

The Fifth-Generation UIUCnet

Charley Kline
UIUCnet Architect
Communications Technologies
CITES



Agenda

- Review Current “Cube” Architecture
- Discuss Current Problems and Issues
- Unveil New Core Design
- Unveil Building Reference Model
- Details



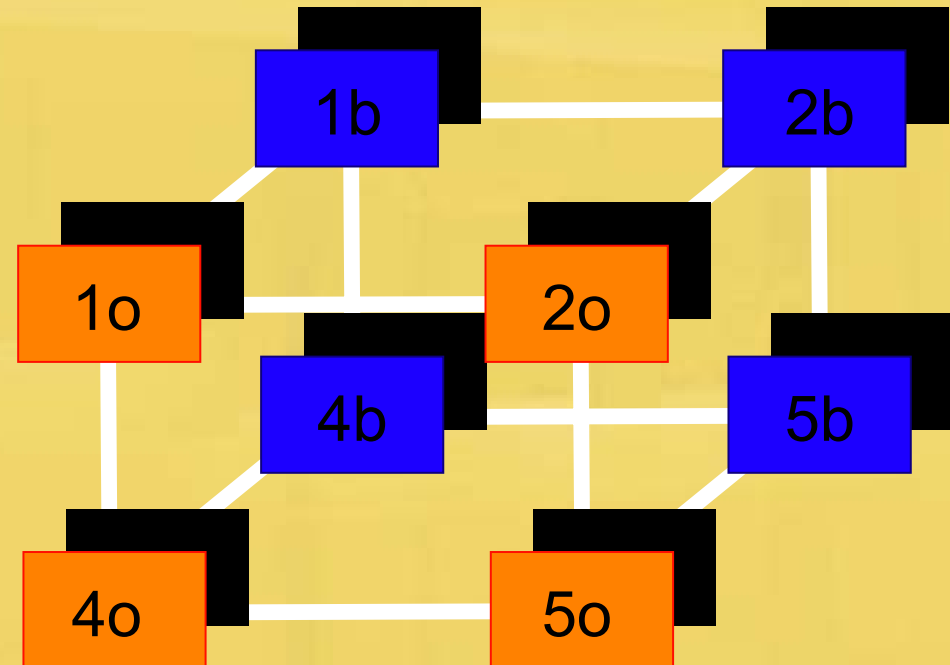
Current Gigabit “Cube”

- Built in 2002 to address the need for a uniform, high-speed design to replace the ad hoc gigabit and 100 megabit Ethernet backbone
- Attempted to address perceived need for redundancy
- Built as two parallel rings spanning the four major nodes



Cube Architecture

- Two devices in each node
- All Gigabit Ethernet
 - Gigabit building links plug directly into these devices
 - 10- or 100-meg links still go to legacy devices dual-connected to each core switch
- Foundry Jetcore technology
- Does VLAN switching as well as IP routing



Issues and Instabilities

- The eight core switches do ALL the work: switching, routing, spanning tree, security, etc.
- Many VLANs are present in multiple nodes, thus have to be configured on all eight core switches, creating a very complex topology for these VLANs
- There are too many VLANs crossing the core, and each must be managed by an instance of the Spanning Tree Protocol

