Right-Sizing the Right Solution

Building a Flexible Core Network for Future ICT Challenges

Michael Harlow,
Network Engineer,
University of Tasmania

Agenda

- Some UTAS Background
- Issues Faced
- Network WAN Solution
- HP Choice and Why
- Implementation
- Platform for the Future

UTAS Background

- 4th oldest University in Australia, and the only university in Tasmania.
- One of the top 10 Australian research universities
- Full spectrum of disciplines

- Staff of ~2500 EFT (Equiv Full-Time)
- Students of ~18000 EFTSU (Equiv Full-Time)
- 3 Primary Tasmanian teaching/research campuses
  - S - Sandy Bay
  - N – Newnham
  - NW - Burnie.
- 15 Smaller teaching campuses
- 15 Smaller Research related sites
- ~20 Health related network presence (field work, residency)

UTAS Collaboration

UTAS Statement: "As the State's only university, UTAS […] works with local and national providers to improve health, education and public policy…

Australia's primary industries are a mainstay of its economy and its culture. UTAS researchers develop and promote agribusiness, agriculture, aquaculture and fisheries, forestry and mining.

The hub for the International Antarctic Institute, UTAS is also the gateway to the Antarctic. It provides the infrastructure necessary for research and collaboration on the great Southern continent, as well as providing much of the ideological impetus.

UTAS is a leader in oceanic policy and maritime security issues...research that touches every nation with a coastline."

My answer: If it is under in or on the water; in under or on the ground; low high or outside the atmosphere; in or on the body or mind; in the past, present or the future; UTAS will have an interest.

UTAS Collaboration

Partnerships with:

- Australian Antarctic Division (AAD)
- CSIRO:
  - Forestry,
  - Marine Labs

Tasmanian State Government departments of:

- Education
- Primary Industries, Parks, Water and Environment
- Health and Human Services

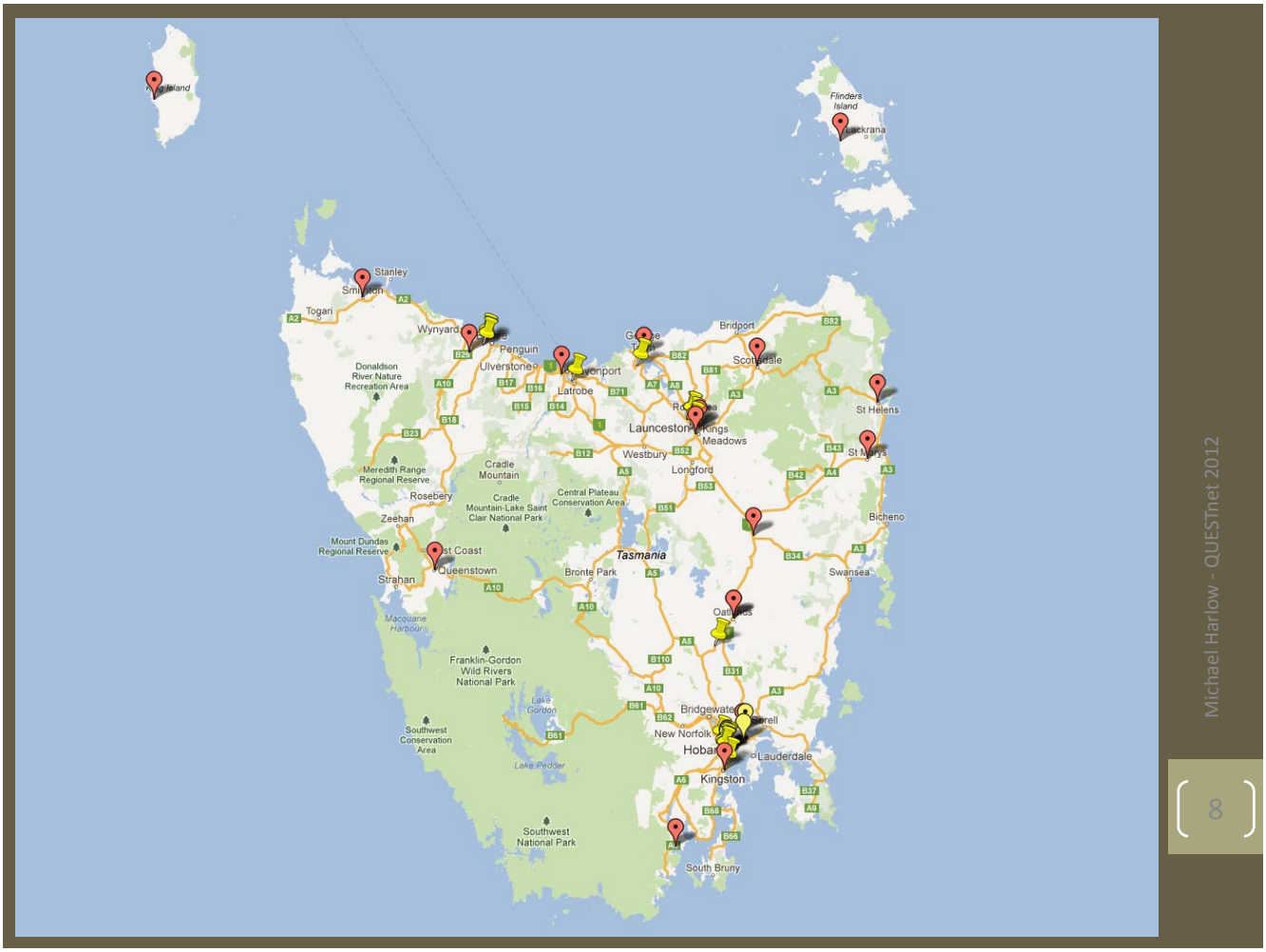
This results in many cases of joint occupancy, shared staff and resources, including ICT resources.

