

CUDN* Update 2016

Bob Franklin <rcf34@cam.ac.uk>
UIS Networks, University of Cambridge

Agenda

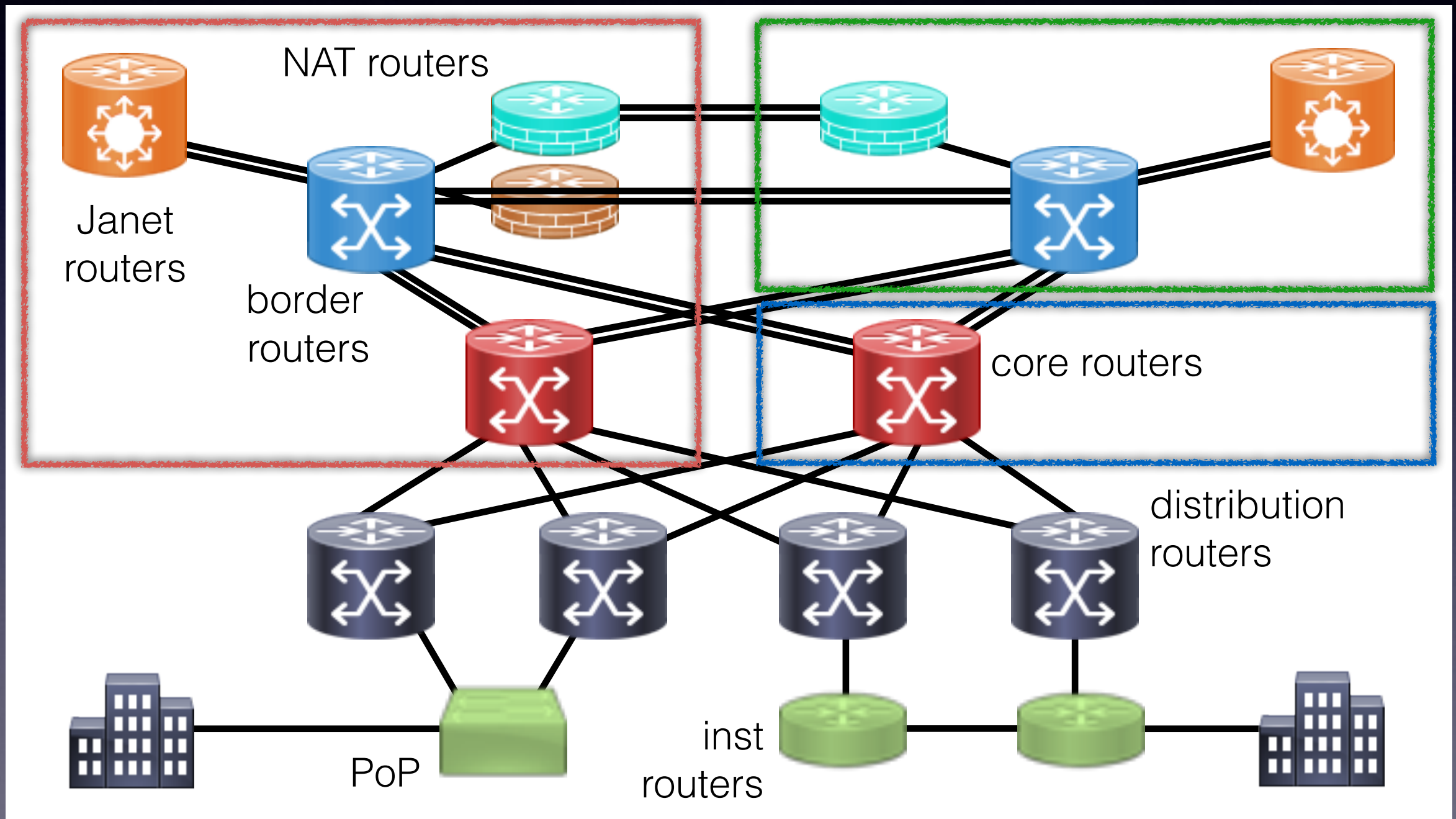
- Post merger update effect on Networks
- CUDN and Networks developments 2015-16
- CUDN upgrade 2016
- PoP upgrade 2016-17

Effect of UIS restructuring

- UCS Network Division has become **UIS Networks** and remained largely unchanged
- Has inherited responsibility for the **ACN** (Administrative Computing Network — the UAS network run by MIS) and staff — current plans include:
 - Refresh edge equipment
 - Absorb backbone into CUDN (probably using MPLS service)
 - Server network will gradually migrate into updated replacement UIS server network (ex-UCS and ex-MIS)
 - Not much effect on the CUDN, except interim workload
- Have also merged with **Hostmaster** group (Tony Finch)

