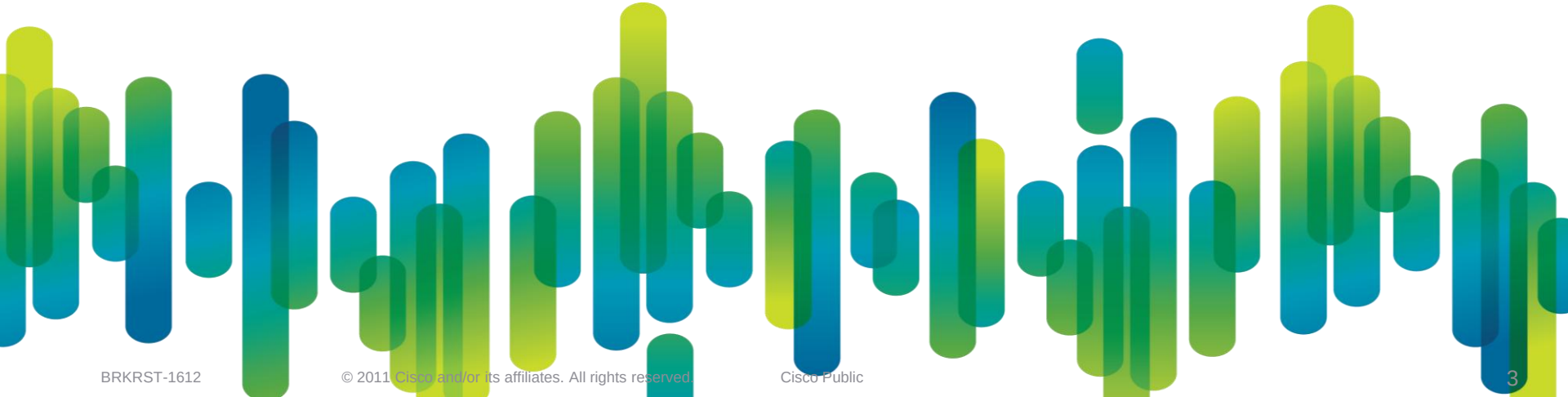
Cisco Switching Platforms – Deployment Case Studies

BRKRST-1612

Agenda

- Problem – which switch?
- Switch Roles
- Traditional Ethernet Switching functions
- Data Centre Switching Requirements
- Case Studies
  - –Recommendations
  - –New Campus
  - –New DC

Ethernet has been around since the early 80s. Why is there more ‘choice’ in switches than there ever has been?



Traditional Switch Functions

Ethernet Switch Functions

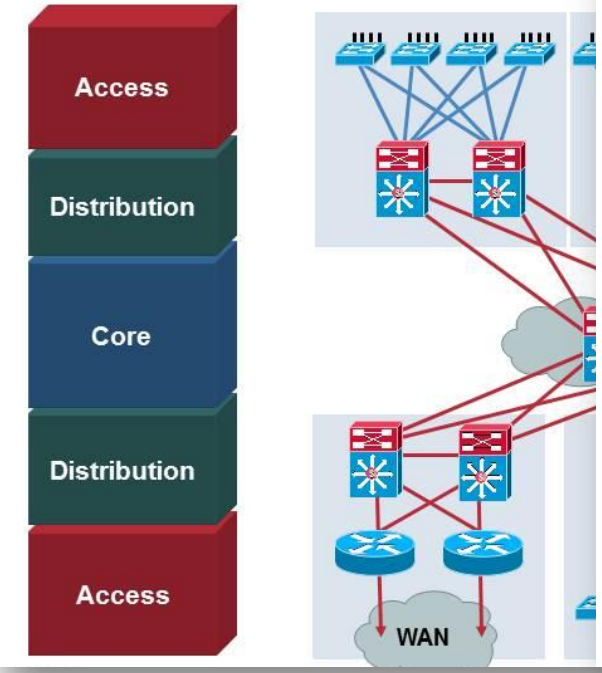
- Layer 2 including Spanning Tree
- Layer 3 - IP
- Network Services - extra services residing in switches such as firewalling or load-balancing or DHCP
- QoS and Multicast
- Power over Ethernet
- Etherchannel* & Stacking

Traditional Switch Choice

- How many ports and at what speed?
- Does it need to route?
- How much redundancy - power and supervisor?
- PoE? If so, how much?
- What about other network services?
- What ‘Role’ does the switch play?

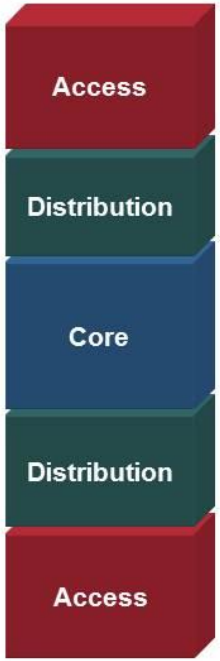
Traditional LAN Design

Hierarchical Campus
Building Blocks



Hierarchical Campus Design
Multilayer L2/L3 Building Blocks

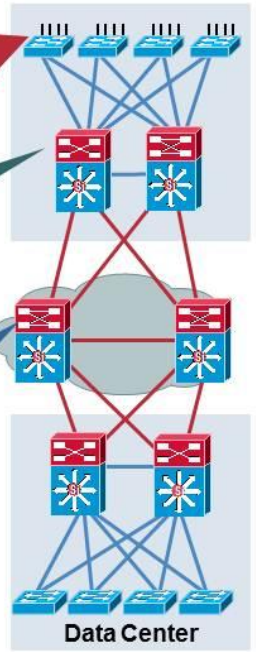
Cisco.com



- Network trust boundary
- Use Rapid PVST+ on L2 ports to prevent loops in the topology
- Use UDLD to protect against 1 way interface UP connections
- Avoid daisy chaining access switches
- Avoid asymmetric routing and unicast flooding, don't span VLANS across the access layer

- Aggregation and policy enforcement
- Use HSRP or GLBP for default gateway protection
- Use Rapid PVST+ if you MUST have L2 loops in your topology
- Keep your redundancy simple; deterministic behavior = understanding failure scenarios and why each link is needed

- Highly available and fast—always on
- Deploy QoS end-to-end: protect the good and punish the bad
- Equal cost core links provide for best convergence
- Optimize CEF for best utilization of redundant L3 paths