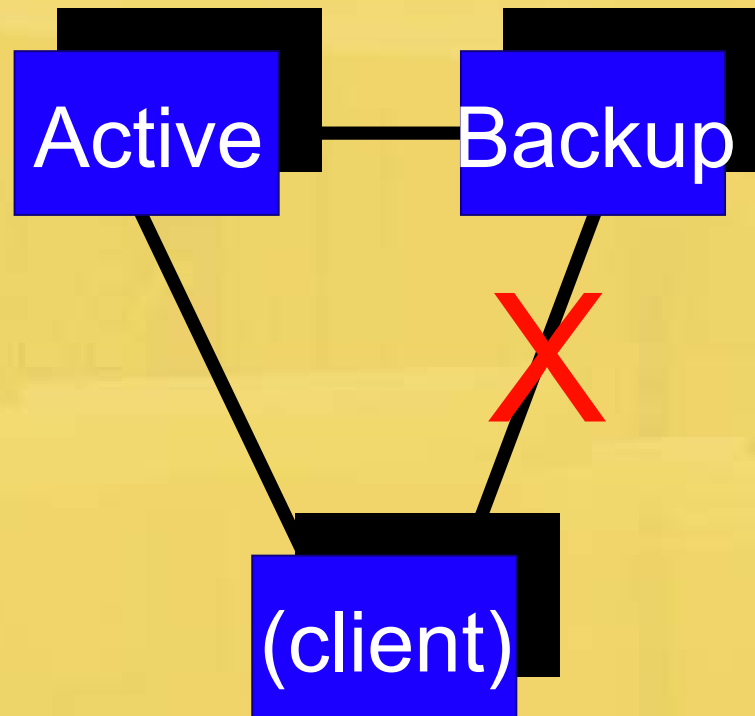


Issues and Instabilities

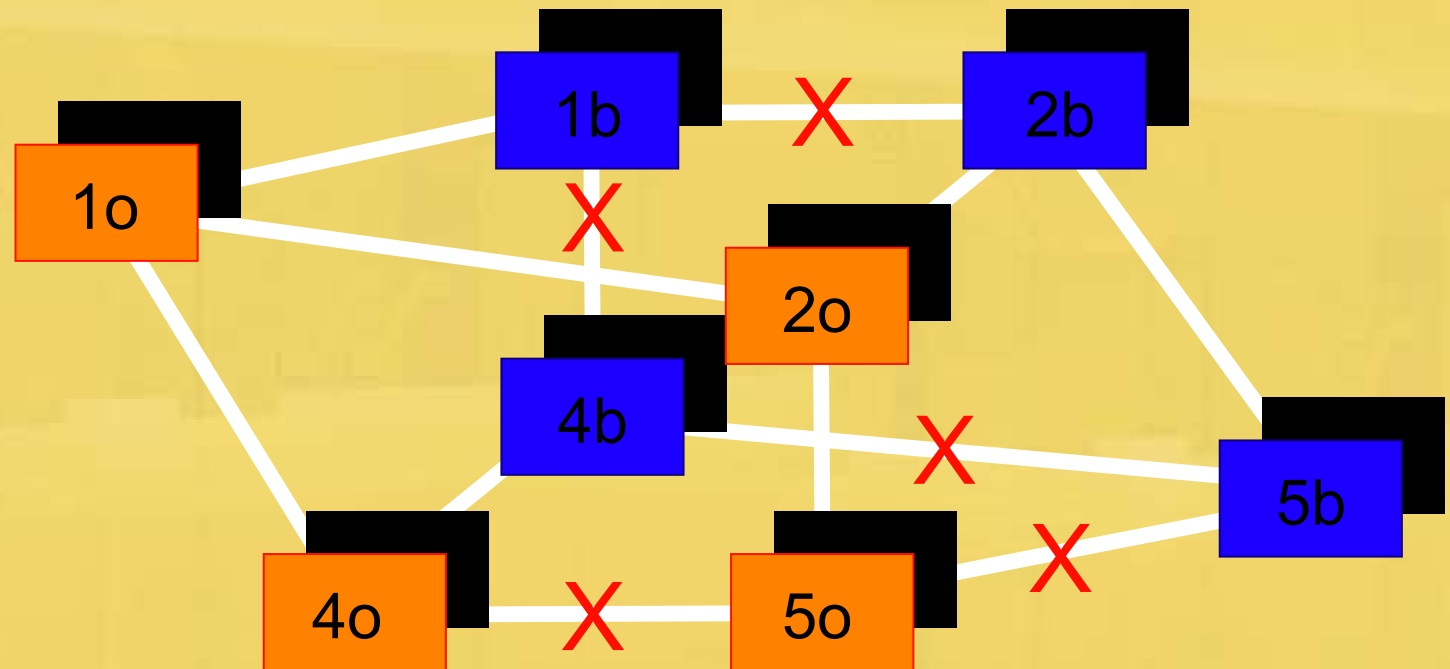
- STP was not designed for arbitrary topologies such as this – there are MANY redundant paths that need to be examined and blocked
- We're exceeding the maximum switch diameter for many large VLANs (creates instability)
- STP extensions such as 802.1w really want to operate on a “working / standby” topology, which this is not



Triangles are Awesome



Triangles are Awesome



Exercise: Find the blocked links



Issues and Instabilities

- The current core switches were procured in 2002, just as the state budget tanked
- Foundry FastIron-series switches were economical but not backbone-class devices
- Tight on hardware cache space due to combination of L2/L3 work



Lessons Learned

- Too much redundancy makes things *less* reliable
- Divide up labor where possible, and leverage economies of scale
- Architecture should reflect topologies where STP can do best
- Use equipment sized for the job



New Design Goals

- Fewer, bigger, beefier boxes with lots of INTERNAL redundancy/availability features
- Larger density → smaller box count
- 10 Gig will be a must
- Advanced features such as IPv6 and MPLS should be provided for



New Design Goals

- As much as possible, try to split up L3 from L2 work—ideally the top-level switches do *only* L3 routing
- Design for STP (working/backup triangles)
- Strive to break up large VLANs, especially where they span nodes
- Allow for L3 routing in building devices where it makes sense



The Plan

- In Fall 2004, Network Engineering began requirements and design meetings for a new UIUCnet architecture to address these goals
- The basic plan was in place by early 2005
- “Building Reference Design” distributed to NDO for use in the upgrade project in July 2005

