



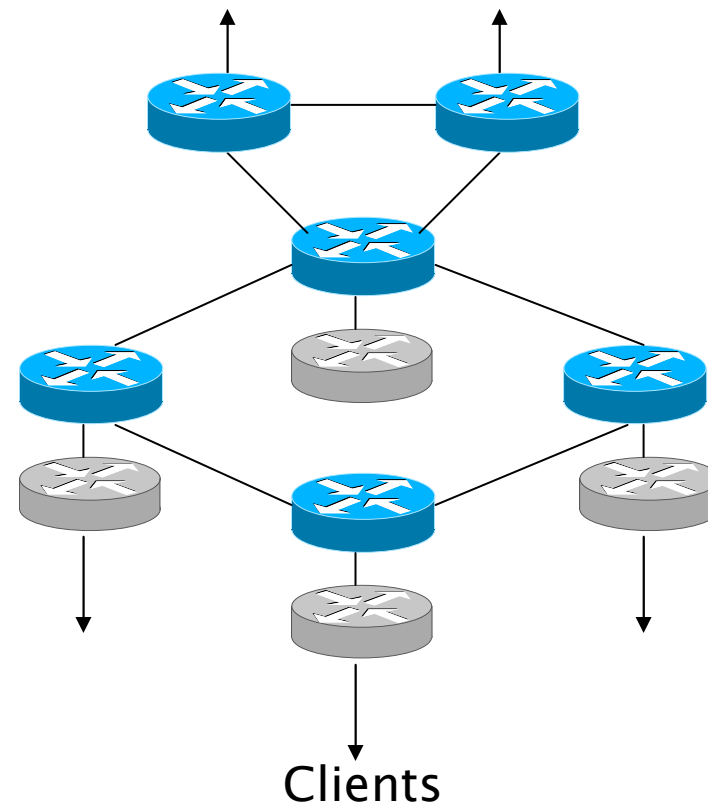
HEAnet's New Network and Working with IOS-XR

Brian Nisbet & Dave Wilson
HEAnet

Network Issues in 2007

- > 1 Gbit/s requires new hardware
- Software upgrades long and intensive
- Feature upgrades require new hardware...

External peers and upstreams



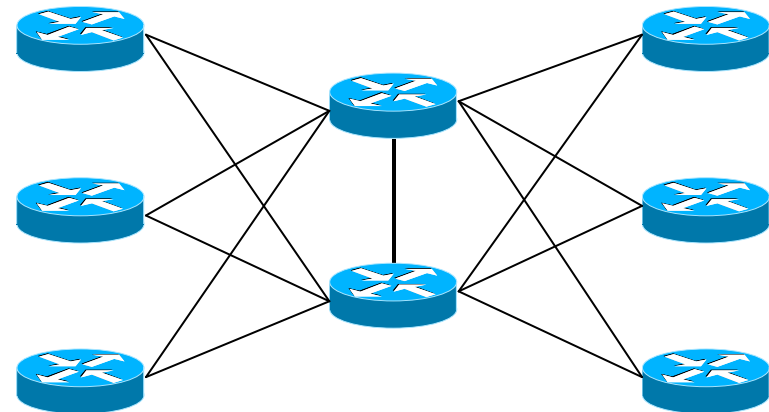
Wrong Kind of Duplication

- The routed network (Rednet) was originally designed to solve the problem of geography.
- New Layer2 network (Bluenet) solved that problem too.
- Port density on the underlying network was cheaper.
- Move towards ubiquitous ethernet.

Collapsed backbone

Now

- New Best Current Practice: Concentrate, then duplicate
 - Concentrating reduces the hardware, and with it failure incidence and scaling cost
 - Duplicating provides resilience, and keeps network consistent
- Model now being followed by many NRENs



New Network Principles

- Bluenet provides backhaul and aggregation.
- Rednet provides resilience and routing.
- Reduce non-ethernet, try to replace completely.
 - Serial and ATM still lingering.

Network Architecture

- Two routers, both performing core and access functions.
 - One each in our two Dublin network PoPs.
- Where possible, client gets connectivity via Layer2 network to *both* routers.
 - Primary/backup connections.
 - Resilience a function of the underlying network.
- Ethernet only, burst up to 10Gbit/s per interface.



BGP Everywhere

- We started going this direction about 2-3 years ago; this cements it.
- Speak BGP to every external connection
- IGP should contain just enough information to bring up BGP:
 - loopback ints and the links to reach them
- All end-user prefixes in BGP

IGP Choices

- We're currently running OSPF and ISIS.
 - only because old network still connected.
- Plan to pick one, and put v4+v6 together.
 - Pro: less to learn, more predictable failure modes.
 - Con: failures more likely to affect both.