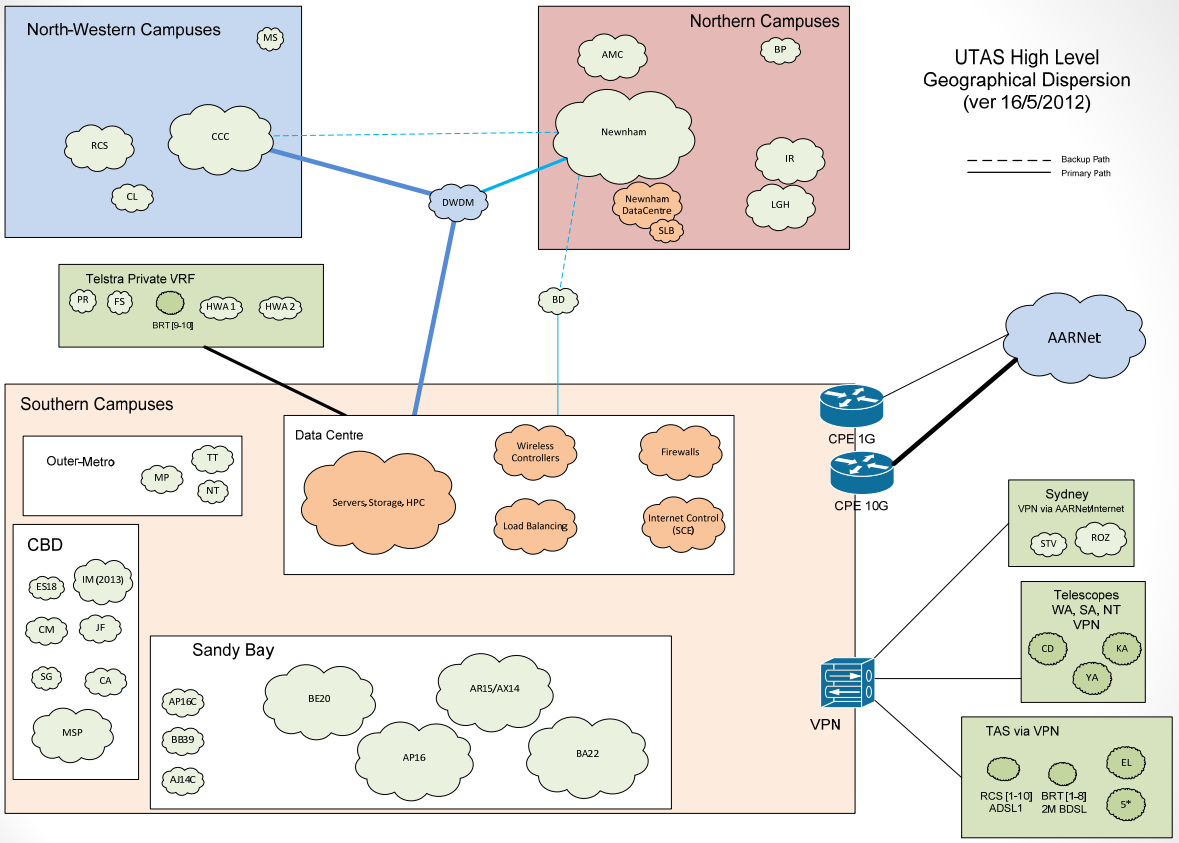
Hospitals, medical centres, educational institutes, research laboratories.

UTAS Collaborative, Specialist and Theme Areas (partial)

- Institute for Marine and Antarctic Studies (IMAS)
  - Integrated Marine Observing System (IMOS)
  - Tasmanian Aquaculture and Fisheries Institute (TAFI)
  - Institute of Antarctic and Southern Ocean Studies (IASOS)
  - Centre for Marine Science
  - Antarctic Climate and Ecosystems CRC (ACE CRC)
- Tasmanian Institute of Agriculture (TIA)
- CRC for Forestry
- Australian Innovation Research Centre (AIRC)
- Australian Centre for Research on Separation Science (ACROSS)
- Radio Astronomy (inc eVLBI and AuScope)
- Centre of Excellence in Ore Deposits (CODES)
- Menzies Research Institute Tasmania (MRIT)
- Australian Maritime College (AMC)
- TPAC – Tasmanian Partnership for Advance Computing (HPC)
- RDSI minor Node
- NeCTAR?