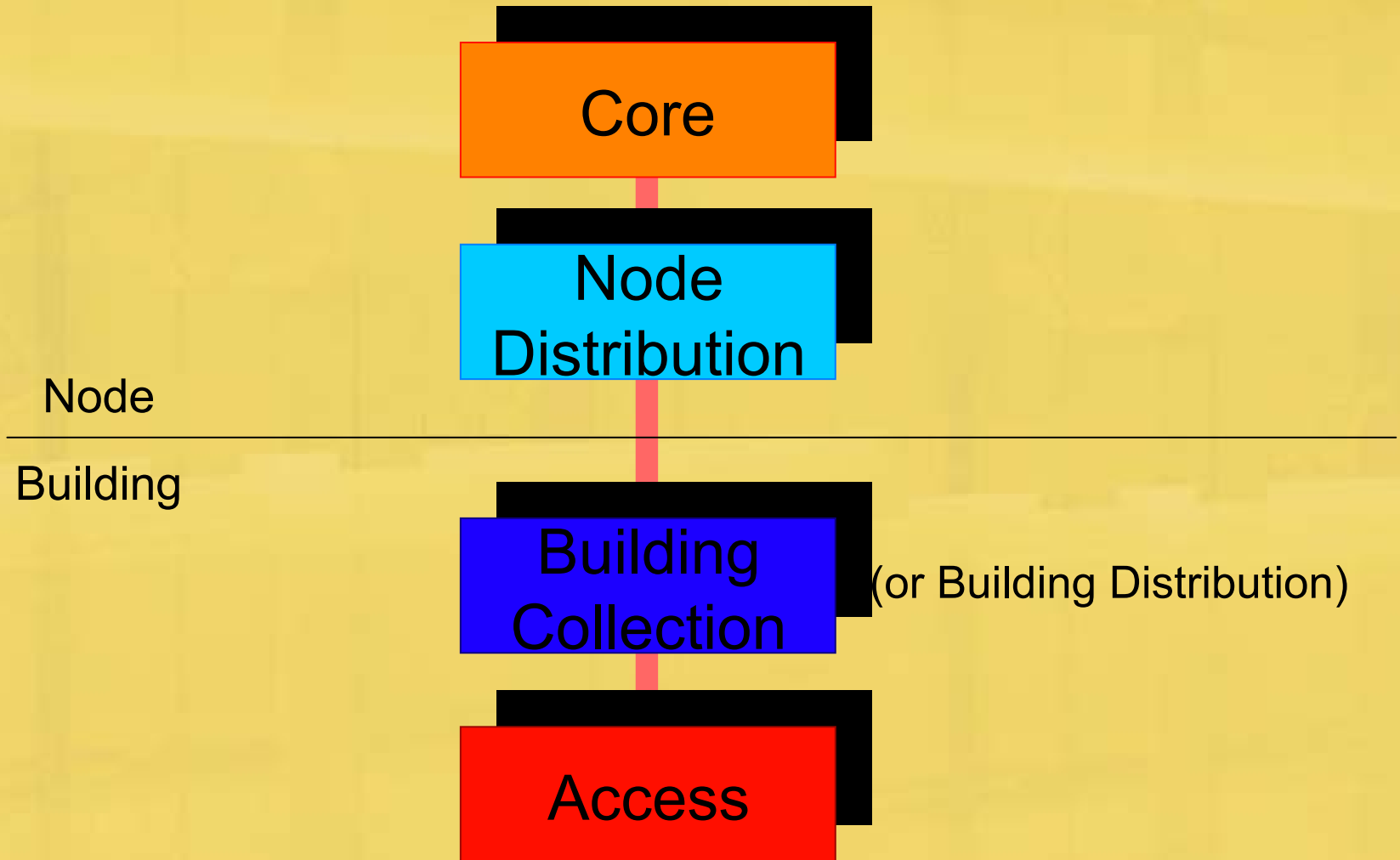


The Plan

- Vendor evaluations proceeded through spring 2005
- RFP process was completed in September
- Now awaiting Board of Trustees approval at their November meeting
- Hopefully, equipment to be installed over winter break with a spring break 2006 cutover

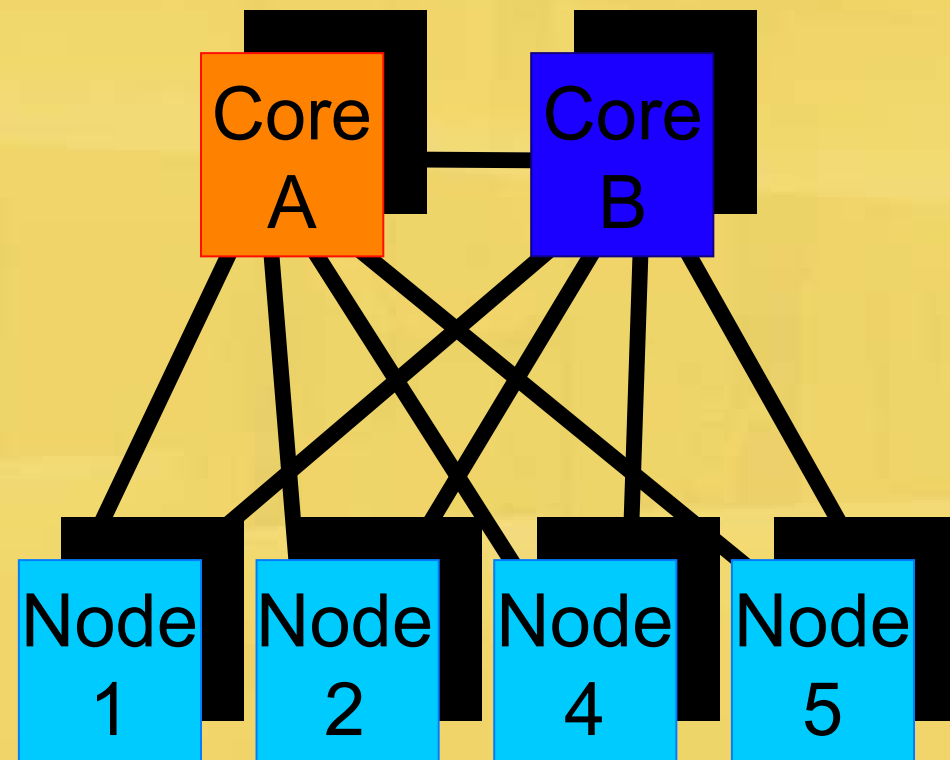


Quick Terminology Hierarchy



Core Architecture

- Based on a tree instead of rings (limits multiple paths)
- Redundancy is simply two interconnected core devices
- Each node distribution switch is dual-connected to both core devices
- Backup STP paths form triangles