CUDN Developments 2015-16

Second Janet connection



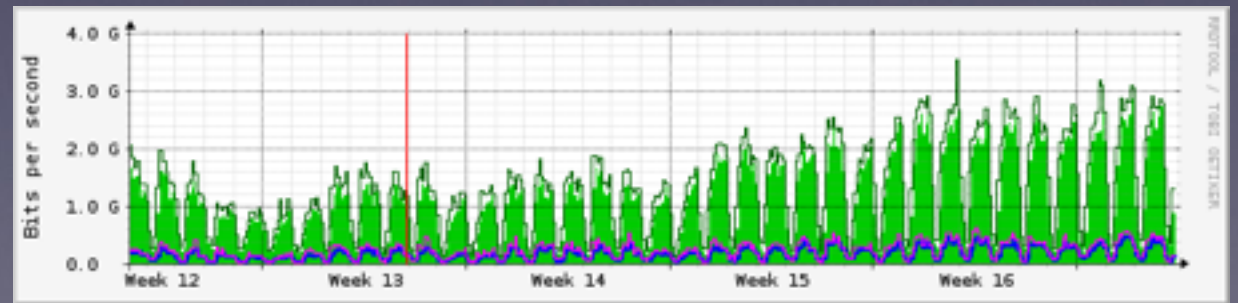
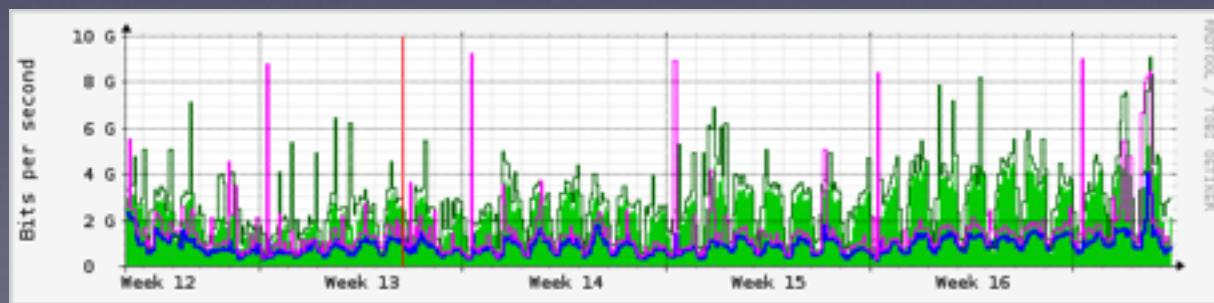
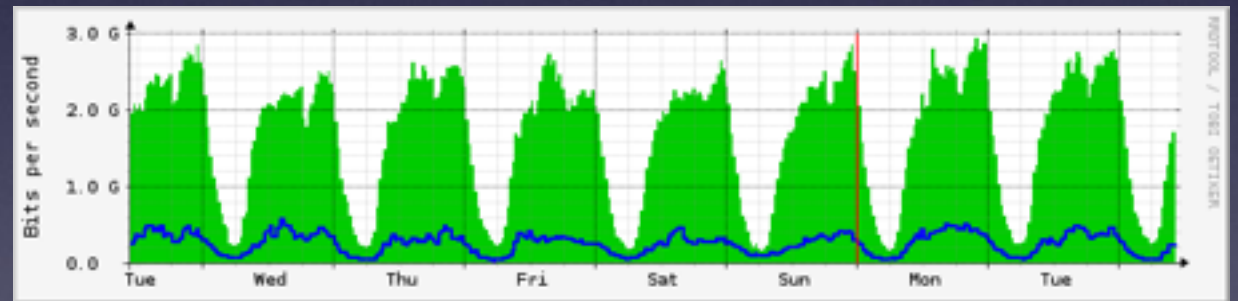
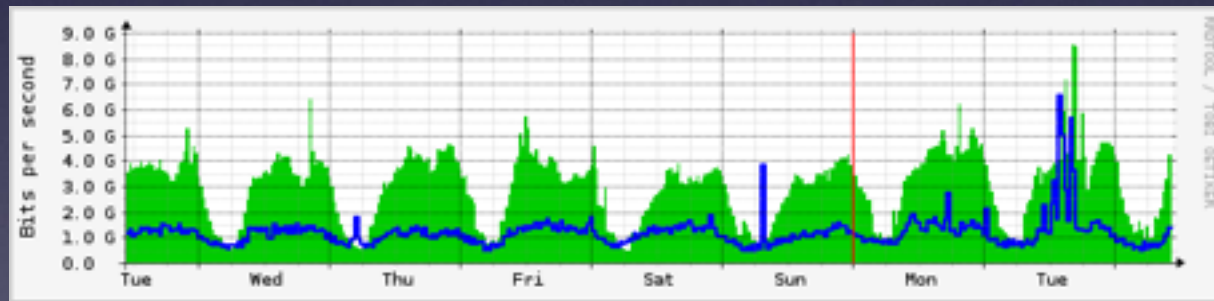
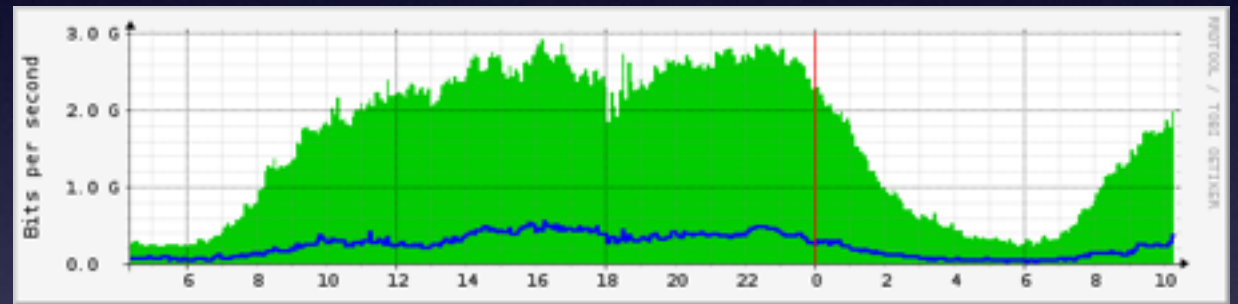
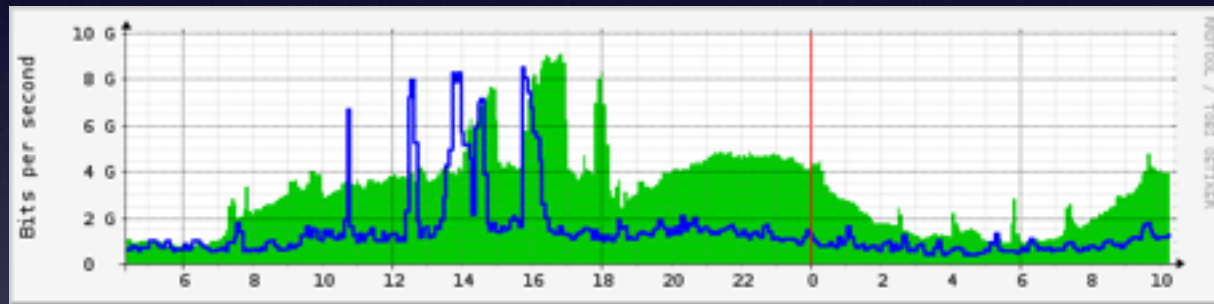
Janet link configuration