Timor Sea
Darwin
Northern Territory
Coral Sea
Queensland
Australia
Western Australia
Brisbane
South Australia
Perth
New South Wales
Adelaide
Canberra
Australian Capital Territory
Great Australian Bight
Victoria
Melbourne

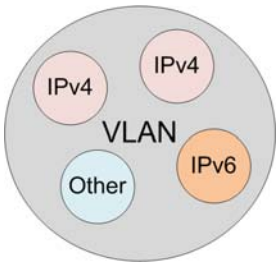

Flinders Island
Stanley
Smithton
Togari
Wynyard
Penguin
Ulverstone
Devonport
Latrobe
Donaldson River Nature Recreation Area
Bridport
Scottsdale
St Helens
Launceston
Kings Meadows
Westbury
Longford
Meredith Range Regional Reserve
Cradle Mountain
Cradle Mountain-Lake Saint Clair National Park
Central Plateau Conservation Area
Rosebery
Zeehan
Bicheno
Mount Dundas Regional Reserve
Queenstown
Tasmania
Bronte Park
Swansea
Strahan
Macquarie Harbour
Franklin-Gordon Wild Rivers National Park
Lake Gordon
Southwest Conservation Area
Bridgewater
New Norfolk
Hobart
Lauderdale
Kingston
Lake Pedder
Southwest National Park
South Bruny



Starting Point

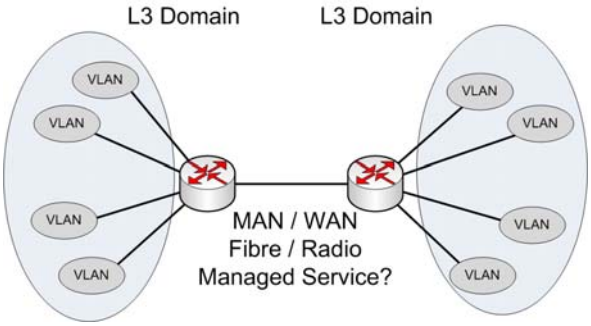
- Networking:
  - 4 permanent staff
  - Cisco networking, for at least the last 18+ years.
  - 500 Wireless AP and 900 radios (mostly dual-band)
  - 800 Cisco switches and 10 routers. Mostly 24-port 1RU switches.
  - Dual Catalyst 6509E in Data Centre with Sup720-3B
  - 50% of switches are greater then 6 y.o. (2900XL, 2950T)
  - Cisco ACE, FWSM, WiSM, VPN, SCE (Service Control Engine)
  - More wall sockets than switch ports by about 2:1
  - Most WAN/MAN links are either dark fibre (1G) or private carrier grade radio (34, 100, 150, 400M).
  - Some WAN links are services at either L2 or L3 from service providers.

Basic Networking 101



VLANS contain one or more protocols:

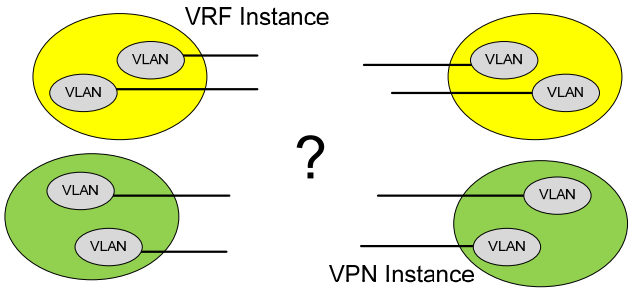
- IP V4
- IP V6
- Multicast
- IPX
- Netbios
- Unknown



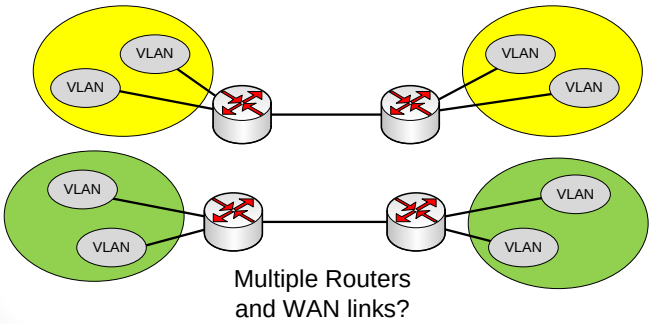
VLANS communicate with each other by routers, if the protocol is “routable”.

All VLAN have access to each other. Security (if any) is via Access Control Lists (ACL) on routers

What if we want two separate isolated networks sharing the same physical infrastructure?