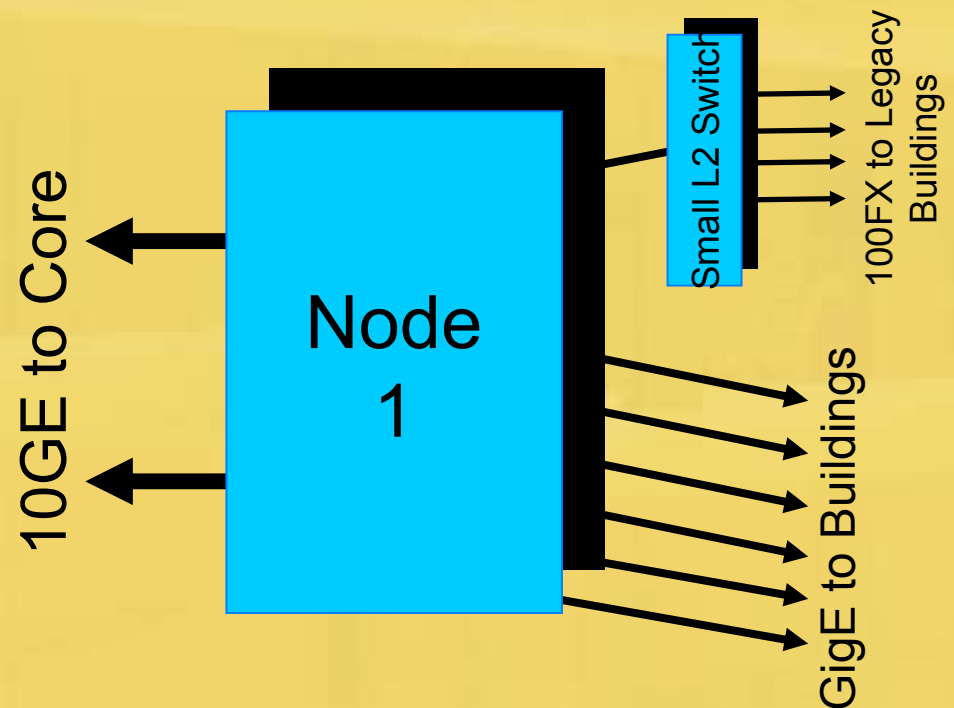
Devices Used

- Foundry NetIron 40G (pending BoT approval)
- 1.28 Tb/s backplane
480 Mpkt/s
- 10 Gig and 1 Gig interfaces
- Dual Management with “hitless failover”
- Lots of line card autonomy (processor and ASIC)



Major Node Design

- One NI40G
- Dual 10GE uplinks to core
- Feeds all gig-connected buildings in that node
- Where 100 meg is still required, small-form-factor Ethernet switches will be used to convert



Building Reference Design

- Describes a “base standard” network design for all network-upgrade buildings
- Can be deviated from to serve special building needs
- Discusses specific hardware recommendations; we’ll just cover the highlights here

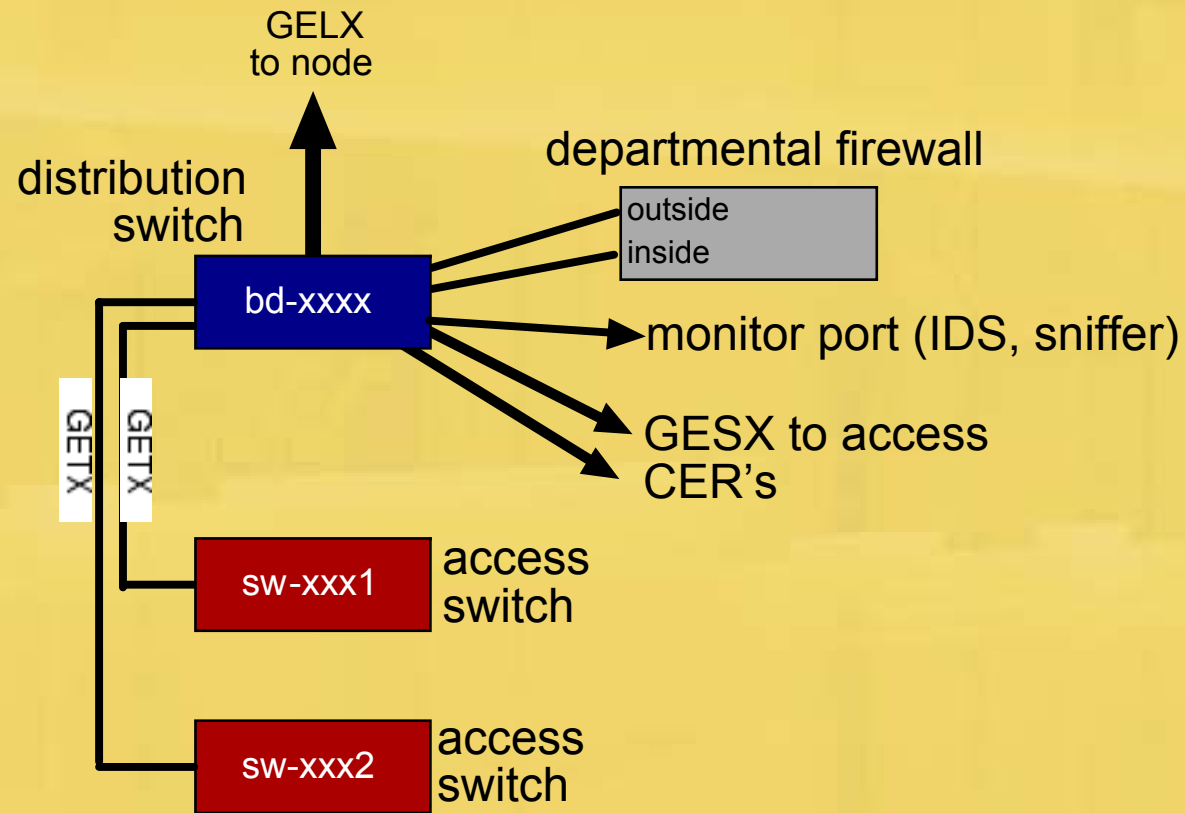


Building Reference Design

- CITES-administered “Building Collection” (L2-only) or “Building Distribution” (L2 + L3 routing) switch in the initial CER of *every* building
- Provides visibility into building infrastructure for monitoring and manageability
- `bc-bldgname.gw.uiuc.edu`
`bd-bldgname.gw.uiuc.edu`