- Most traffic goes in and out of the primary link:
 - All non-NATed global addresses
 - IPv6
 - Multicast (IPv4 and IPv6) — usually
 - Bottom half of the NAT (172.16.0.0/13) + UniOfCam
- Secondary link is used for the “upper half” of the NAT (172.24.0.0/13) + eduroam
- Traffic flow is symmetric in most situations

Janet link utilisation

Primary @ ~4-9/20Gb

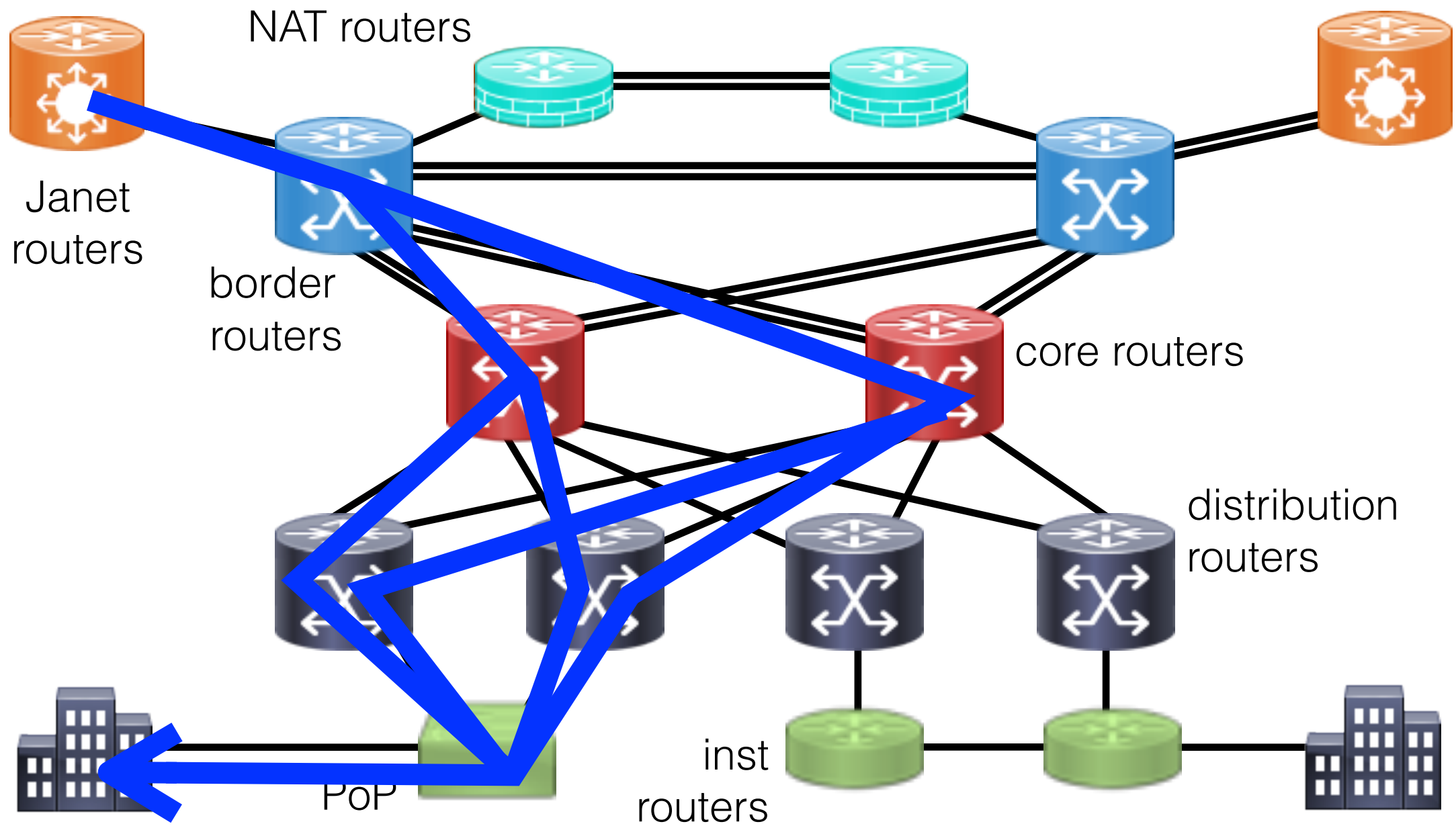
Secondary @ 2-3/20Gb



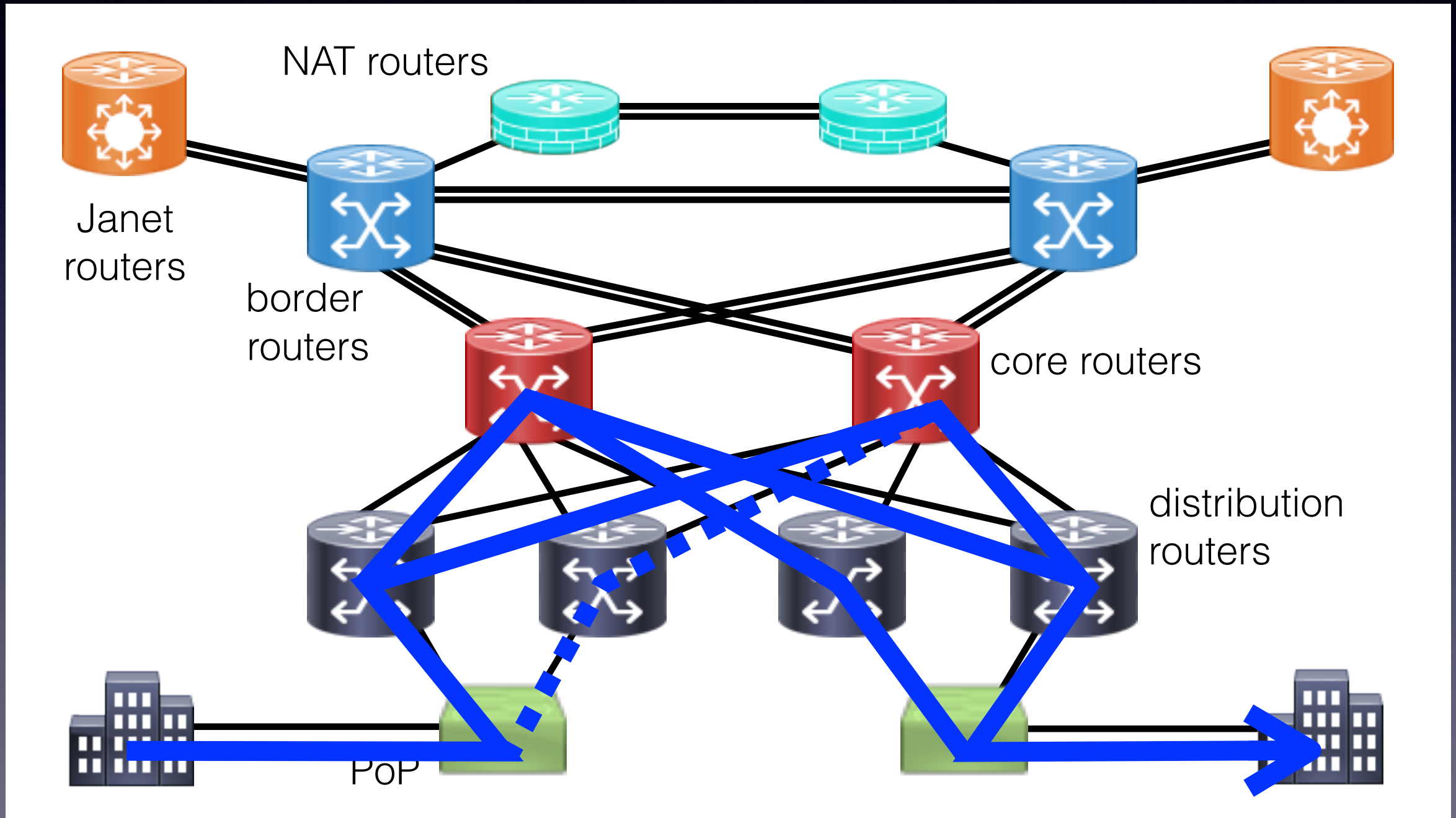
Multipathing

- Traffic used to be carefully steered to use *most* links on the CUDN:
 - Institution ↔ Janet traffic via one core
 - UCS servers ↔ Janet + Institutions; Institution ↔ Institution via the other
- Traffic now takes **all paths**:
 - Both links down to an institution utilised (up to of 2Gbit/s down to a PoP)
 - (Still only a single link used up from a PoP at present)
 - Path is consistent for a particular flow
- Some subtle issues needed resolving (DHCP broke)