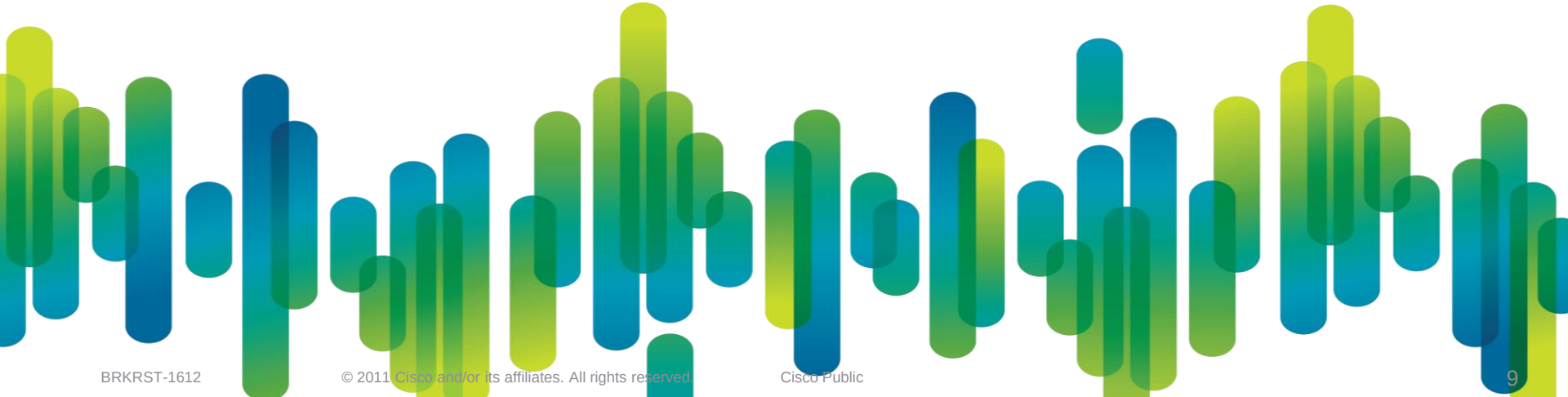


Newer switch requirements in the Campus

- MPLS
- 10GE
- Multi-Chassis Etherchannel (VSS)
- ‘Green’
- More tuning and automation (EEM, Auto…, Medianet, etc)

How much of this is driven by user or business requirement?

Data Centre Switching Requirements



Why is Data Centre a special case?

Business case justification

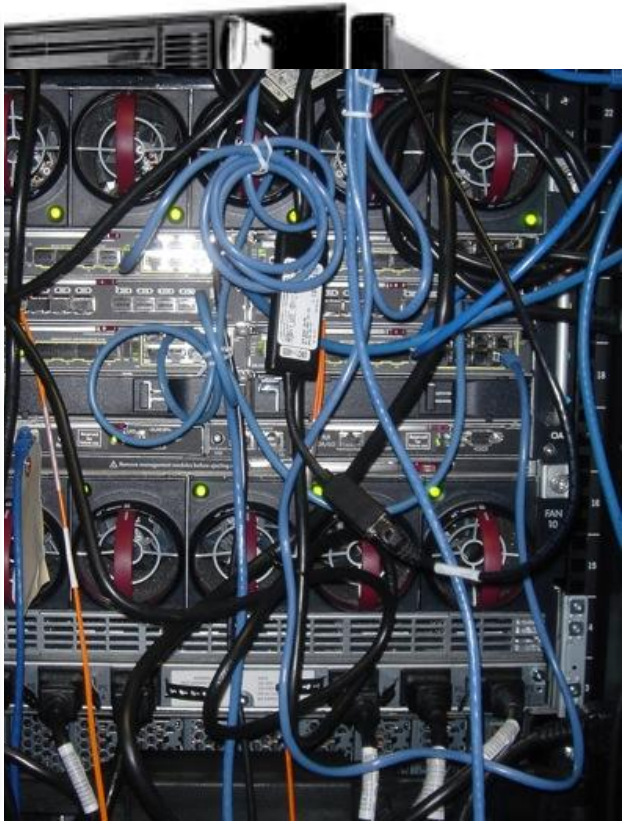
- As ‘servers’ have grown and evolved their networking requirements have changed – more rapidly than that of humans in an office environment
- Better ways have evolved
- What are some these ‘problems’ we need to fix?

Since 2004, the compound annual growth rate in workloads has been about 16 percent

If a company had 100 servers in 2004, in 2010 they would have had 243.

Importantly, each of those servers would have been connected to 4 networks on average

Server Evolution



Data Centre Switching Requirements

Server Networking Needs aka Problems for Networks

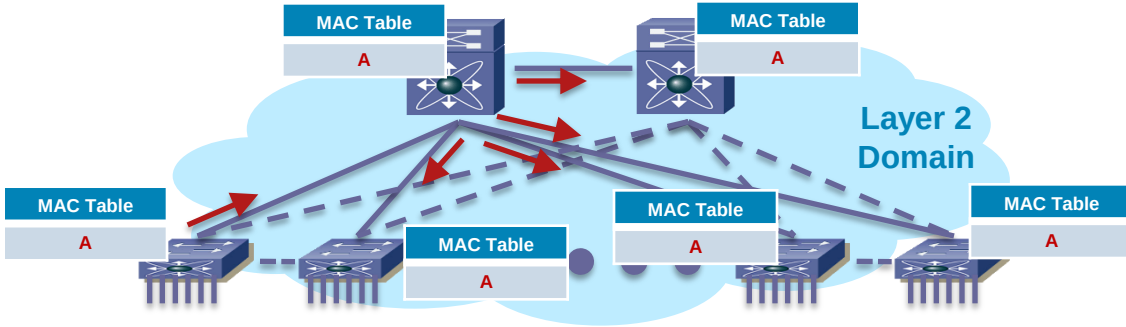
- Scale – bandwidth an ever-present issue
- Scale – port counts create further issues:
  - ‘Fabric Infrastructure’ costs:
    - lots of uplinks and downlinks especially since STP isn’t very efficient
    - lots of switches (bridges in STP)
    - lots of cables
- Server Virtualisation (VMWare) :
  - Virtual NICs
  - Machine mobility

Data Centre Switching Technologies

- Fabric Extender (FEX) - fixing a physical problem, cabling
- VirtualNetwork Link (VN-Link) N1000v - fixing a control problem caused by virtualisation
- DCB and FCOE - improving on a scale problem
- Virtual PortChannel (vPC) – improving efficiency (STP)
- FabricPath - improving efficiency and scale of LANs
- Overlay Transport Virtualisation (OTV) – bridging a L3 link

Are any of these useful in a Campus environment?

L2 MAC Address Scaling Issues



- Mac addresses facts:
  - –There are billions
  - –They have no location associated to them, no hierarchy
  - –They are not “registered” by the hosts to the network
- A routing table is impossible at Layer 2: **default** forwarding behaviour is **flooding**
- A filtering database is set up to limit flooding
- The whole mechanism is not scalable

L2 Requires a Tree

Branches of trees never interconnect (no loop)

- The Spanning Tree Protocol (STP) is typically used to build this tree
- Tree topology implies:
  - Wasted bandwidth -> over-subscription exacerbated
  - Sub-optimal paths
  - Conservative convergence -> failure catastrophic