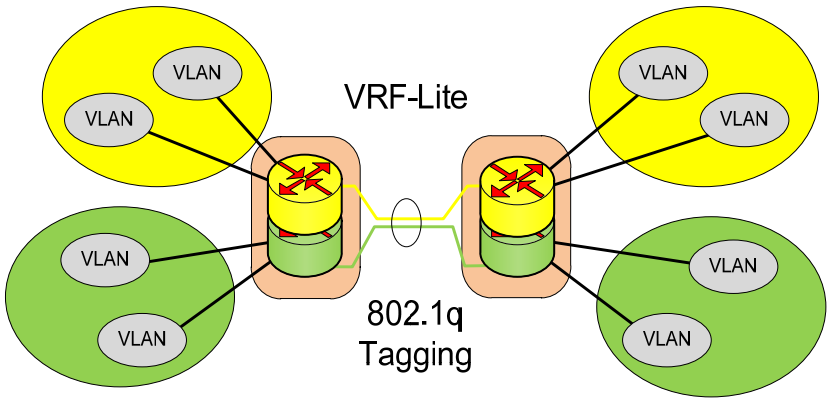


How to segregate the network, including across WAN links?



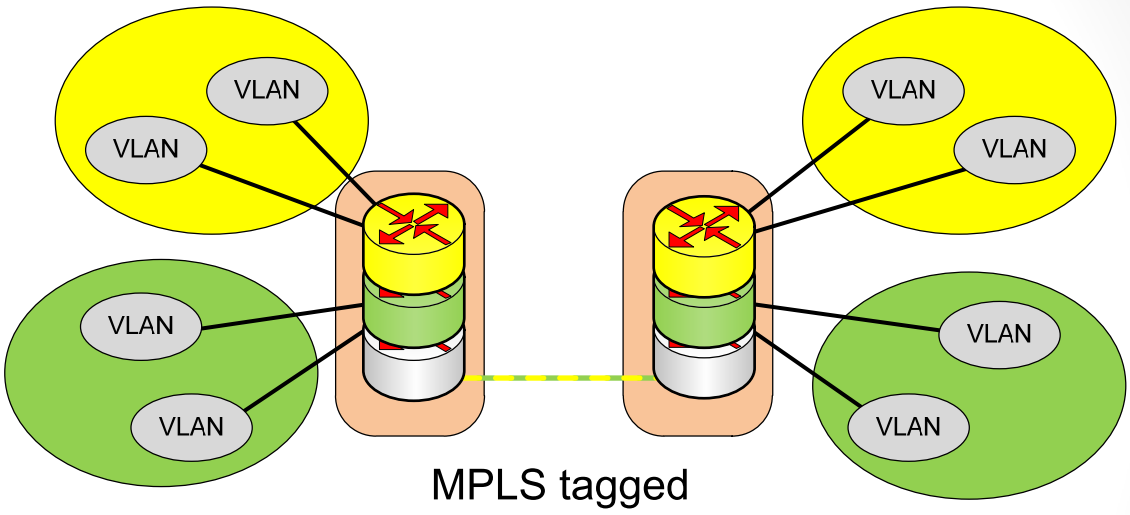
Really have two networks, routers and WAN links are duplicated!

Isolated Networks – The Lite Way



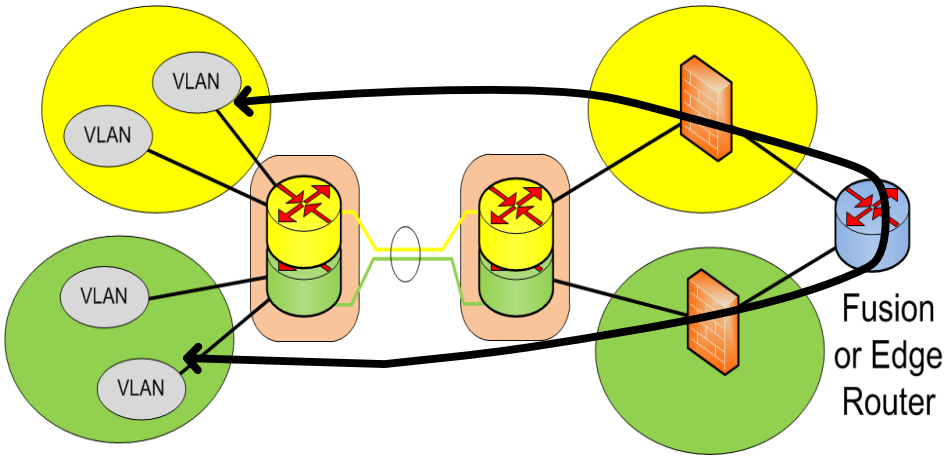
Does the WAN link support 802.1q tagged packets?
How scalable? 15 VRF times 30 sites = 450 vlans
Management overhead!

The MPLS (Multi-Protocol Label Switching) Way



- Not so fast….. MPLS needs extra protocols, like LDP, LSP, MP-BGP, ISIS or OSPF etc. Not as simple as 'tick a box' and hit OK.
- But with this complexity comes functionality. Transmit any labelled packet across the MPLS link, the MPLS network does not need to understand the inner packet.

Isolated Networks – The FW Effect



- Inter-VRF traffic can involve multiple firewalls, and not necessarily the shortest path.
- Has implications for sharing of services (DNS, DHCP, Fileservers, App servers)
- NAT should occur on exit of firewalls when RFC 1812 used. Fusion is public/Internet IP

MPLS Capabilities

Multiprotocol Label Switching (MPLS) is a method for engineering traffic patterns by assigning short labels to network packets that describe how to forward them through the network. MPLS is independent of routing tables or any routing protocol and can be used for unicast packets.

