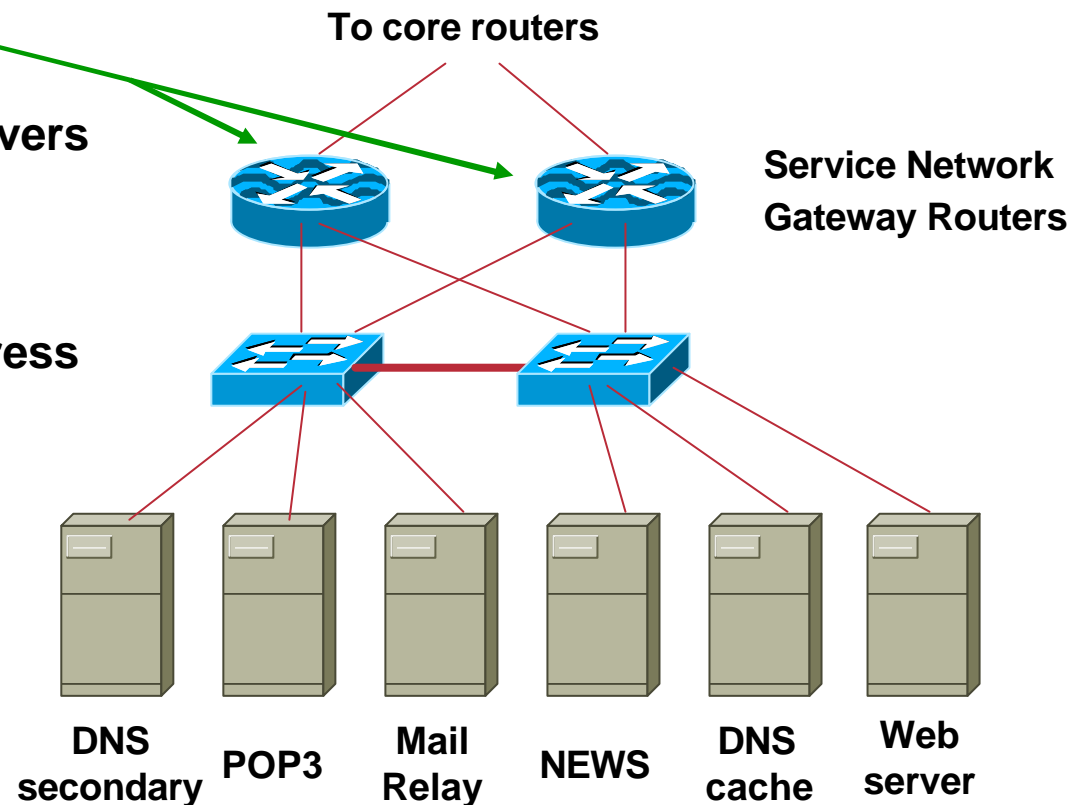   protect network from hosted servers using routers with strong filters

    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**

To core routers

Service Network
Gateway Routers

| DNS secondary | POP3 | Mail Relay | NEWS | DNS cache | Web server |

ISP Infrastructure Security
Hosted Server Protection

Access-list examples:

Inbound

Allow tcp/established to all servers
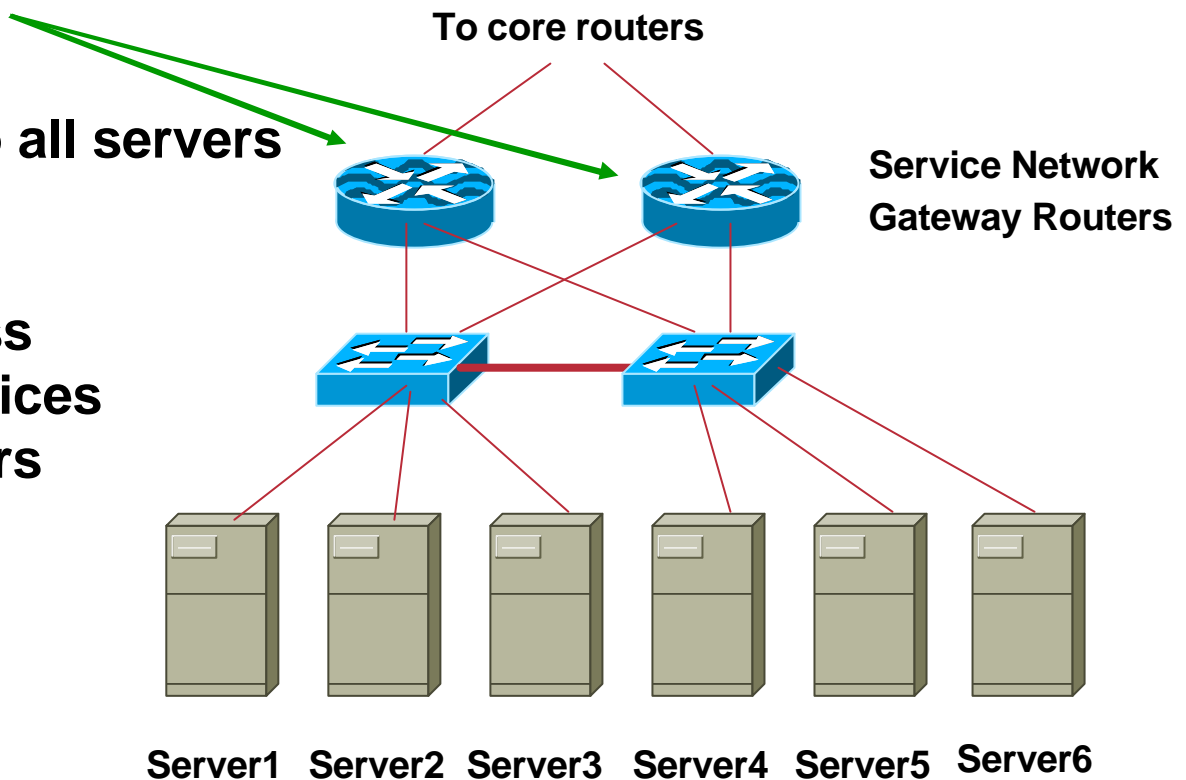**ICMP**
Web server: tcp/80
SSH for customer access
Any other ports for services
**   sold to customers**