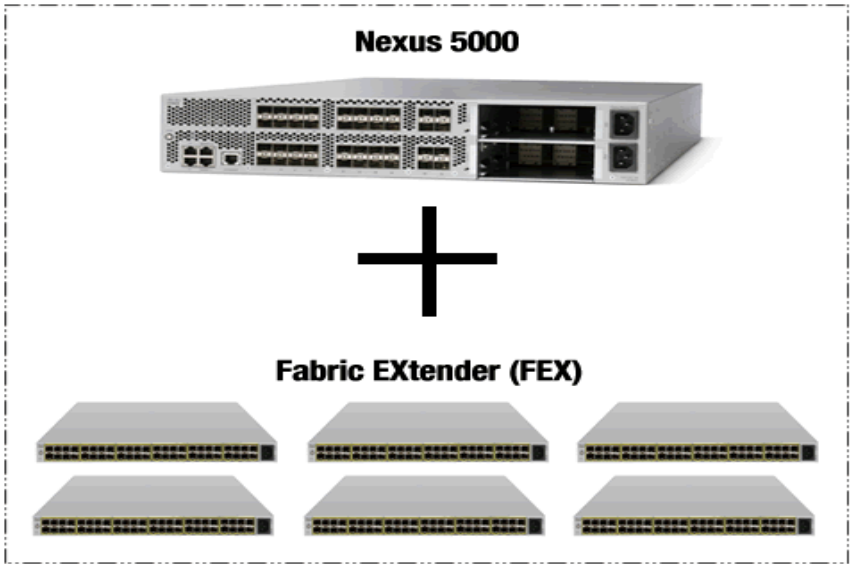
Cisco STP Implementation

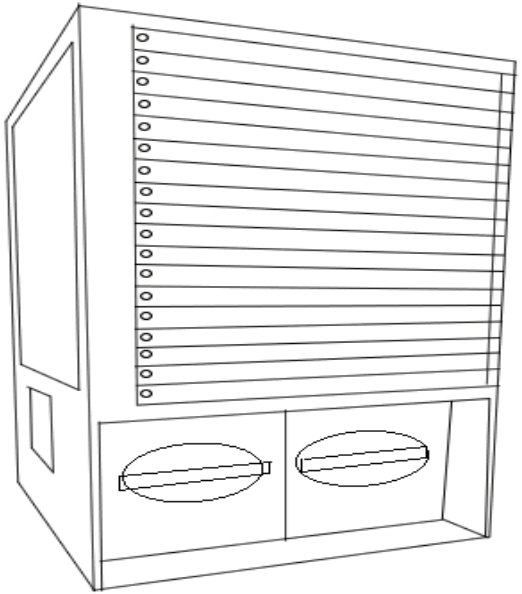
Feature Rich

	Stability	Speed	Policy Enforcement	Scale
MST		+		+
Rapid-PVST		+		-
Dispute	+	-		
Bridge Assurance	+	-		
Loopguard	+	-		
RootGuard	+	-	+	
BPDUGuard	+		+	
Global BPDU filter				+

Great features, but all working around problems

Fabric Extender (FEX)

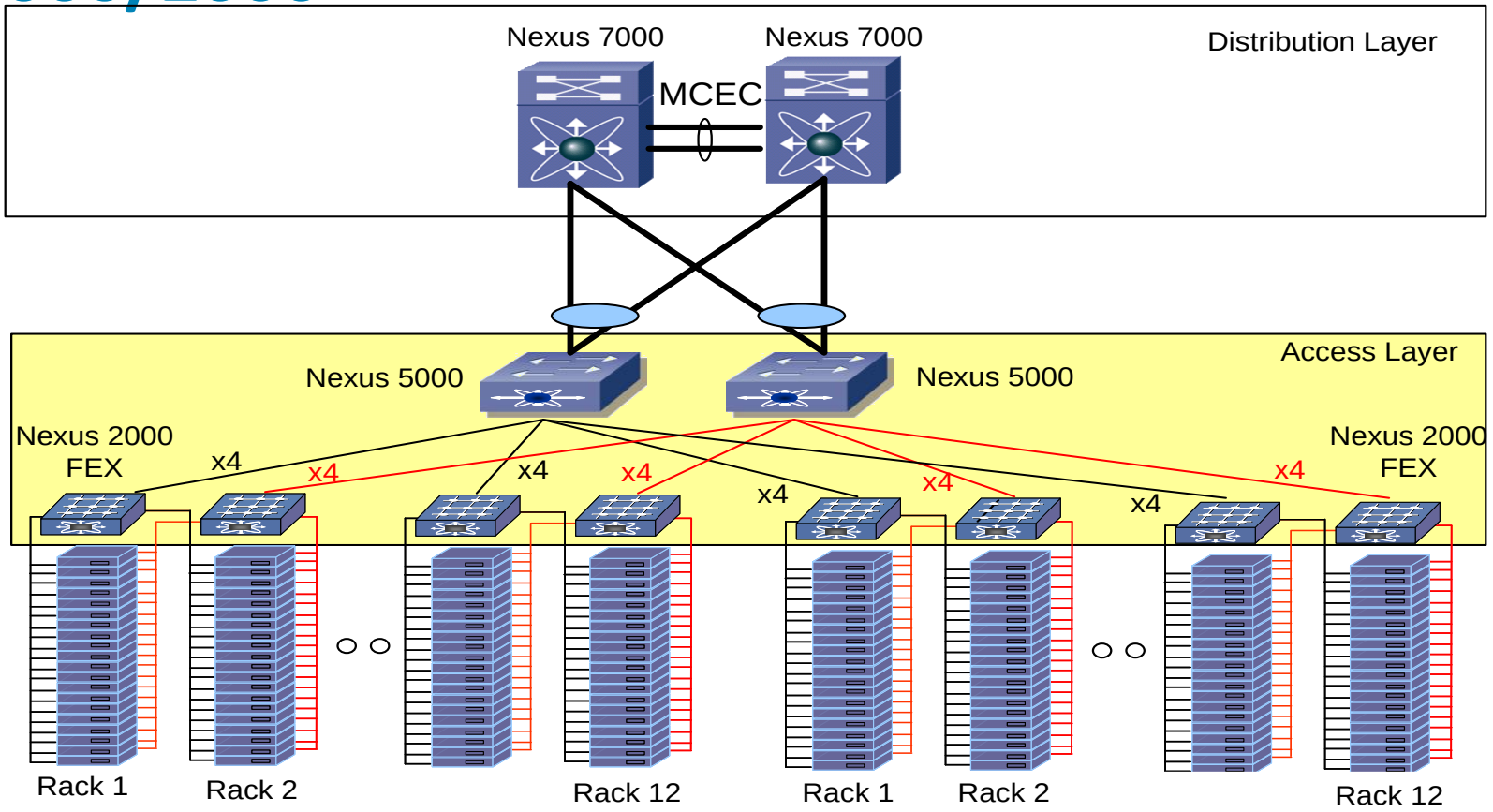




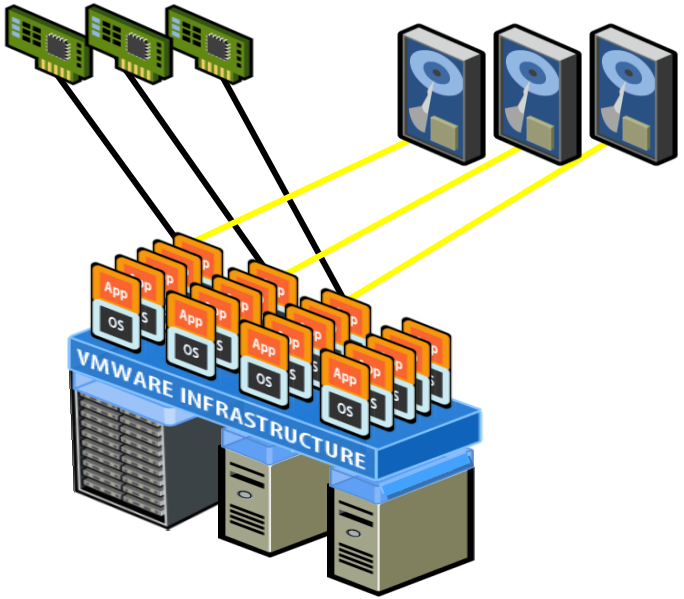
Nexus 5000 Fabric EXtender (FEX)