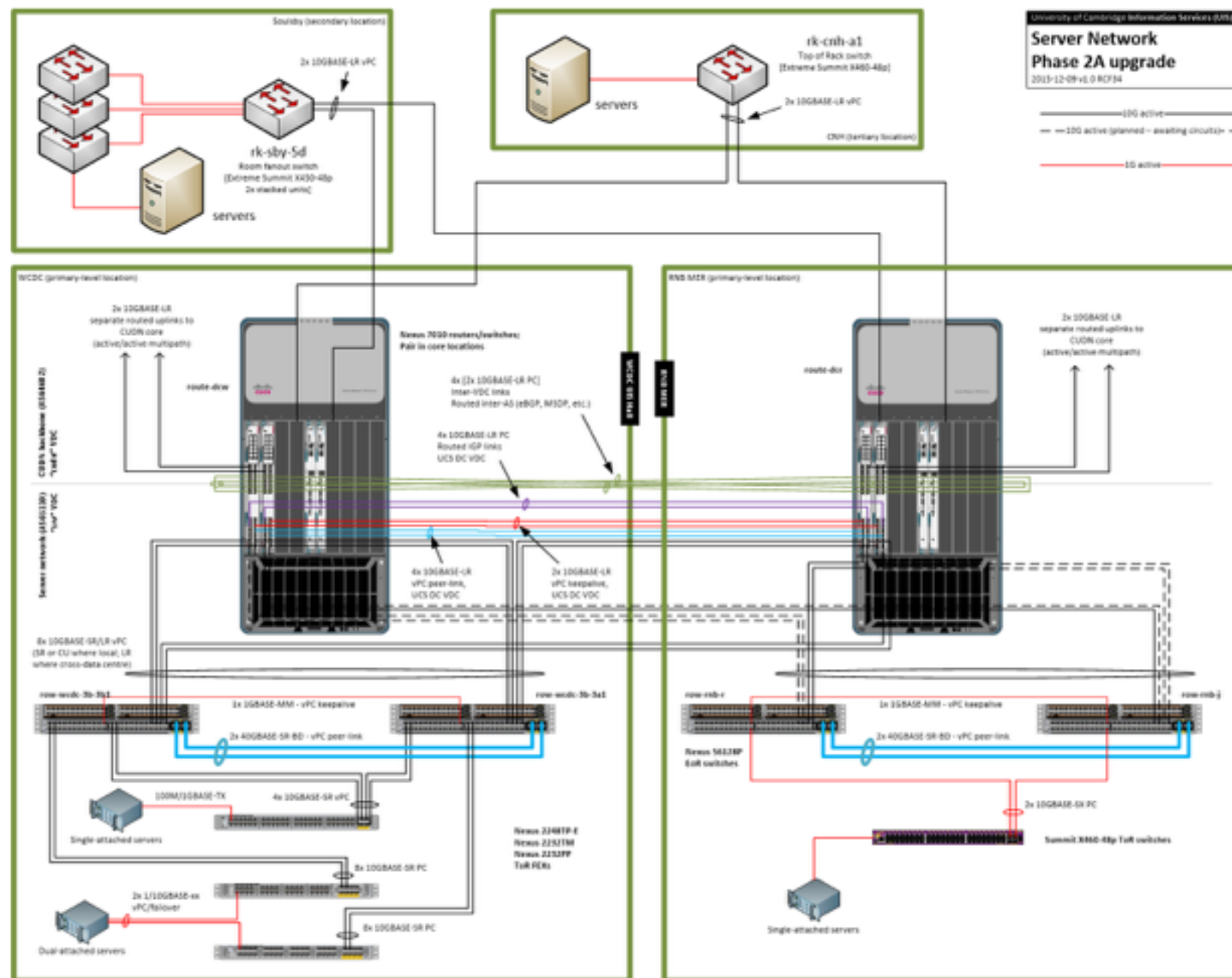
Multipath inbound



Multipath between insts



Upgraded data centre



DNS service upgrade

- Servers were 3x Sun Solaris boxes each doing authoritative/recursive — 2x active, 1x standby for manual failover / upgrades
- Upgraded February 2015 to:
 - 4x physical Linux recursive servers split across 4 UIS data centre locations with automatic failover
 - 4x virtual Linux authoritative servers split across locations — 2x active, 2x manual standby
- Off-site secondaries at Imperial College and ISC anycast service

VPN service replacement

- December 2014 introduced new **VPN Service**
 - Based on open source VPN software running on Linux hardware
 - Support built-in clients on common OSs
 - Uses same *Network Access Token* as eduroam
 - Greater performance
- March 2015 introduced **Managed VPN Service**
 - Has dedicate client IP address pool
 - Restricted to subset of users (by Lookup group)
- April 2015 retired old VPDN service

CUDN Upgrade 2016

Upgrade philosophy

- "Tick-tock" process — decouple equipment upgrades from topology changes
 - Replace equipment in same configuration — can deal with problems with equipment and purchase equipment to resolve issues based on experience of running the network
 - Change topology using same equipment — resolve issues with familiar equipment
- Avoids big bang upgrades and minimises chance of potential problems
- At a "tick" phase — changing the equipment (2011 was a "tock" phase: same equipment but new topology)

Current equipment

- Based on **Cisco Catalyst 6506/6509**
- Originally bought in 2004
- Minor refresh in 2010 (E-series chassis, new fans, upgraded Supervisor 720 cards)
- Some modules have gone End of Support; remainder in next ~12 months

New border routers

- **Cisco Catalyst 6880-X**
- Bought in May 2015
- Based on Catalyst 6807-XL, Supervisor 2T with 16x 10GE port line card
- First installed during flooding incident in July 2015 at secondary location
- Second installed in November 2015 to replace flood-damaged unit at primary location



Data centre routers/switches

- **Cisco Nexus 7010, Nexus 56128Ps and Nexus 2Ks**