Distribution CER



Building Reference Design

- Access switches are Layer-2 switching only devices
- Usually low-port cost but reliable devices
- Ideally, these are `sw-bldgnameN.gw.uiuc.edu` and are under CITES administration with departmental management via Iris

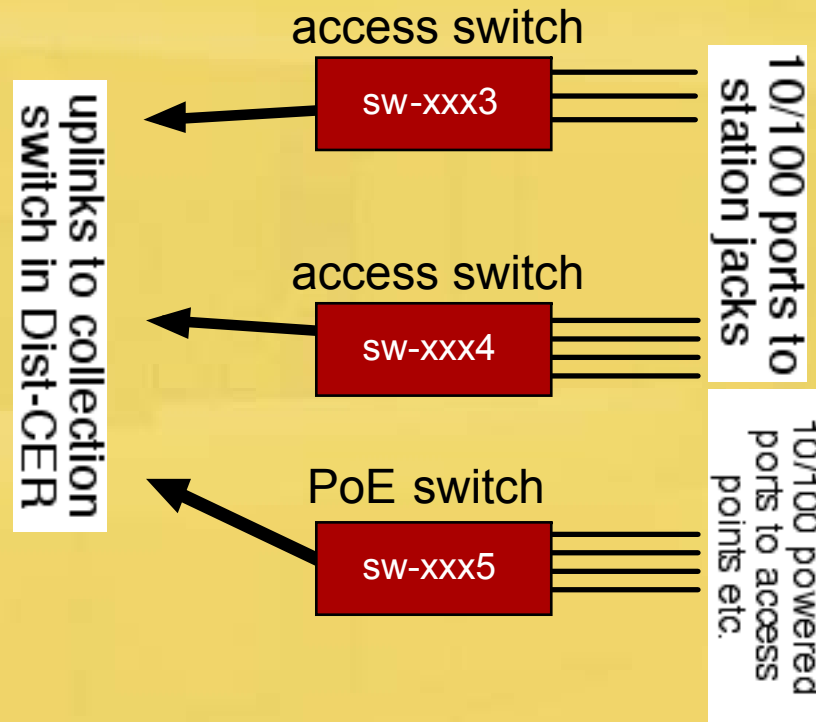


Building Reference Design

- Access switches may be under departmental control
- Departmental-controlled switches can be multi-VLAN, but only VLANs controlled by a single department
- Multiple access switches must be used when multiple department need control in a single CER



Access-Only CER



Building Routers?

- He said “Building.” Then he said “Routing.”
- Yes, we’re returning to building routers