The Nexus 5000 Fabric Extender (FEX) acts as a remote line card (module) for the Nexus 5000, retaining all centralized management and configuration on the Nexus 5000, transforming it to a Virtualized Chassis

Top of Rack FEX Deployment with Nexus 5000/2000

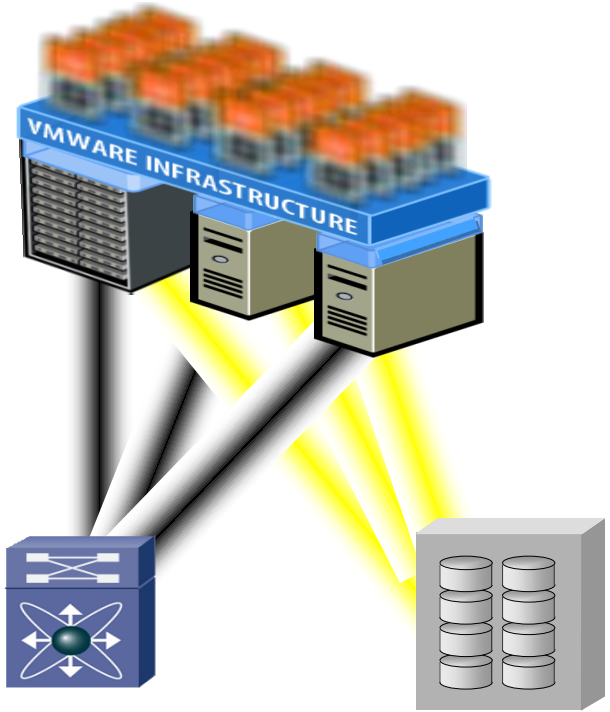


Transparency in the Eye of the Beholder



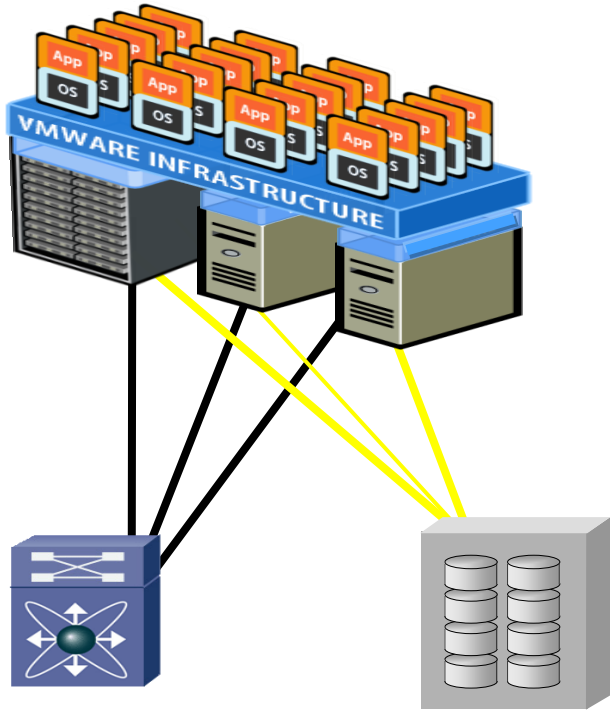
With server virtualisation, VMs have a transparent view of their resources…

Transparency in the Eye of the Beholder



…but its difficult to correlate network and storage back to virtual machines

Transparency in the Eye of the Beholder

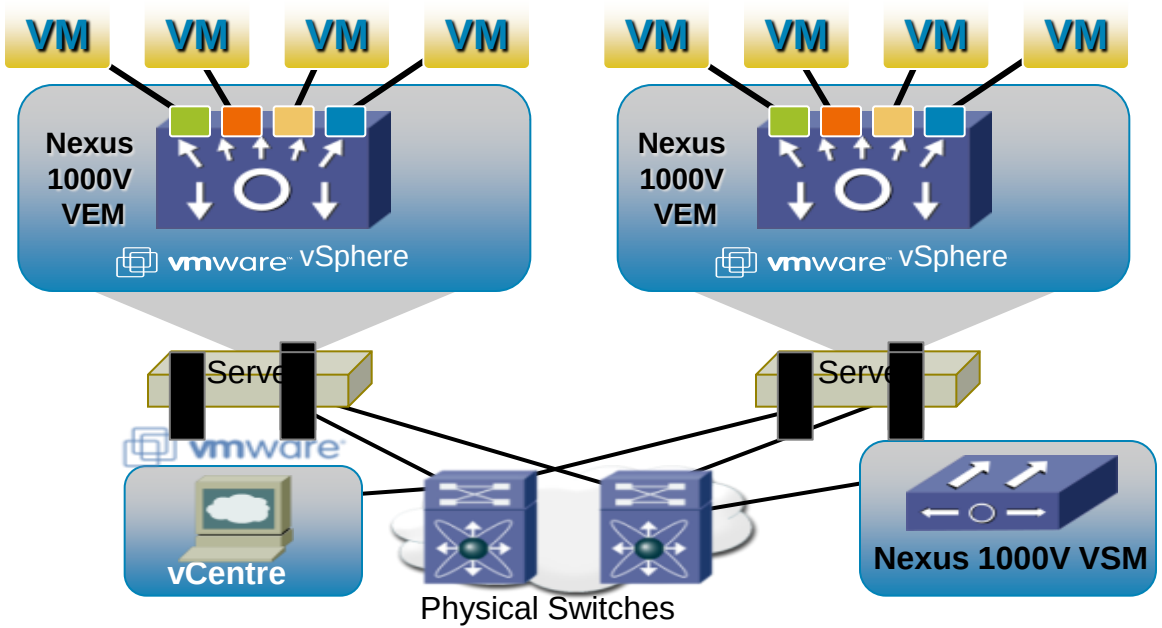


Scaling globally (across ESX hosts) depends on maintaining transparency while also providing operational consistency