The MPLS framework supports traffic engineering and the creation of virtual private networks (VPNs). Traffic is engineered (controlled) primarily by the use of signalling protocols to establish label-switched paths (LSPs). VPN support includes Layer 2 and Layer 3 VPNs and Layer 2 circuits.

- L3 & L2 VPN (Martini, Kompella)
- VPLS Virtual Private LAN Service (L2) Point to Multi-point
- Pseudo Wire (point to point)
- LDP – Label Distribution Protocol
- LSP – Label Switched Path
- LSR – Label Switching Router
- MPLS-TE – Traffic Engineering
- MP-BGP – Multi-Protocol extensions for BGP
- OSPF or ISIS – Underlying link-state routing protocol
- Transmit any protocol across the network. Inc multicast, IPv6.
- Reuse of address space (RFC)
- ALL are standards, multi-vendor.

Issues Faced

- Request for PCI-DSS environment
- Co-location of Student Residences/bedrooms on UTAS campuses
- No separation of Academic, Student, Research and Corporate networks
- WAN links are L3, limiting ability to provide layer-2 private services
- Shared facilities, request to transit other parties across internal network
- High volume Internet data transfers for HPC and Telescope eVLBI
- Perceived need for 10G capacity internally.
- Increased Data Centre port counts, 10G capacity
- Equity of network functionality at all locations (Multicast, PXE, WoL, IPv6)
- Hardware fresh of oldest equipment (2900XL, 2950)
- IPv6 deployment

Shopping List

Decided an MPLS network would provide the foundation for a flexible core WAN network to support the known, and yet to be known requirements of the University. Support of tunnelling would allow network to be stretched to as many remote campuses as possible, to provide a uniform functionality and egalitarian experience to all.

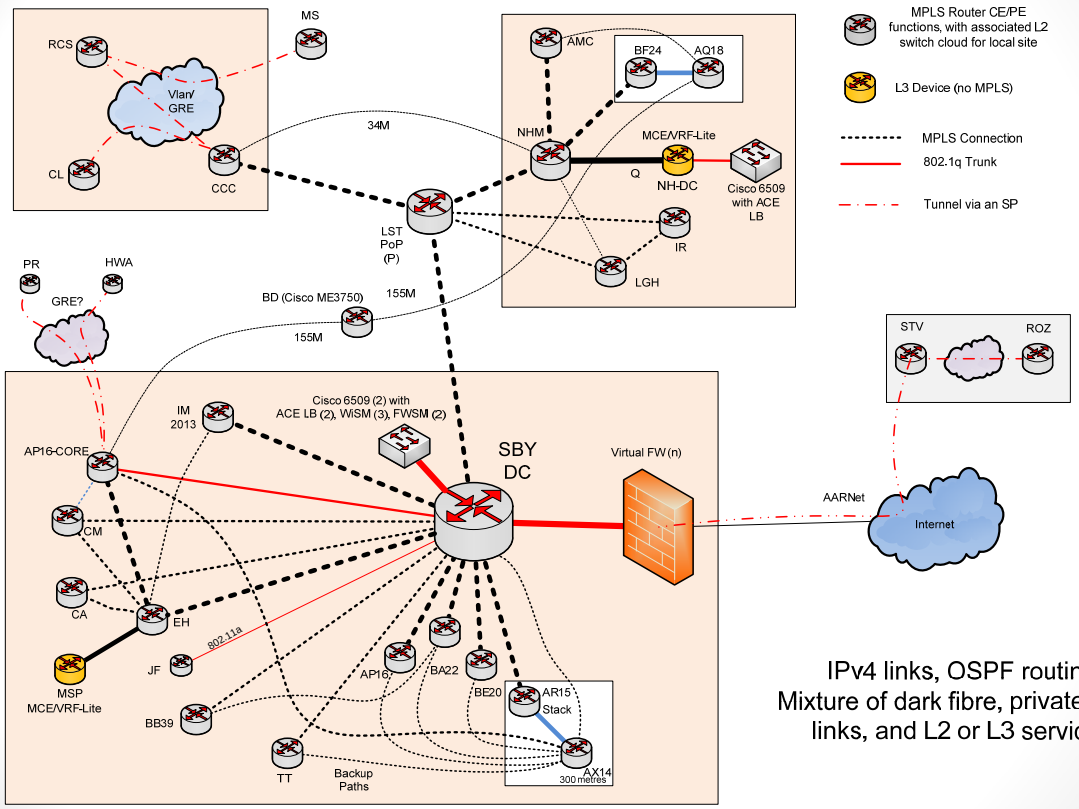
So, start to consider ourselves to be an SP (Service Provider), and each functional business unit as discrete customers, with their own security needs, firewalls, and SLA level.

MPLS usually means moving into SP equipment. And price range!