Hardware

- Chosen equipment: Cisco CRS-1
- Scalability
 - up to 100Gbit/s interfaces
 - 1.5+ Terabits per second total
- Full support for current and future services
- New operating system platform: IOS-XR



The HFR

- The 8 slot is big. The 16 slot is very big indeed.
- Other than its size, nothing terribly remarkable about physical installation.
- Cabled all ports on day one back to a patch panel to make future cabling easier.





CRS-1 & IOS-XR

- Bigger? Faster? Better?
- All experiences on a Cisco CRS-1.
- Overall a large number of improvements.
- Definitely many steps in the right direction.
- Standard leap-frogging, now very similar in look and feel to JUNOS.



Major Design Differences

- v4 & v6 treated largely the same.
- Commit!
- Editable lists (vi or emacs).
- Route Policy Language (RPL).
- Config sections grouped in a nicer way.
 - Routing protocols especially!



Commitment Issues

- Nothing changes until you commit.
- This is likely to be one of the main causes of, “I made a change, but it isn’t working.”
- Commit confirmed gives you a chance to change your mind.
- Commit comment – documentation is good.
- Commit replace – **Extreme Caution!**

RPL

- No more route-maps.
- Proper if/elseif and Parameters.

```
route-policy schools-in($pset, $pref)
  if destination in $pset then
    if as-path in (ios-regex '^65002$') then
      set local-preference $pref
    else set local-preference 245
    endif
    set community (1213:2000)
  endif
end-policy
```



Configuration Changes

- Only one way to configure. (But *conf t* still works.)
- Everything in sections.
- Line/login details at the top.
- Much more flexibility in defining user rights.
- Access Lists and route policies before protocols.



- IGP Config all in the one place!

```
router ospf red
router-id 193.1.238.129
nsf cisco
address-family ipv4 unicast
area 0
dead-interval 6
hello-interval 2
interface Loopback0
!
interface Loopback9
!
interface GigabitEthernet0/12/0/2
network point-to-point
mtu-ignore enable
!
interface GigabitEthernet0/12/0/5
cost 20000
network point-to-point
```




- BGP config now per neighbour. All config for v4 in one place, all for v6 in another.

```
neighbor 62.40.125.101
  remote-as 20965
  password encrypted
  description GEANT 1GE Backup
  address-family ipv4 unicast
    send-community-ebgp
    route-policy geant2-in in
    route-policy geant2-out(aggregate-v4) out
    send-extended-community-ebgp
    remove-private-AS
    soft-reconfiguration inbound
  !
  address-family ipv4 multicast
    send-community-ebgp
    route-policy geant2-in in
    route-policy geant2-out(aggregate-v4) out
    send-extended-community-ebgp
    remove-private-AS
    soft-reconfiguration inbound
```



Little Changes

- ip now needs to be specified as ipv4 or ipv6.
- sh ip bgp sum -> sh bgp [ipv4|ipv6] [uni|mul] sum
- “sh ip bgp neighbor <addr> [route|adv]” -> sh bgp [ipv4|ipv6] [uni|mul] neighbor <addr>[route|advertised-routes]
- Routing table now updates after config changes, even without clearing session.
- No policy = no routes exchanged (will get a warning).

The Code

- It's clearly a very new code-base.
- Current latest versions: 3.5.4, 3.6.1, 3.7.
- Large number of files, one for the base and then one each for additional processes/features.
- We've been doing some fairly intensive debugging for a few issues.
- 3.7 doesn't fit on the installed 1GB flash card. Requires RP removal to replace.
- A number of issues ending up with the developers.

So, You Really Want to Upgrade?

- IOS-XR is not designed to be upgraded like IOS.
- Upgrade from 3.5.2 -> 3.6 took 2 ½ hours.
- Multi-stage process with several points at which processes are interrupted.
- Cisco do make it clear that upgrades are a very intricate process, but reality outstripped the warnings.



Avoiding Reloads?

- Software Maintenance Upgrades (SMUs) are intended to patch the software without a full upgrade.
- Some, such as any that modify the hfr-base code, still require a reload. 30 minutes per RP.
- Others cause a process restart, others again are hitless and fine in service.
- We've been told to avoid reloading if at all possible, but it hasn't been as possible as we'd like.



Other Niggles

- Not just a case of “different to IOS.”
- Login is much slower, box less responsive.
- Reload messages are sometimes misleading and/or spurious.
- Multicast RPF checks now rely purely on multicast routing table (this may be changing).
- Path MTU discovery not working as expected.
- Vast mounts of SNMP information (CPU on every line card), but MIBs initially very hard to obtain.
- Names working far better for SNMP than ifindexes.

HEAnet



IRELAND'S NATIONAL EDUCATION & RESEARCH NETWORK

Questions?