Cisco Nexus 1000V Architecture

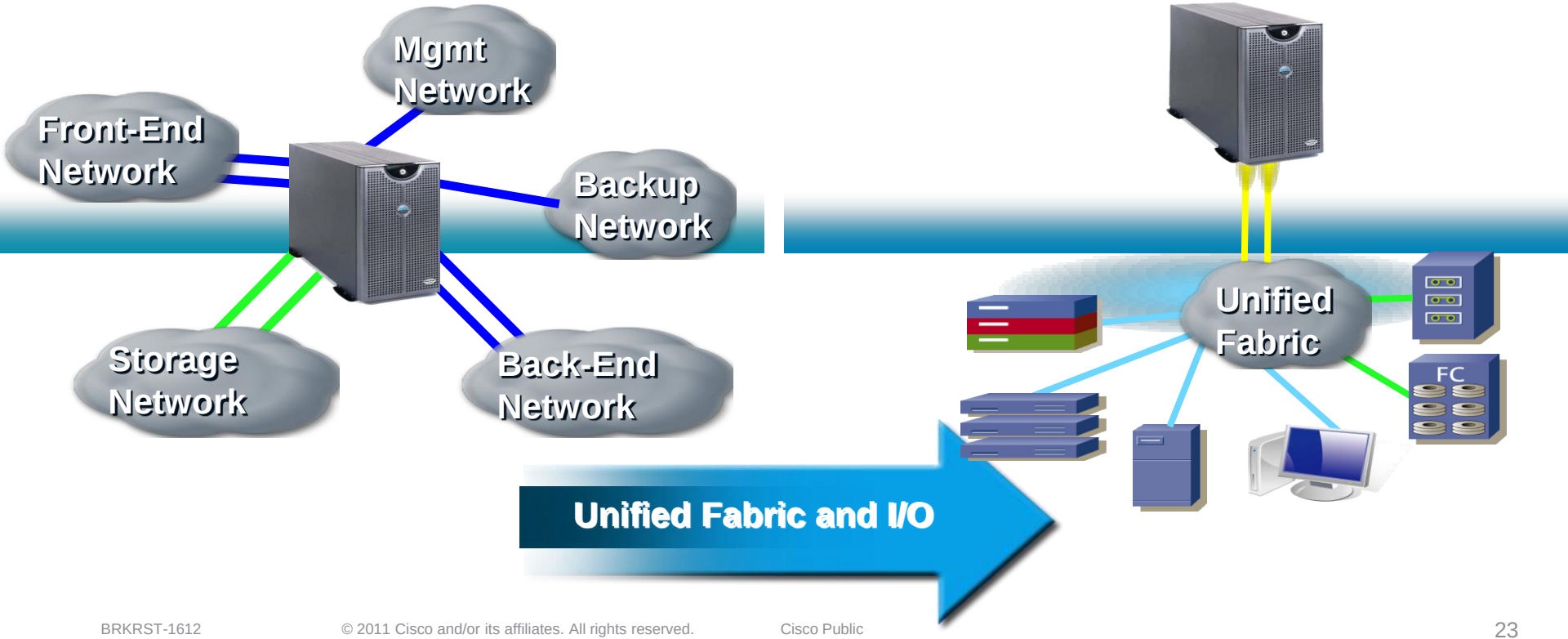
Installation

- ESX & ESXi
- VUM & Manual Installation
- VEM is installed/upgraded like any ESX patch



FCOE

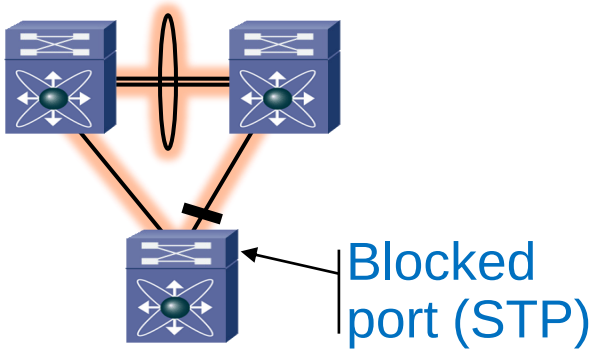
Server I/O – educe cables and interfaces



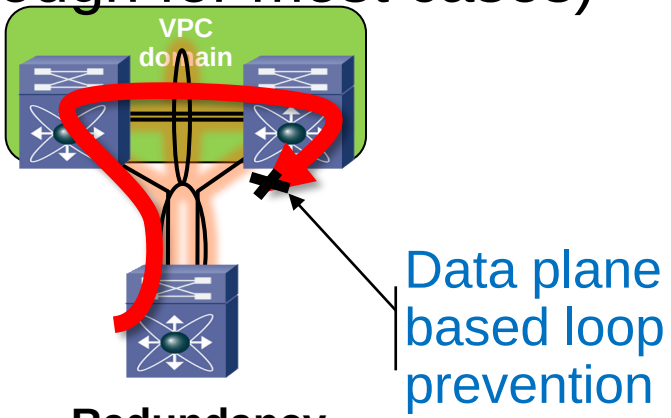
Virtual Port Channel (vPC)

To improve uplink efficiency

- Introduces some changes to the data plane
- Provides active/active redundancy
- Does not rely on STP (STP kept as safeguard)
- Limited to pair of switches (enough for most cases)



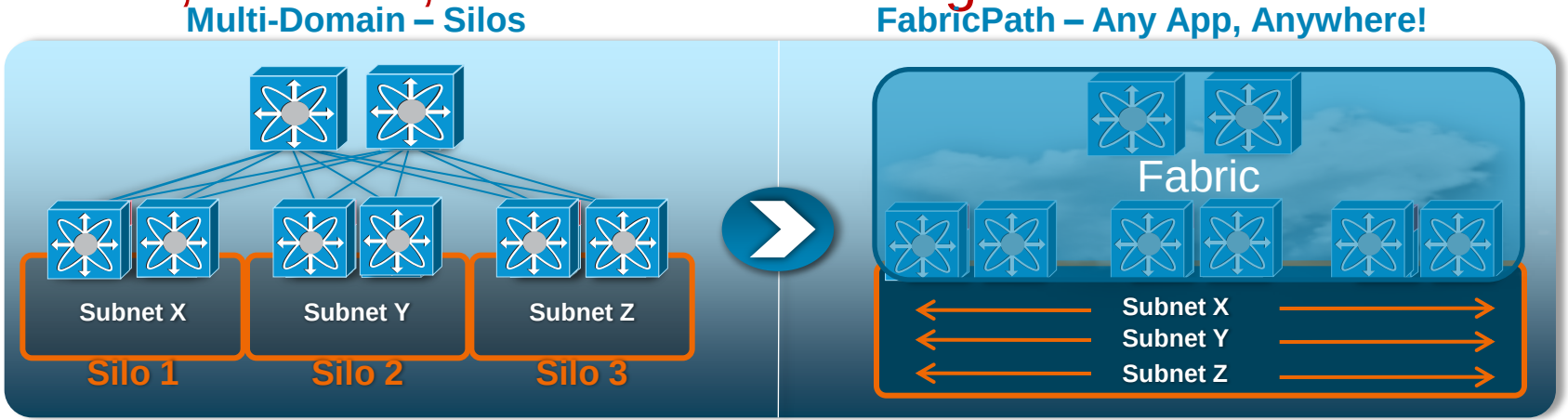
Redundancy handled by STP



Redundancy handled by vPC

FabricPath: Simple from the Outside

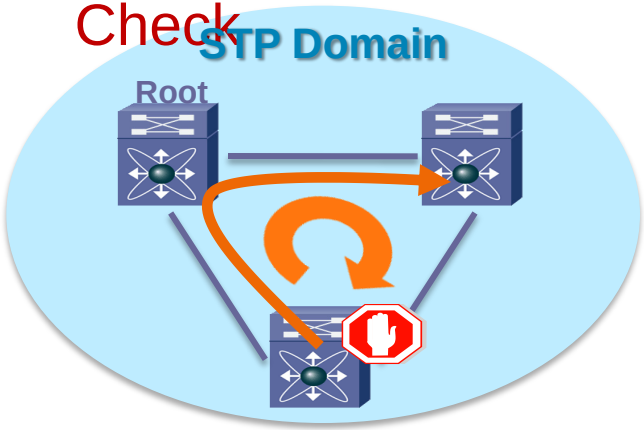
- Scalable, Uniform, Efficient Forwarding