Outbound
ICMP
Allow DNS udp/53 and
**   tcp/53**
Block all access to ISP
**   address range**

To core routers

Service Network
Gateway Routers

Server1 Server2 Server3 Server4 Server5 Server6

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)**

ISP Network Security

- **Denial of Service Attacks**

    eg: "smurfing"

    see http://www.denialinfo.com

- **Effective filtering**

    network borders – see IOS 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 then the address you have allocated to them!
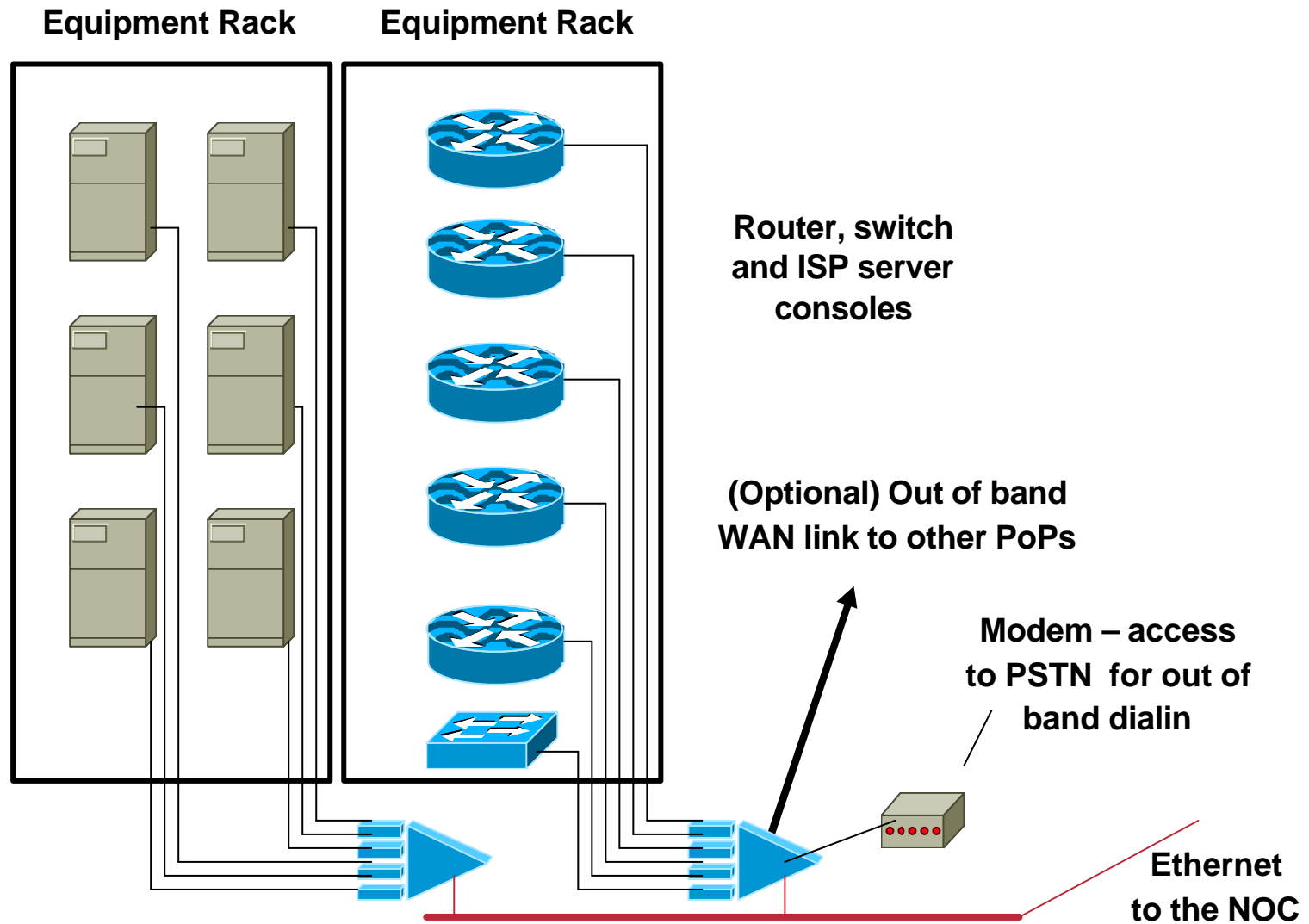
Out of Band Management

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

Equipment Rack   **Equipment Rack**

**Router, switch
and ISP server
consoles**

**(Optional) Out of band
WAN link to other PoPs**

**Modem – access
to PSTN for out of
band dialin**

**Ethernet
to the NOC**

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!

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