- Replaces Catalyst 6509-Es and Extreme Summit X460s
- Bought in February 2015, deployed March 2015 (WCDC room level), August 2015 (core level) and January 2016 (RNB room level)
- Cisco platform designed specifically for data centre deployments
- 40Gbit/s to CUDN; 80Gbit/s to each location (all active)
- Moved from CNH to WCDC (+ RNB)



New core/distribution routers

- **Cisco Catalyst 6807-XL**
- Ordered April 2016; expecting delivery May 2016
- Will be installing during summer vacation 2016
- Replacing one router at a time; using inherent CUDN redundancy to maintain service
- Some preparation needed



Supervisor 6T

- **Catalyst 6800 Supervisor 6T**
- “CPU” of the router
- Brand new
- 2x 40GE ports for uplinks to core routers (eventually)
- 8x 10GE ports for institutional or other connections
- 1x in the distribution routers;
2x in the core routers



32-port 10GE cards

- **C6800-32P10G line cards**
- 32x 10GE SFP+ ports; also support 1GE connections
- Used for institutional connections
- Local distributed forwarding module to improve performance
- Consistent module across all routers — simplifies spares



Backbone links

- Initially will run the core↔distribution links at the same 10Gbit/s speed
- Late 2016 will be upgrading to 40Gbit/s using 40GE links on the Supervisor 6T
- Multipathing configuration means core bandwidth is between 40 and 80Gbit/s
- (Can add additional 40GE line cards to distribution routers if this becomes an issue)

PoP Upgrade 2016-17

Current PoP equipment

- **Cisco Catalyst 3560G-24PS** switches with **RPS-2300** redundant power unit
- 24x 1GE RJ-45 ports w/ PoE, 4x 1GE SFP ports
- Bought in early 2008 as preparation for the VoIP project
- Have recently gone out of software support
- APC UPS equipment

Current PoP issues

- Some 10GE customers: all getting bespoke service and need to justify need due to limited capacity — 10GE upgrade requests now common
 - Even if general load is under 1Gbit/s, peaks can result in occasional packet loss
- RPS is annoying extra unit (bigger than the PoP!)
- UPS not always desired and batteries need replacing regularly
- BGP customers pay the same as a PoP customer but get no switch