- MPLS
- VPLS
- Tunnelling (GRE at least)
- IPv4 and IPv6
- Switch Stacking
- V4/V6 Multicast
- 10G interfaces

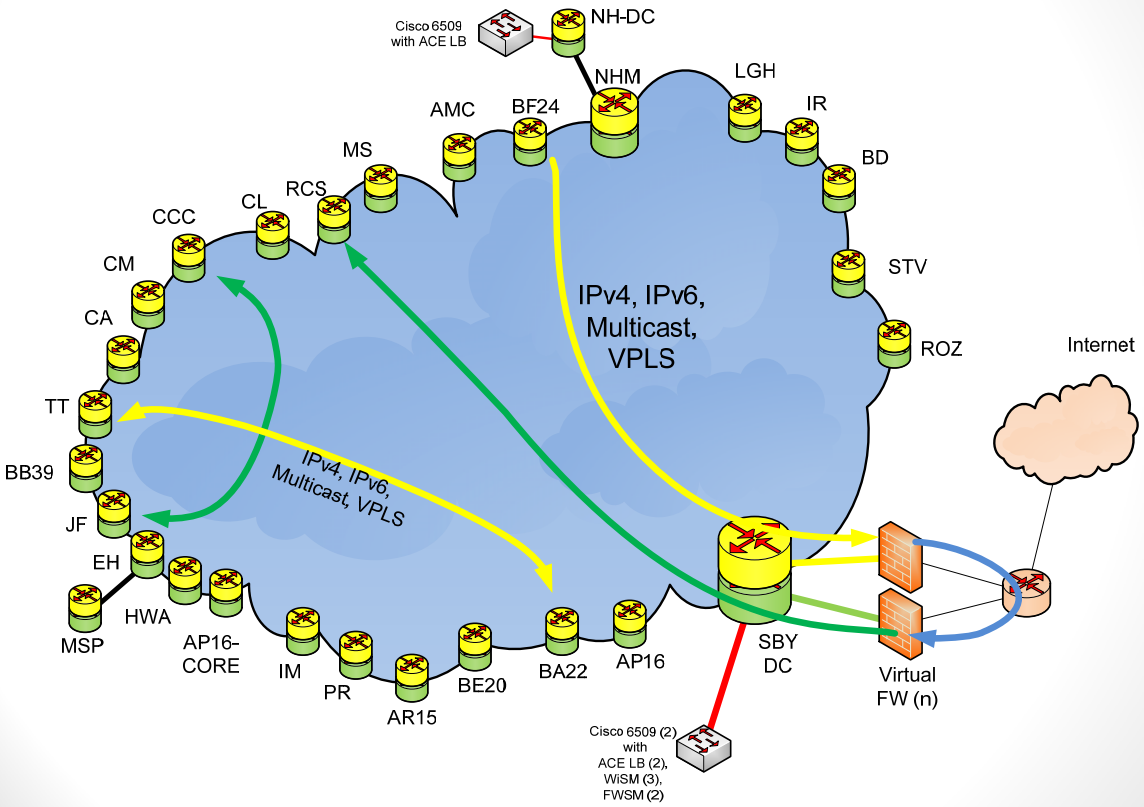
- Single product family
  - Minimise software versions
  - Ease of "sparing"
- Energy efficiency
- Small form factor (1 or 2 RU)
- Cost effective annual maintenance

The Hidden Underlying MPLS Network



IPv4 links, OSPF routing
Mixture of dark fibre, private radio links, and L2 or L3 services

User or VPN/VRF view of the Network



What would our MPLS VPN/VRF be?

Break existing big single network into separate VPN/VRF instances behind individual firewalls:

- PCI DSS
- Corporate Functions (HR, Admin, VC, Exec, ITR)
- Security (Cameras, door control)
- BMS (Building Management, HVAC)
- **Residents (Bedrooms)**
- Network Management (switches/WAP)
- **Student Wireless/Wired (802.1x) access**
- **Academic (Most remaining areas)**
- **Partner organisations**

Future

- Sensitive Research
- Anything…

First Steps

- Planning days with Cisco and HP/Frontline
- Discussion of our objectives, and possible products.
- Two rough designs, asked many questions, received most answers.

- Obtained HP products for PoC (Proof of Concept)
- Under CAUDIT/HP arrangement, limited to three resellers/integration specialists. Chose “Frontline” to partner with for pre-sales and sales support. Nationwide (except TAS).

- At all times, Frontline and HP have been very responsive and eager to support our endeavours. Access to specialists and engineers has been forthcoming, and several onsite visits and planning days. Deliveries have been as advertised, and faster than we are used to.