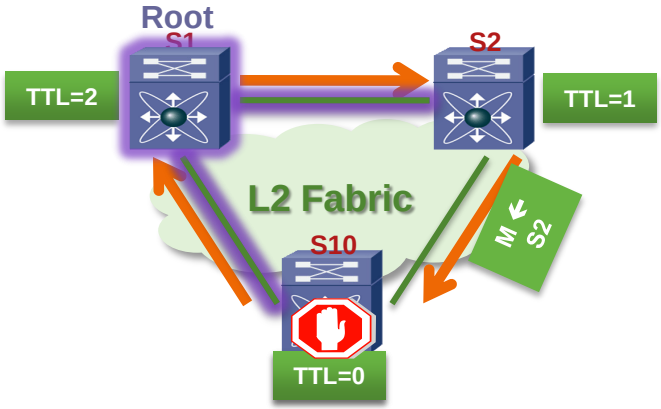


- FabricPath provides a Fabric that looks like a switch =>
  No silos, workload mobility and maximum flexibility

Loop Mitigation with FabricPath

- Time To Live (TTL) and Reverse Path Forwarding (RPF) Check



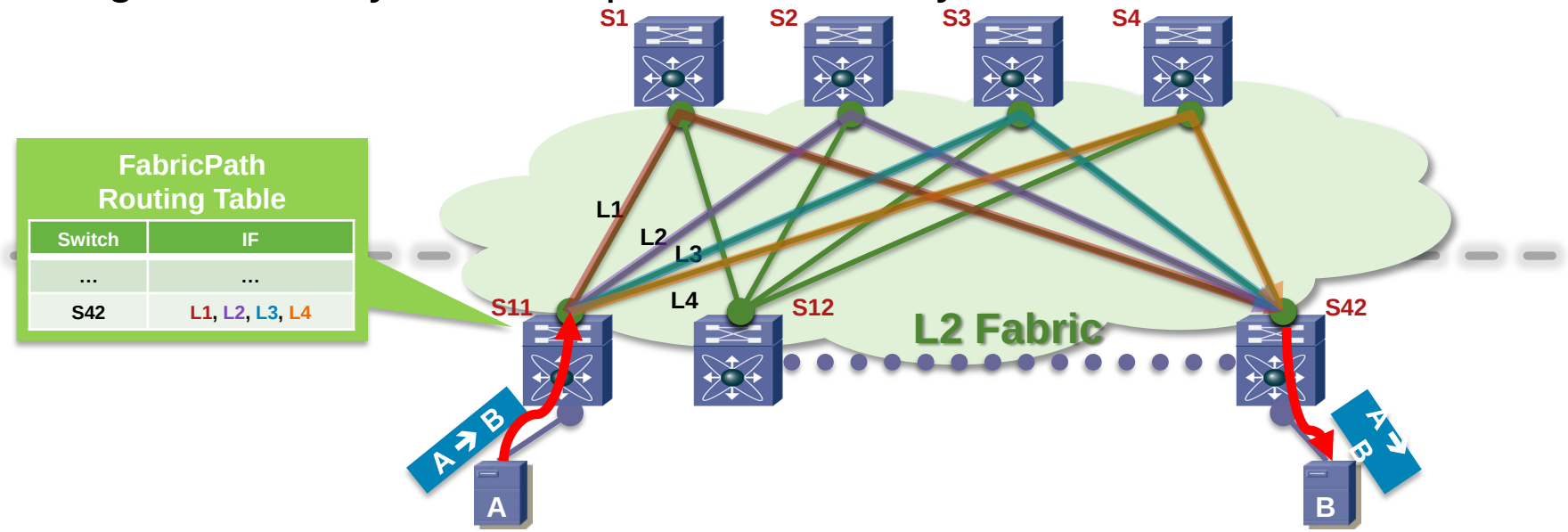


- Control protocol is the only mechanism preventing loops
- If STP fails -> loop
  - no backup mechanism in the data plane
  - Probable network-wide melt-down

- TTL in FabricPath header
- Decrement by 1 at each hop
- Frames with TTL =0 are discarded
- RPF check for multicast based on “tree” info

FabricPath is Efficient

- Shortest path, Multi-Pathing, High-availability
- Shortest path for low latency
- Up to 256 links active between any 2 nodes
- High availability with N+1 path redundancy



OTV – Overlay Transport Virtualisation

Intelligent LAN extensions over any transport

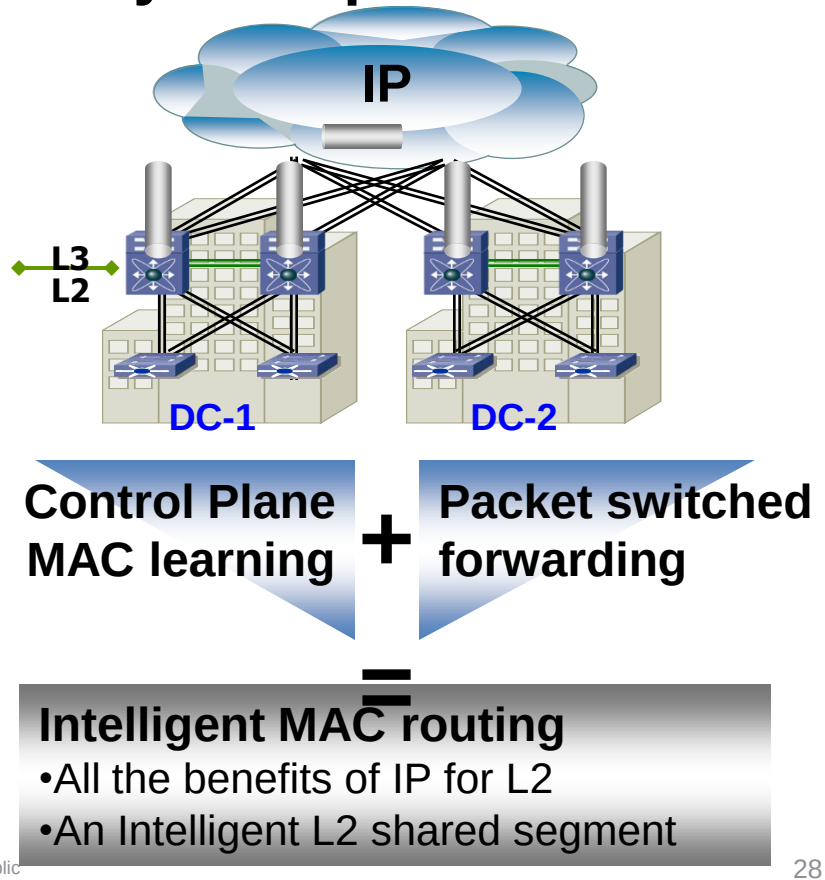
Zero Impact to existing network design

- **Failure isolation**
  - – Preserve L3 boundary failure containment with L2 protocols remain localised by site

Optimised Operations

Scalability

- Full BW utilisation & optimal traffic replication
- Stretch a VLAN between DCs in a protected way



Case Studies

Switching Platform Positioning

- General Position:
  - Users plug into Catalysts, new Servers plug into Nexus

- If a user PC plugs into the switch then a Catalyst is likely best fit

Cisco Catalyst

- If it's going into a new, large or virtual Data Centre then it is probably a Nexus

Cisco Nexus

Enterprise Core Positioning

The Core of the LAN is the main place of ‘overlap’ where both Nexus and Catalyst have fitting products.