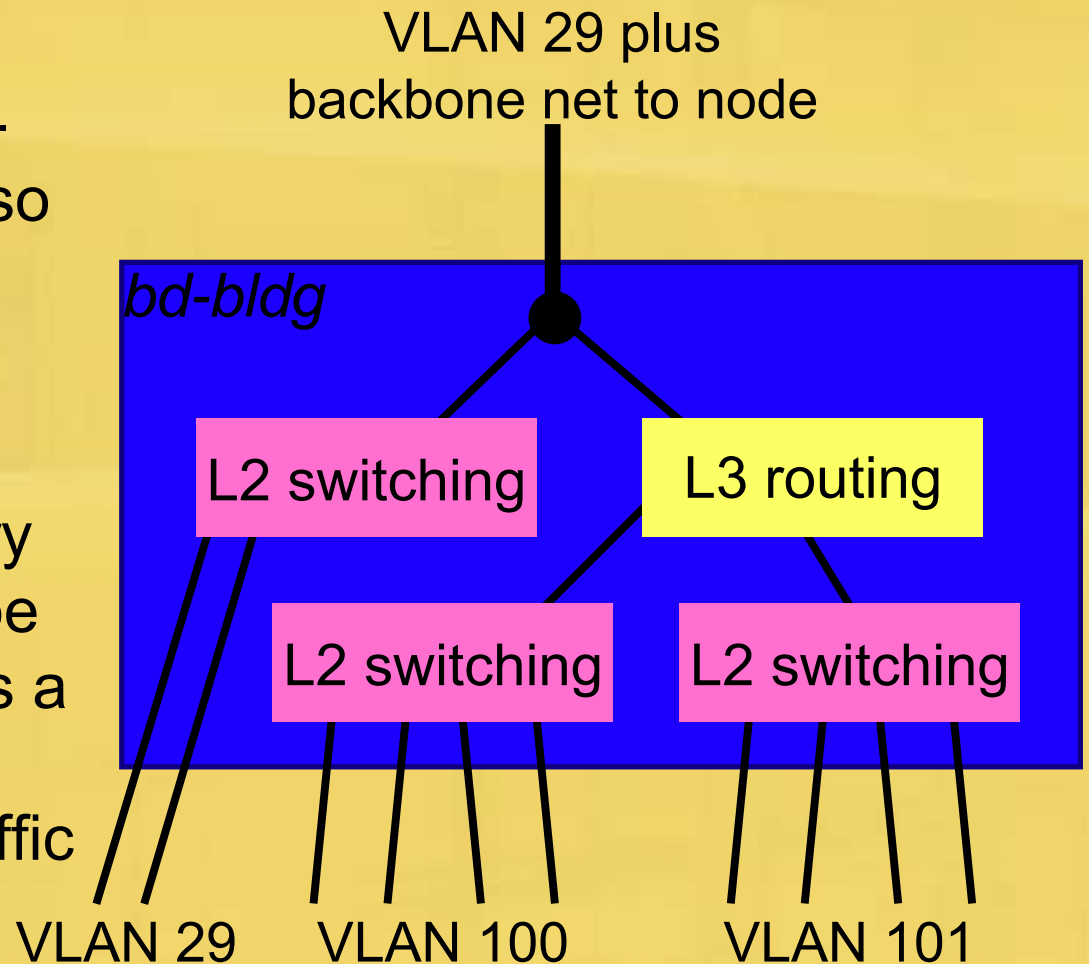
Building Routers

- Where a VLAN is present in only one building (happens quite often), the BD switch for that building also routes for that VLAN
- Improved scaling (Node switches no longer have to see every single MAC address on campus)
- Improved reliability (routing between VLANs in a building no longer depends on the Node)



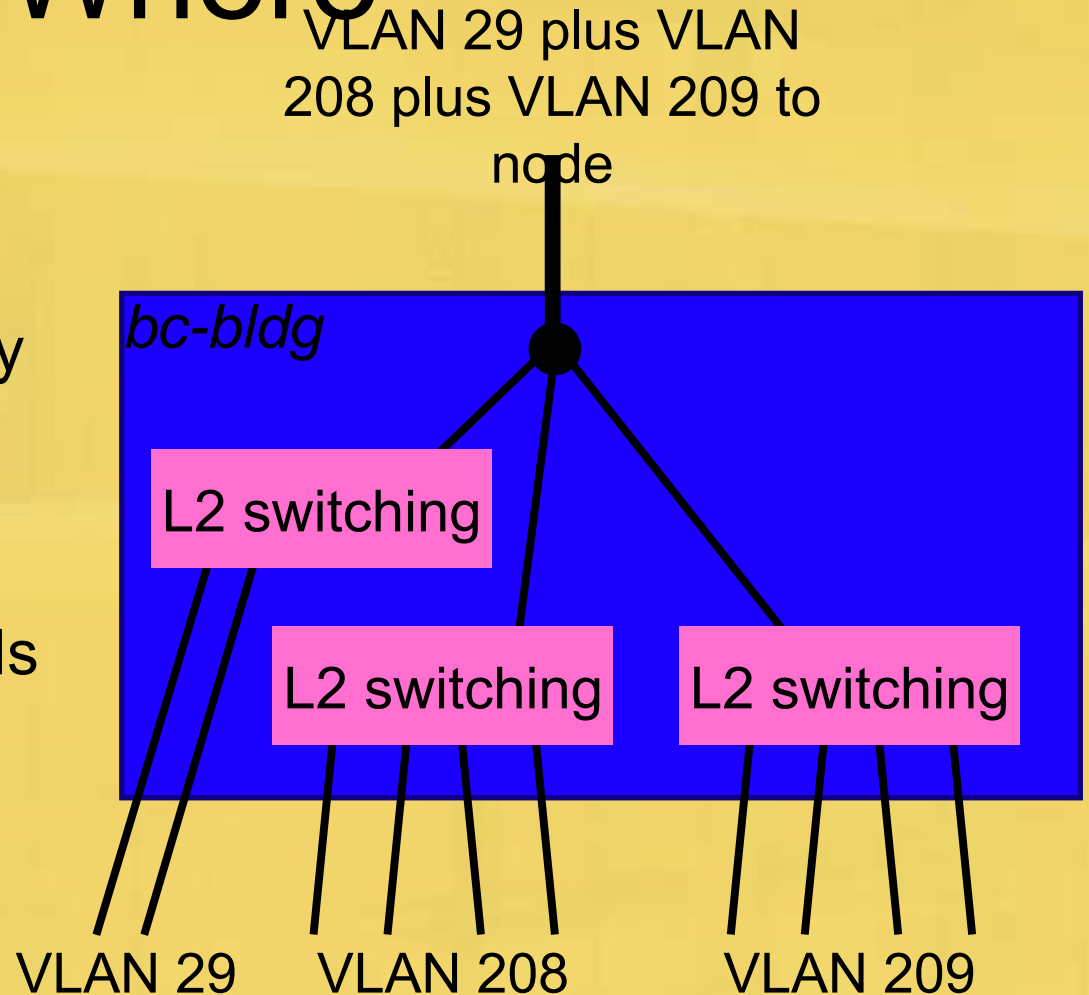
Details – What Happens Where

- A BD switch does intra-VLAN switching and also routes between VLANs present only in this building
- Uplink to node can carry L2-switched traffic (to be routed in the node) plus a backbone subnet to transport the routed traffic