New PoPs

- Will commence procurement during summer 2016, whilst the backbone rollout is in progress
- A number of options (1G and 10GE) depending on bandwidth requirements of the end institution
- Institutions will be required to confirm their choice for the next few years to avoid procuring an excessive number of unwanted switches
- Currently unsure of the exact model/specification
- Expect details and choice to be made end of CY 2016

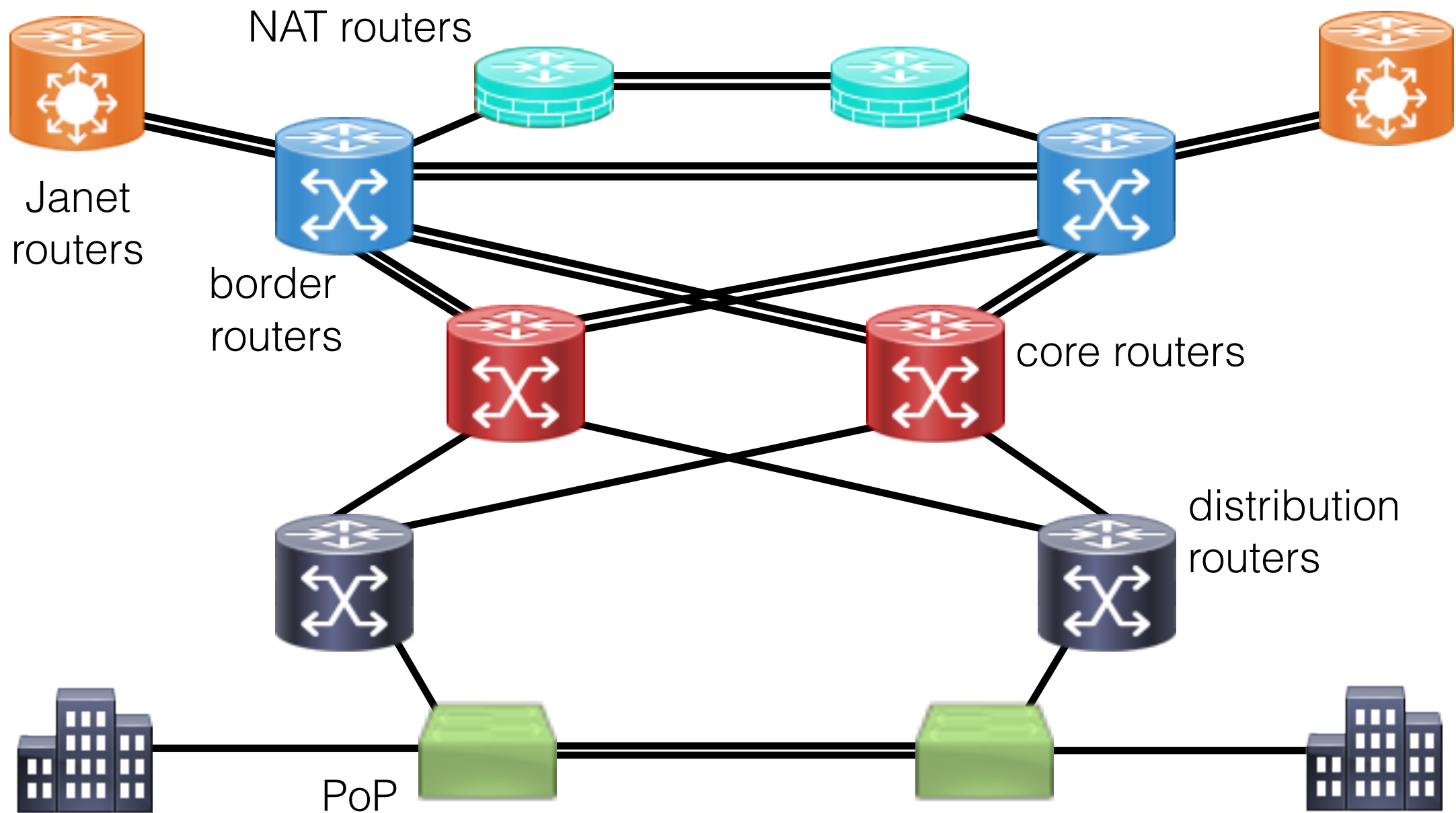
Other PoP developments

- RPS expected to be integrated into the switch
- UPS will be optional
- BGP will be priced lower to reflect the lack of a switch
- PoP configuration likely to change re: spanning tree, etc. to increase filtering of layer 2 problems — Techlink to follow!