HP A-Series Genesis

- 2003 – H3C Joint Venture - Huawei and 3Com
- 2006 – 3Com executes option to buy Huawei out.
- 2009/Aug – CAUDIT 8 person study tour includes a visit to H3C
- 2009/Nov – Announcement of HP purchase of 3Com
- 2010/Apr – HP purchase completed
- 2010/July-Sept – CAUDIT Technical Group evaluate products
  - 5 Universities. 107 Criteria. 3-day hands-on demonstration/testing.
- 2010/Oct – Technical Group report presented to CAUDIT
- 2010/Oct – HP makes offer to CAUDIT of pricing structure

HP Networking

- The 3Com acquisition products were branded as A-Series, and existing ProCurve relabelled as E (Essentials) and V (Value) Series. Tipping Point becomes the S-Series (Security)
- This unfortunately has since been change and it is now harder to separate which products are ProCurve genes, and which are H3C/3Com genes.
- The A-Series are targeted at Enterprise markets, and given their H3C genes, any comparison with the ProCurve range is in my opinion a disservice to them.
- As a result, past experiences, rumours, stories about the ProCurve range should be ignored (fingers in ears and chanting). Not relevant.
- Also, dealing with HP Networking division, is not dealing with the printers, or servers or desktops or laptop divisions, sales or support. Leave experiences at the door.
- I'm always willing to allow new organisations a chance to prove themselves. So far, so good.

HP A5800 1RU Layer 3 Switch

- A selection of 1RU switches, with 24 or 48 ports, includes 4 front SFP+ 10G interfaces, and a module slot for expansion or media flexibility. Some models are dual PS, others can have RPS. PoE+ Options.
- Modules: (one slot) 16x 1G SFP or 16x 1G TX or 4x 10G SFP+
- MPLS, VPLS, IPv6. Line-rate. 256Gbps/211Mpps
- Tunnelling: QinQ, 4 over 4, 4 over 6, 6 over 6, GRE, 6 over 4 (manual/6to4/ISATAP)
- Stackable with "IRF", using ANY available pair of 10G port
- Fully feature licensed by default – "Drive-away" pricing
- 2RU model that has a "services" slot (VPN, FW, WLC)
- Usual features of a Layer 2 and 3 device
- Maximums acceptable: frame 10Kb, MAC Address 10K, VLAN 4K, ARP 16K, SVI 1K, Routes IPv4 16K, IPV6 8K.
- QoS: 8kbps granularity, CAR, Eight output Q per port, WRED, SP, WDRR, WFQ, and SP+WDRR, shaping, marking, filtering. HQoS.
- No comparable Cisco product. Combination of 3750X, ME3600X and 6500 in 1RU!

HP A5120 (L2 Only)

- 24 and 48 port 1RU switches
- 2 Module slots, each slot can handle 2x 10G ports
- Modules include XFP, SFP+ and copper CX (Infiniband?)
- Full line rate
- PoE+ Option
- Stackable with IRF, if optioned with 10G modules
- CX module for very cost effective short range stacking or interconnecting.
- Maximum stack of 4 switches
- The usual L2 features, including 802.1x, security etc
- CapEx is low enough that we buy the CX for stacking without thought. Also CapEx is low enough that we are able to increase switch port density to reduce wall-socket to switch port ratio closer to 1:1, and achieve the holy grail of fully patched. Roll on dynamic vlan assignment and less visits to the field.
- Direct comparison with Cisco 2960S

HP Networking Product Warranty & Support (Feb 2012)

Fixed Configuration Switches (like A5800 (L3) and A5120 (L2))

- Warranty Duration:
  - Lifetime (for as long as you own product)
  - Lifetime includes integrated fans and power supplies.
- Advanced Replacement Delivery:
  - Next Business Day
- Tech Support (Phone/Email/Electronic Case):
  - As long as owned
- Software/OS Releases:
  - As long as owned
- Cost:
  - Included in purchase $0

[Cisco L3 switches must have SmartNet. L2 covered by E-LLW, similar but with some differences to the HP terms and conditions.]

Decision Factors

IRF Stacking (Intelligent Redundant Framework)

- Based on 10G full-duplex interfaces, much like VSS™ and FlexStack™.
  - 10G vs 16G of StackWise+™.
- Any 10G port, any media type (Cu or FX), any module. Also on 40G ports were available. Stacking 1m or 300m apart!
- Like VSS™, FlexStack™
  - Distributed Device Management
  - Distributed Resilient Routing
  - Distributed Link Aggregation
- Stacks only within same switch family, but all families can stack.
- 4 member stacks now including large chassis!
- 9 members on some platforms (A5800)

Module Slots – flexible option

- 5800 (L3) – 1 slot of 16 SFP or 16G TX, or 4x SFP+/XFP)
- 5120 (L2) – 2 slots, each 2x 10G (SFP+, XFP, CX4)

Decision Factors

- Rich feature sets. All the features we need and then some.
- Drive away pricing, no license levels. All features available. No reboots to change licenses.
- Specifications/Capacity (mac, arp, routes etc) adequate for our network size, still with substantial headroom for expansion, without being over specified.
- REUSE of existing optical transceivers (1G SFP)
- 3rd Party DA (Direct Attach) SFP+ (TwinAx) cable support
- Energy efficiency, low power per port. Cost effective RPS options.
- 128 port aggregations (etherchannels) per switch or stack.
- Line-rate performance, 1:1 backplanes.
- Documentation for switches adequate. Approx 3500 pages each. Quite descriptive and educational really.
- Build Quality