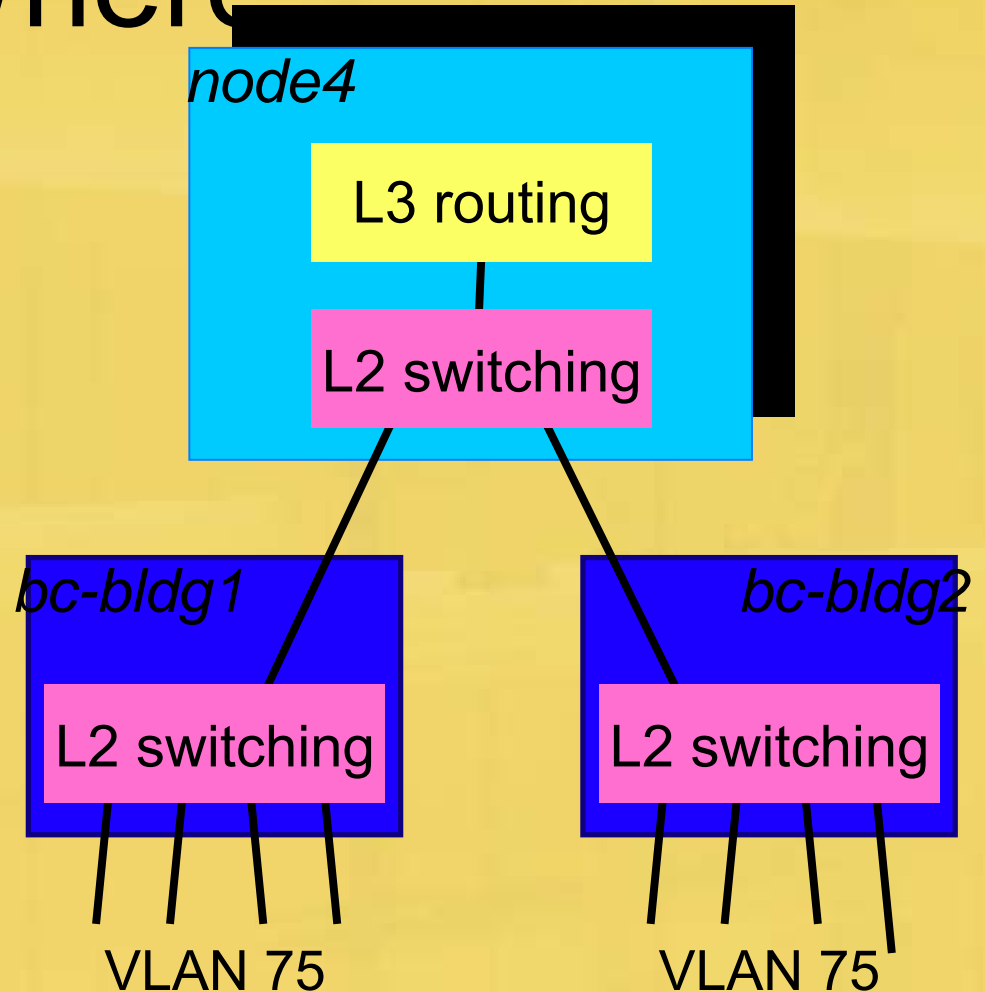
Details – What Happens Where

- A BC switch does only intra-VLAN switching and transports the switched traffic to the node, where all VLANs will be routed by the node distribution device