- Core outside Data Centre
- PoE or WAN services
- VPLS/ Avanced MPLS

Decision Criteria

- Core inside Data Centre
- High 10GE density
- Data Centre Interconnect

- Cisco IOS consistency
- Two terabit scalability
- Service modules
- VSS

Cisco Catalyst 6500 Series

- NX-OS capabilities
- Highest scalability
- Resiliency (hitless ISSU)
- Virtual Device Contexts (VDCs)

Nexus 7000 Series

Switching Positioning

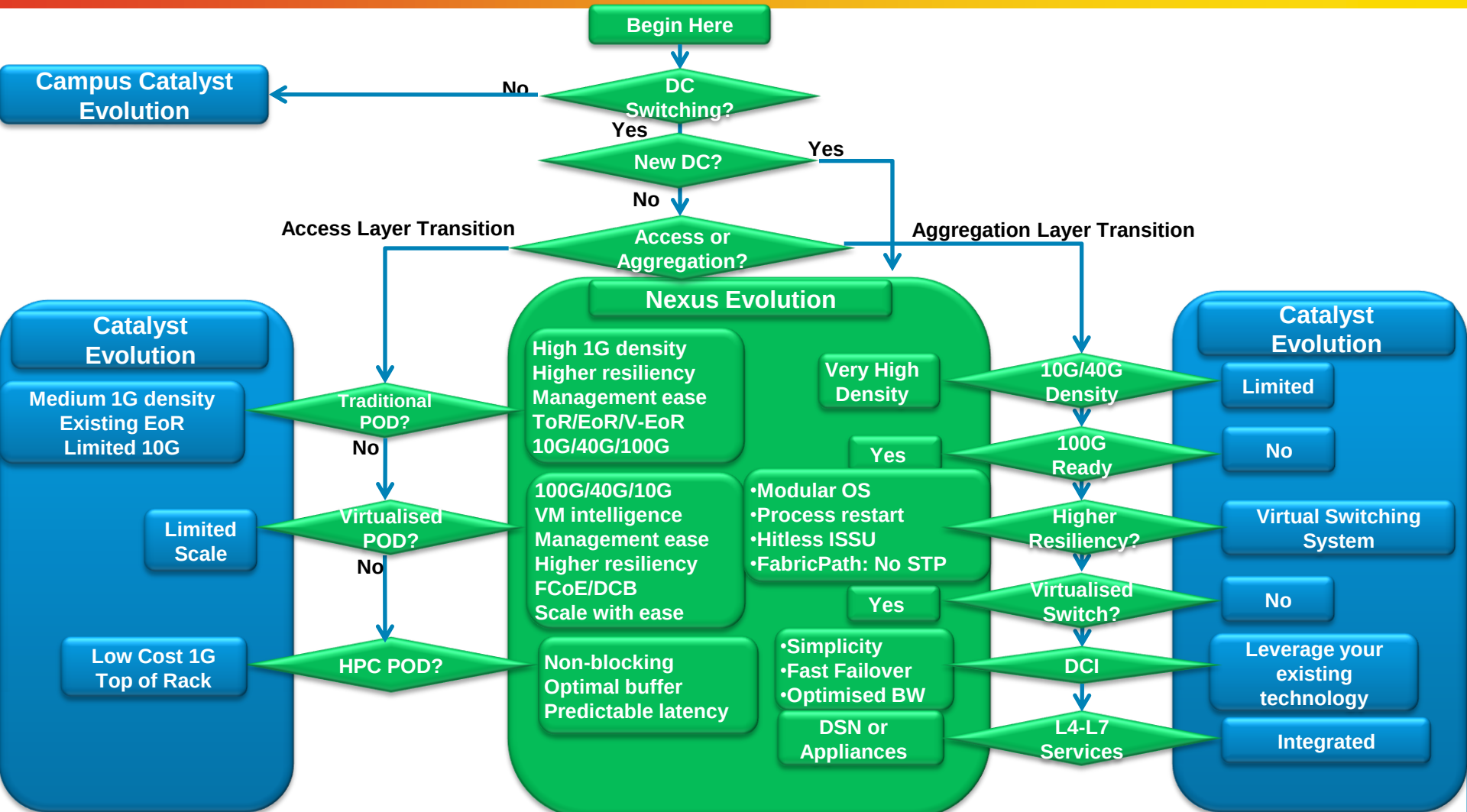
- Catalyst Innovations for Campus networks
- Nexus Innovations for the next generation Data Centre

- Energywise
- StackPower
- Medianet
- VSS

Cisco Catalyst

- Fabric Extender
- DCB/ FCOE
- Unified Ports
- VPC & OTV & FabricPath

Cisco Nexus


Begin Here
DC Switching?
No
Campus Catalyst Evolution
Yes
New DC?
Yes
No
Access Layer Transition
Access or Aggregation?
Aggregation Layer Transition
Catalyst Evolution
Medium 1G density
Existing EoR
Limited 10G
Traditional POD?
No
Limited Scale
Virtualised POD?
No
Low Cost 1G
Top of Rack
HPC POD?
Nexus Evolution
High 1G density
Higher resiliency
Management ease
ToR/EoR/V-EoR
10G/40G/100G
100G/40G/10G
VM intelligence
Management ease
Higher resiliency
FCoE/DCB
Scale with ease
Non-blocking
Optimal buffer
Predictable latency
Very High Density
Yes
•Modular OS
•Process restart
•Hitless ISSU
•FabricPath: No STP
Yes
•Simplicity
•Fast Failover
•Optimised BW
DSN or Appliances
10G/40G Density
100G Ready
Higher Resiliency?
Virtualised Switch?
DCI
L4-L7 Services
Catalyst Evolution
Limited
No
Virtual Switching System
No
Leverage your existing technology
Integrated

Case Study 1 – a new Campus for MediaInc

What would you use?

- 1600 users across 8 floors in a new ‘green’ building
- All areas have new GigE phones and WLAN coverage
- Single riser and Comms room per floor
- Small Computer Room on Ground floor
- Remote Data Centre connected over dual dark fibre
- Security and Loading Dock to connect also

Case Study 1 – a new Campus

Access Layer/ Wiring Closet

- Chassis or Stack both valid options
- Dual 10GE uplinks (power users)
- Redundant Power supplies (or PowerStack)

- Use Compact Switches for small spaces

Case Study 1 – a new Campus

Core & Distribution

- Able to collapse Distribution into Core?
- Multi-Chassis EtherChannel (vPC or VSS) would be ideal for a faster, simple topology
- High-availability a must and routing between sites

Case Study 2 – refitting a Data Centre Facility

What would you use?