Decision Factors

- Cheat-Sheet. Large document that compares Cisco IOS commands to Comware and Procurve equivalents
- Uniform OS (ComWare) experience, uniform command syntax across all models.
- Quick delivery time-frames.
- Standard warrantee gives the equivalent as having paid maintenance with other vendors. Most of our existing switches do not have maintenance
- Lower CapEx results in budget stretching to stocking of spares for better MTTR than the best maintenance can achieve.
- Lower CapEx results in “test lab” being outfitted with identical parts, not obsolete hand-me downs hoping they are similar enough to adequately validate designs.
- iMC (Intelligent Management Centre) software quite powerful. Multi-vendor support so monitoring/management of HP, Cisco, APC, Powerware, WAPs etc. Bits of Nagios, bit of Cacti functionality as well. Replaces our CiscoWorks LMS 2.6. iMC does Cisco support almost as well as older LMS.

Implementation

- From order being faxed to HP, pallets started arriving in a fortnight, in Hobart.
- All parts arrived within approx 4 weeks.
- All switches could be “spun up” and configured in the test lab simultaneously. The entire network was active and tested. BGP, OSPF, MPLS, the lot.
- Sites were migrated after-hours, new switches at 10G connected to new core MPLS device. Experience improved the process to the point that the latter devices were installed with less than a 5 minute outage for the remote campus.
- 80% of campuses moved to the new MPLS core network within 10 weeks of the order being placed.
- **All existing users, subnets, and vlans initially placed in a single VPN/VRF, so no impact on functionality or security. At appropriate times vlan/subnets will be repositioned into individual VPN to enhance the security and provide the isolation. Consultation process and impact to discussed first.**

Has the Sky fallen?

- Extra Staff needed to support multi-vendor network?
  - Nope. Just very busy like in any refresh cycle.
- Higher Training Costs?
  - Nope. Didn’t even bother with ”delta” training in the end.
  - “Networking” skills transferred just fine.
- Higher OpEx due to failure rates
  - Nope. Similar MTBF. If anything more reliable, as younger.
- Is it more complex to manage?
  - Not really. Like switching from driving on the left to driving on the right. A few miles under the wheels and it feels natural.
  - Configurations have been cleaned up, simplified. 18 years of tweaks retired. Standard template applied.
  - Configs are “smaller” as more features are enabled by default
- Slow to deploy?
  - Not really. Takes time to negotiate outage windows. The time saved by “auto-configuration” is minor compared to the transport, unpacking, attaching brackets, finding space, moving cables, using screw driver. Only takes a few minutes to cut’n’paste a standard config template in via the console anyway. And you know it is right.

Come on, really, what issues were there?

- Missing Cisco proprietary protocols not the end of the world.
- LLDP can replace a lot of CDP. And there is a CDP listen mode.
- GARP/GVRP can do some of VTP. Or use iMC management suite to bulk config change switches to add new vlans.
- Turn off VTP Pruning. Really. Turn it off. Now.
- Turn off UDLD aggressive. It is propriety and will stop links coming up.
- Proxy Arp is off by default on HP. Should not matter if all devices are configured correctly. Are they? How can you tell with proxy arp on by default with Cisco? Know what defaults change between vendor.
- Do changes in clusters, buildings, campus. Don't use a shotgun approach, makes it harder to trace and identify issues.
- Biggest issue was spanning tree. Oldest Cisco don't play nicely or support latest standards, the HP don't bend backwards that far with support either. The 2900XL have to go. Worked around it.

What would you change?

- Displaying interface counters are in byte! What networking person doesn't work in bits?
- Unable to use SNMP to access counters for SVI (vlan interfaces)
- Need some deep training. Using debugging a bit hazy.
- HP Web site a bit slow. No where the same depth of networking resources. More product and support than education or general knowledge.
- Ummm…

HP Support

- We are not in Kansas anymore Toto.
- Support is not Cisco TAC. This can be good and bad.
- Approach with open mind. Expect things to be different.
- It has been satisfactory. It will take time to generate as much experience with HP support so as to be able to compare with years of experience with other support organisations. So far, it has worked.
- Good people work everywhere. Find them and use them.
- Critical larger equipment is on SmartCare 24x7 support. Responses and actions have been acceptable. It is Tasmania after all.

Future Vendor?

Are we moving away from Cisco after 18+ years?

Short Answer: Not intentionally

Long Answer: Each functional unit(#) of the network should undergo an competitive evaluation when being replaced, on a broad range of criteria, including but not limited to:

- Meeting operational requirements, functionality, capacity
- Integration/training costs
- Value for Money
- Interoperability

Results in best option for University, maximising ROI, minimising CapEx, efficient OpEx, and meeting University requirements in a practical and cost effective way.

Keeps them (vendors) working for your affection/dollars.

Bundling is evil. If I really want some steak knives, I'll go buy some good quality ones!

(#) VPN, Wireless, Firewalls, Internet Traffic Control, Load Balancing, Access Layer, Distribution Layer, Core, Data Centre Switching, Identity Management, UC, VC.

The end, or a new beginning?

- Questions?