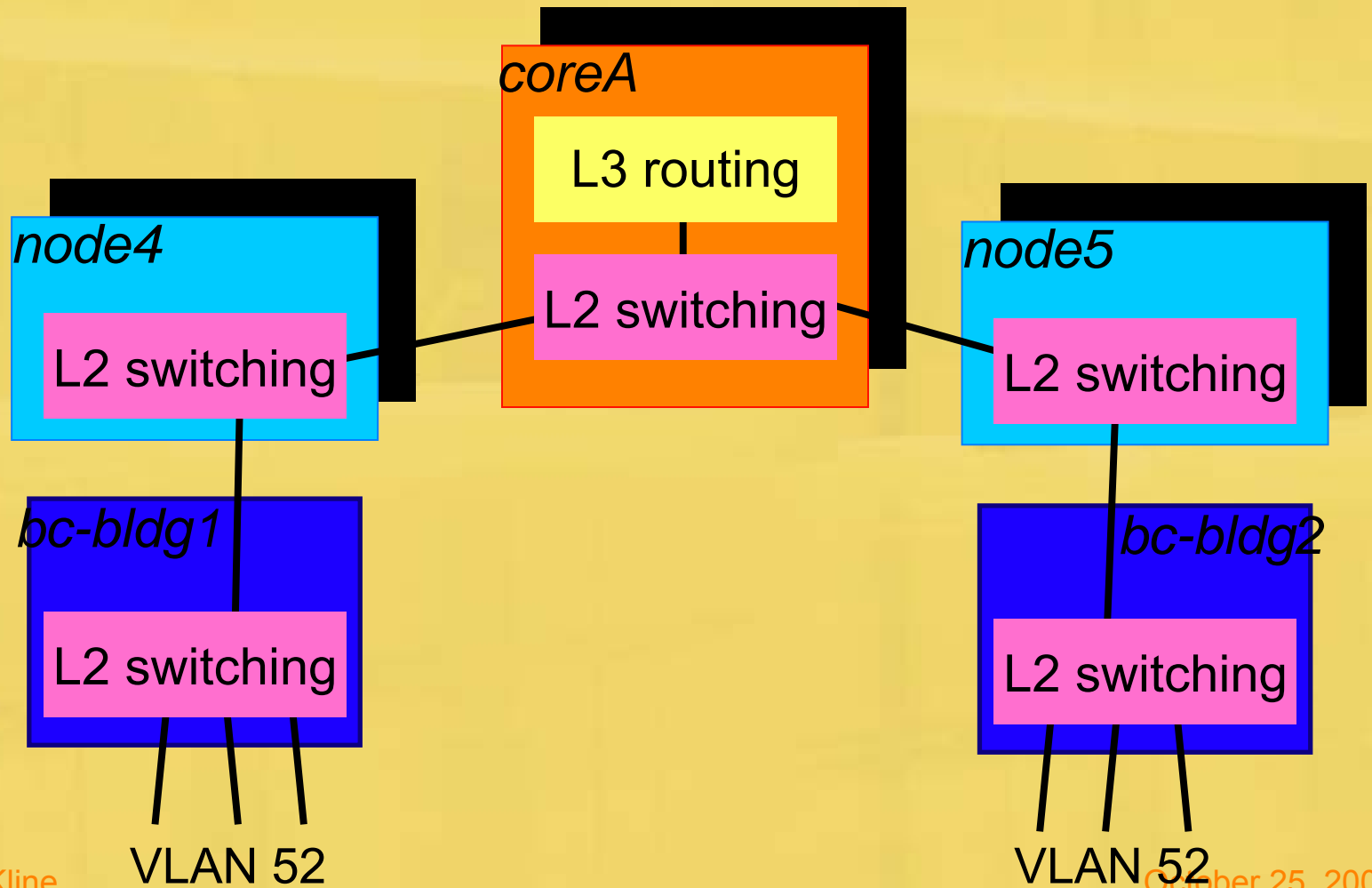


Details – What Happens Where

- A VLAN which spans multiple buildings, but remains within the same node, is routed in the node distribution device
- [next slide] A VLAN which spans multiple nodes must be routed in the core device (we want to avoid this)

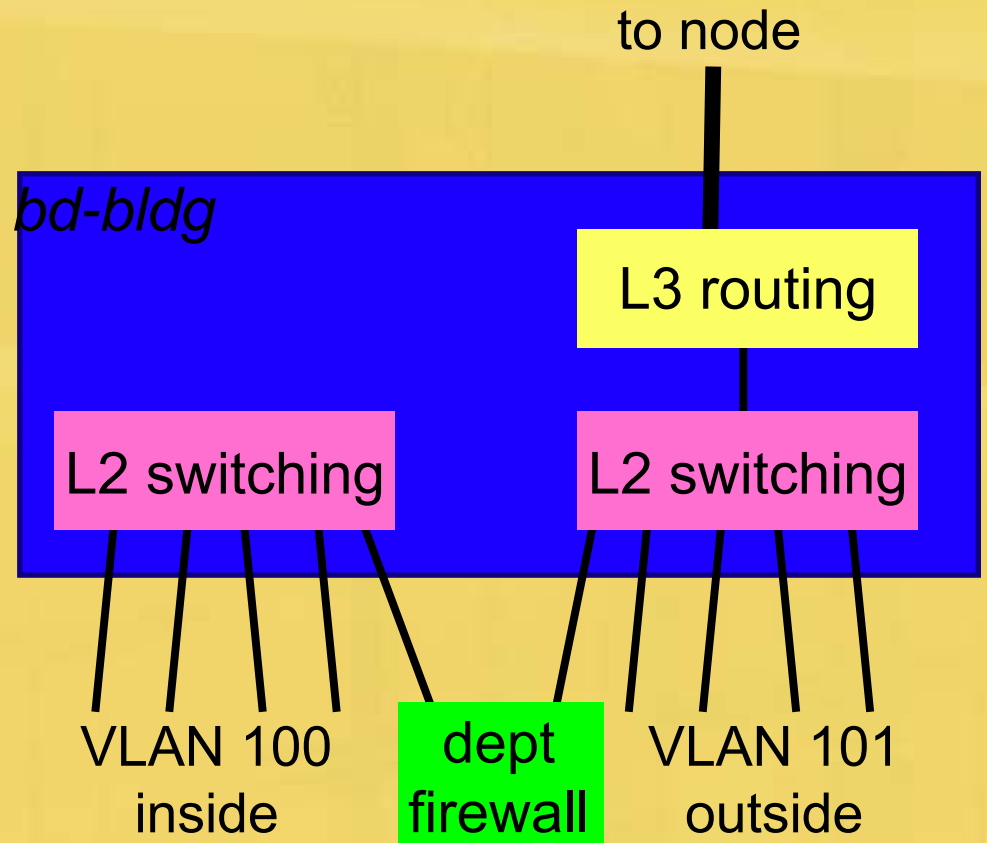


Details – What Happens Where



Details – What Happens Where

- The BC/BD device makes a natural and supportable place to connect a departmental firewall
- Requires a second VLAN which is not routed and does not leave the building
- The firewall then provides the “only way out”



Where Are We Now?

- NDO is already using the Building Reference Design for the network upgrade project
- All BC devices for now; BD functionality will require the new network core
- So far no deviations have been necessary



Where Are We Now?

- Core RFP complete except for Board of Trustees approval
- Projected install over winter break
- Projected turn-up and cutover begins en masse over spring break
- Conversion to BD switches and in-building routing will start in the spring and proceed en masse over next summer



Where Next?

- The new core equipment can route IPv6 – investigating whether to enable this day one
- 10-gigabit core links mean very high speed research applications can be supported
- MPLS, VLL, and L3VPN are being investigated as ways to provide advanced services and eliminate L2 switching and STP once and for all (requires board swapout)



Questions and Discussion