- Business is virtualising computing onto 20 new blade chasis but support of existing environment must remain
- 100 racks of a mix of blade servers, 1RU servers and a couple of mainframes all connected to iLO, LAN(s) & SAN
- Existing switching is Cisco 6500 and MDS including heavy use of ACE and FWSM
- Business critical mainframes ideally on separate network
- Business Continuance planning calls for DR Data Centre in the future

Case Study 2 – a new Data Centre Facility

Aims

- Minimise cabling to reduce costs via 2 methods:
  - Localise cable runs using Top-of-Rack (ToR) switching
    - Utilise low-cost 10GE cabling options where possible (Twinax/CX1, FET, LRM)
  - Reduce server cabling via Unified Fabric/FCOE where possible
- Flat topology to support Vmotion and aid in future DC Interconnect
- Continue use of ACE modules and FWSMs

Case Study 2 – a new Data Centre Facility

Options - Access

- Virtual switching – leverage Nexus1000v on all ESXi hosts to gain granularity and consistency
- Connecting Blade Chassis – preferably Unified Fabric/ FCOE using converged switch module (UCS, Nexus 4001i)
  - Alternatively pass-thru modules to Nexus ToR switch can achieve same outcomes
- Where suitable, migrate existing cabling to ToR using FEX to distribution layer of Nexus 5500

Case Study 2 – a new Data Centre Facility

Options – Distribution

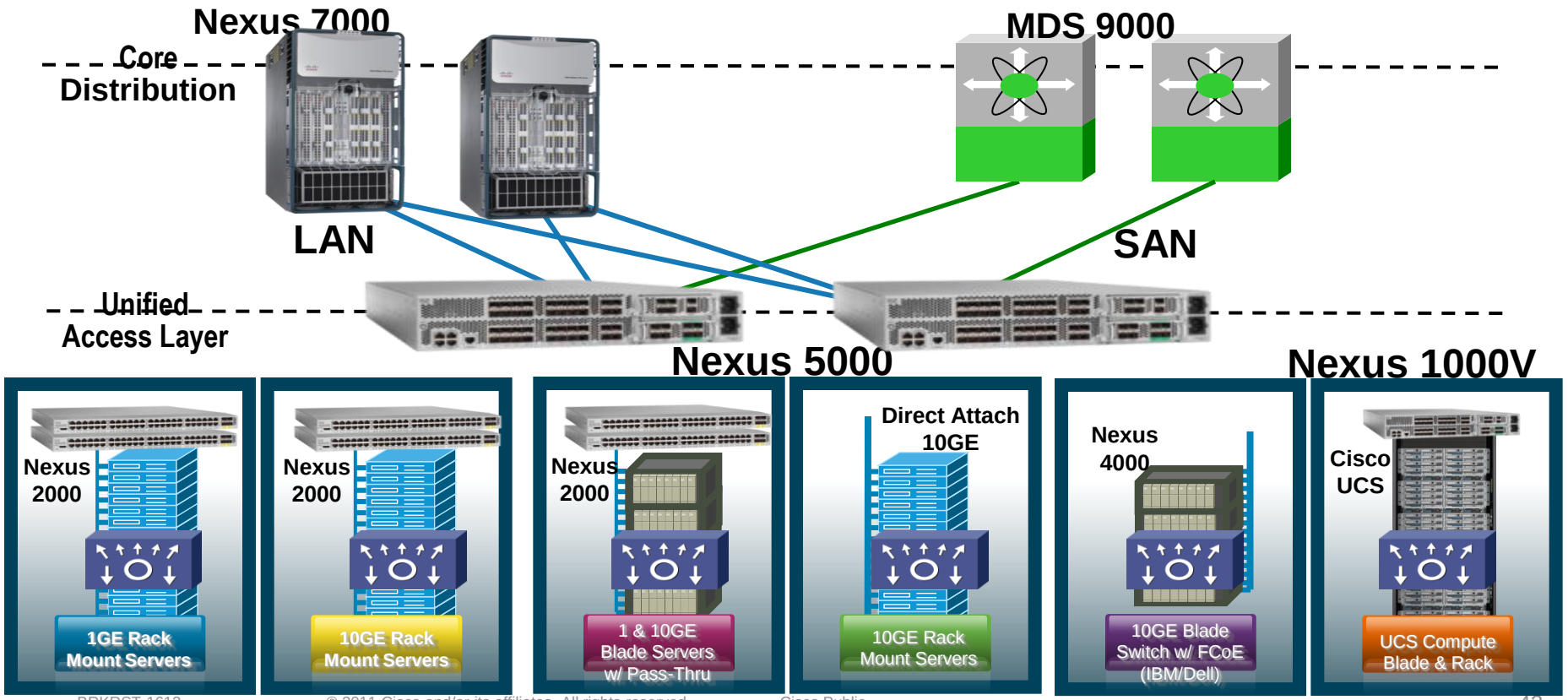
- Distribution provided by Nexus 5548s x 8 upstream from blade switching and FEXs*
- SAN connection to existing MDS from 5548s FC ports
- Use vPC in pairs for ‘STP-blocked’ free environment

Case Study 2 – a new Data Centre Facility

Options – Core

- 40G usable (non-blocked) bandwidth to each ‘Pod’ of 2 Nexus5548s
- Retain existing Catalyst6500s and move to ‘side’ for Services (ACE and FWSM)
- Nexus 7000 seems the logical choice for 10G port density and resilience along with OTV feature for future DC interconnect and VDC capability for segmenting network

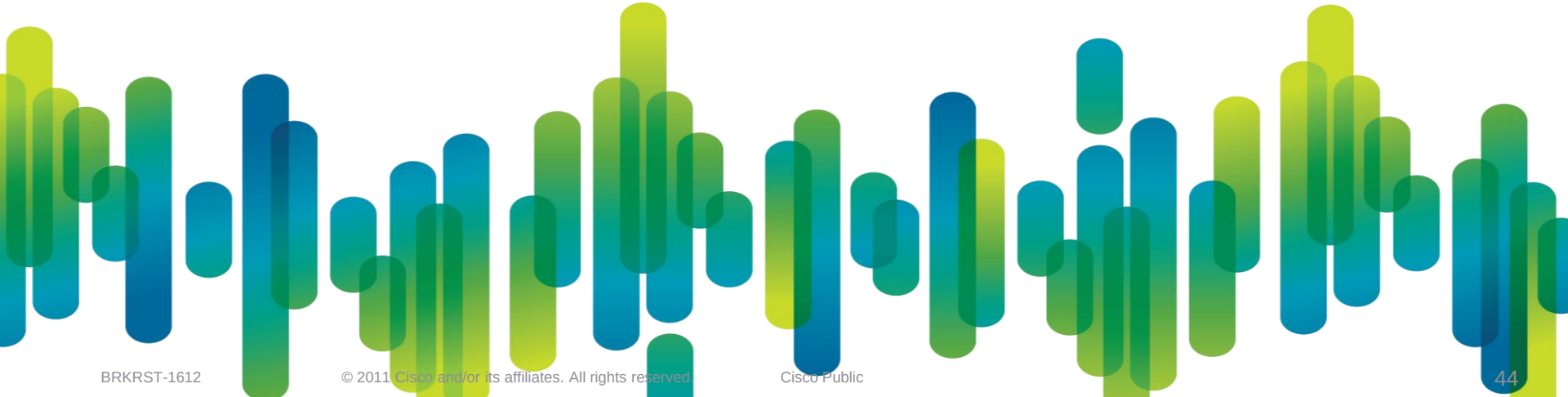
Data Centre Access options



Summary

- Traditional user switching continues to evolve with the Catalyst family
- New Data Centre switching technology is available on Nexus family
- Interoperation between the two segments is seamless
- …but it is still a case of matching business/network needs

Q & A



Complete Your Online Session Evaluation

Complete your session evaluation:

- Directly from your mobile device by visiting www.ciscoliveaustralia.com/mobile and login by entering your badge ID (located on the front of your badge)
- Visit one of the Cisco Live internet stations located throughout the venue
- Open a browser on your own computer to access the Cisco Live onsite portal


CISCO