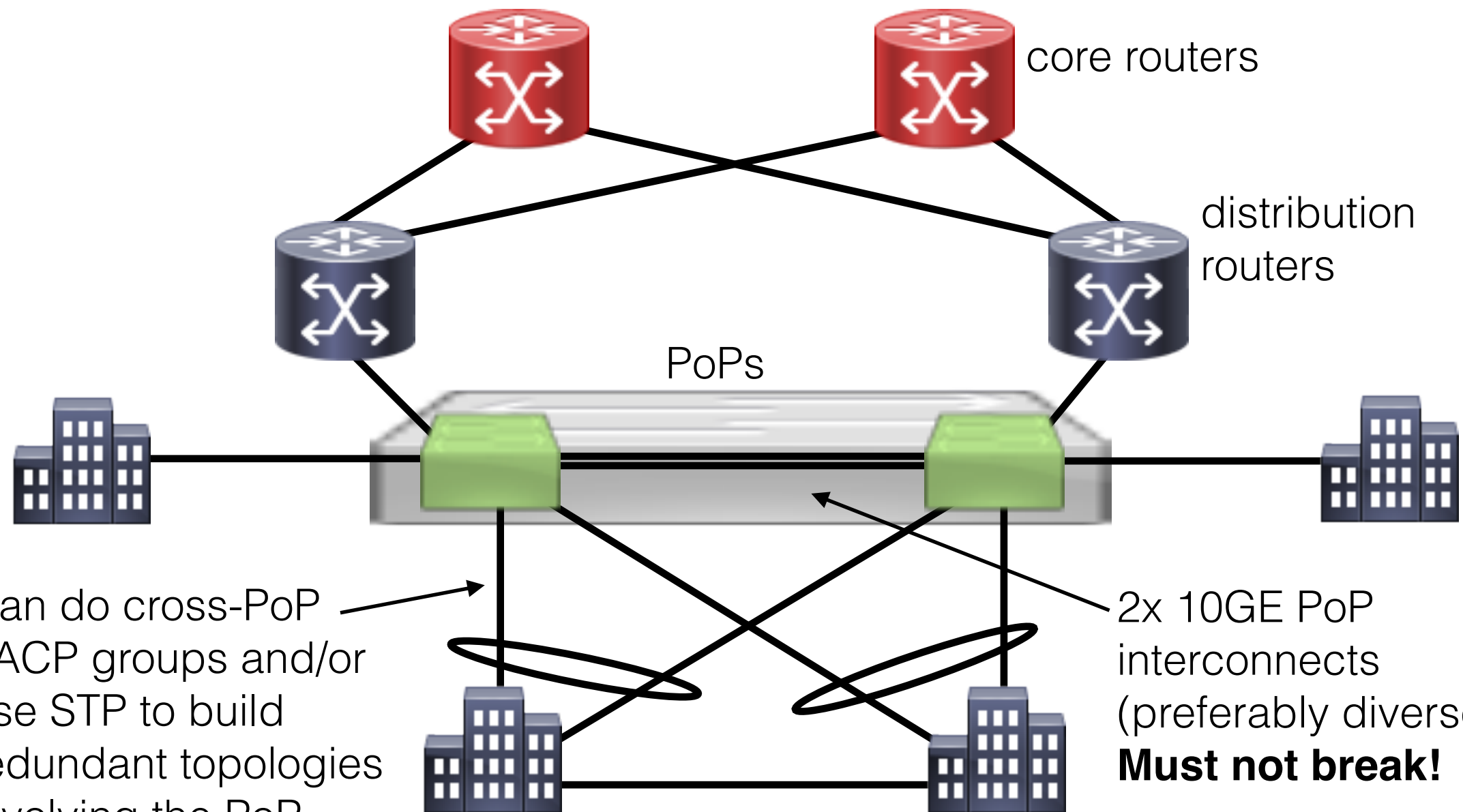
Split PoPs

- Expect to be able to offer two PoP switches split across sites but functioning as a single logical switch — LACP, same VLANs, etc.
- Gives physical redundancy (location and hardware)
- Uplift in price will be:
 - Additional switch hardware
 - Cost of fibres + transceivers to interconnect sites
- Not as independent as BGP solution (e.g. combined logical switch and propagate software faults between units; major software upgrades still sometimes disruptive)
- Currently not guaranteed but possibly early CY 2017

Split PoP arrangement



Split PoP arrangement



Can do cross-PoP LACP groups and/or use STP to build redundant topologies involving the PoP

2x 10GE PoP interconnects (preferably diverse)
Must not break!

PoP choices (likely)

Model	Uplinks	10GE downlinks	1GE downlinks	Single PoP	Split PoP
1GE	2x 1GE	-	Multiple (~20x) RJ-45 PoE+	£2,700	£3,400 (+ £700)
10GE	2x 10GE	2x SFP+	Multiple (~20x) RJ-45 PoE+	£5,900	£7,500 (+ £1,600)
Multiple 10GE	2x 10GE	Multiple (~8x) SFP+	-	£5,600	£7,000 (+ £1,400)
BGP 1GE	2x 1GE	-	-	£4,000	-
BGP 10GE	2x 10GE	-	-	£4,500	-

End

The deadline for
Chest orders for
wireless APs is
end of May!