Data Center 3.0
Technology Evolution
Session Goal

The focus of this seminar is on the latest technologies some of which can already be used in today’s deployments and some that will become available in the next 1 or 2 years. These include: virtual Port-Channeling, Fabric Extender, Layer 2 multi-pathing (read IETF TRILL), VN-TAG, FCoE, IO Virtualization.

Attaching Physical and Virtual Compute resources to the networking is described including blade switching, pass-thru, virtualized adapters and Hypervisor based switching. The Cisco compute platform, the UCS, is analyzed as part of the Blade Servers discussion in terms of network connectivity.
Understand Evolving Design Requirements
Data Center Inflection Points

First Inflection Point

- Mainframe
- Minicomputer
- Client Server

Second Inflection Point

- Distributed Computing
- Web
- Virtualization

The Evolving Data Centre Architecture
Dec 2011 Gartner Data Center Conference: CIO Feedback

1) Budget in 2011
   a. 33% Flat
   b. 42% Higher (greatest was up 5%, lowest was up 10% or more)
   c. 25% Lower

2) Biggest Challenge you Face in 2011
   a. 29% Power, Cooling and Energy Efficiency
   b. 16% Cloud Computing
   c. 14% Meeting Business Needs with IT
   d. 5% Dealing with limited budget

Technologies CIO’s selected as one of their top five priorities
   Cloud
   Virtualization
   Mobility
   Infrastructure
   IT Manager
The Evolving Data Centre Architecture

Data Center Evolution Path

- Efficient use of Compute, Storage and Network
- Flexibility
- Provisioning Speed
- Tighter integration between servers and the network
- Virtualized Workloads
- Network/Server demarcation moving inside of the server
Impact on the Data Center
Operations & Maintenance Now ~80% of IT Budgets and Growing

Virtualization makes things worse

Source: IDC
The Datacenter Today

Trusted
Control
Reliable
Secure
Cloud Computing?

- Trusted
- Control
- Reliable
- Secure

Flexible
Dynamic
On-demand
Efficient
The Goal:

- Trusted
- Control
- Reliable
- Secure

Flexible
Dynamic
On-demand
Efficient

Cloud Computing

Virtualization
Security
Applications
Information
Data Center 3.0
Business Driven, Fabric Architecture
The Evolving Data Centre Architecture
Data Center 2.0 (Pre-Virtualized)
The Evolving Data Centre Architecture
Data Center 2.0 (Physical Design == Logical Design)

- The IP portion of the Data Center Architecture has been based on the hierarchical switching design
  - Workload is localized to the Aggregation Block
  - Services localized to the applications running on the servers connected to the physical pod
  - Mobility often supported via a centralized cable plant

- Architecture is often based on optimized design for control plane stability within the network fabric

- **Goal #1:** Understand the constraints of the current approach (De-Couple the Elements of the Design)

- **Goal #2:** Understand the options we have to build a more efficient architecture (Re-assemble the elements into a more flexible design)
Cisco’s Holistic Fabric-Based Approach
High-Performance Infrastructure Delivering Architectural Flexibility
What is Cisco Unified Fabric?

**ATTRIBUTES**

- **CONVERGENCE**
  - Ethernet Network
  - Storage Network
  - Data Center
  - OS & Mgmt

- **SCALE**
  - Simplified Infrastructure
  - Intra Data Center
  - Inter Data Center
  - Multi-Tenant

- **INTELLIGENCE**
  - Simplified Operations
  - Inter Data Center
  - Multi-Tenant

**BENEFITS**

- Reduced Cost
- Integrated App Delivery
- Fluid VM Networking
- Secure VM Mobility

Network-Based Approach for Systems Excellence
Cisco Unified Fabric Technologies NX-OS
Continued Innovation for Differentiation

Unified Ports: Deployment Flexibility
DCB/FCoE: Consolidated I/O
VDC: Virtualizes the Switch
LISP: Scalability and Mobility
FabricPath: Architectural Flexibility
vPC: Active-Active Uplinks
FEX Architecture: Simplified Management with Scale
OTV: Workload Mobility
IO Accelerator: Replication and Backup
SME/DMM: Compliance and Workload Mobility
Building Next Generation Data Center Fabric
Evolving Data Centre Architecture
Focus of this Session

- The Tightly Coupled Workload Domain
- Layer 2 Scaling
  - FEX – Scale-Up
  - FabricPath – Scale-Out
- Common Storage Fabrics
  - SAN, NAS, iSCSI, CIFS
- Virtual Machine Connectivity
  - vSwitch, Nexus 1000v, VM-FEX
- Policy and Services in a vMotion environment
Evolution of the Data Center Architecture
Tightly Coupled Workload—Active/Active

- The domain of active workload migration (e.g. vMotion) is currently constrained by the latency requirements associated with storage synchronization.

- Tightly coupled workload domain has specific network, storage, virtualization and services requirements.
Evolving Data Centre Architecture
Tightly Coupled Workload (e.g., vMotion and Clustering)

- Hypervisor based server virtualization and the associated capabilities (vMotion, …) are changing multiple aspects of the Data Center design.
- How large do we need to scale Layer 2?
- Where does the storage fabric exist (NAS, SAN, …)?
- How much capacity does a server need?
- Where is the policy boundary (security, QoS, WAN acceleration, …)?
- Where and how do you connect the servers?
Evolving Data Centre Architecture
Design Factor #1 to Re-Visit

- As we move to new designs need to re-evaluate the L2/L3 scaling and design assumptions
- Need to consider VLAN Usage
  - Policy assignment (QoS, Security, Closed User Groups)
  - IP Address Management
  - Some factors are fixed (e.g. ARP load)
  - Some factors can be modified by altering VLAN/Subnet ratio
- Still need to consider L2/L3 Boundary Control Plane Scaling
  - ARP scaling (how many L2 adjacent devices)
  - FHRP, PIM, IGMP
  - STP logical port count (BPDUs generated per second)
- **Goal:** Evaluate which elements can change in your architecture
The traditional L2 vs. L3 debate has been based on a number of issues:
- Scalability
- Availability

Requirements for the High Availability design moving forward is a scalable, highly available switching fabric with the advantages of both L2 and L3.
Architecture Flexibility Through NX-OS

Spanning-Tree

- Single
- Up to 10 Tbps

vPC

- Dual
- Up to 20 Tbps

FabricPath

- 16 Way
- Up to 160 Tbps

Layer 2 Scalability

Infrastructure Virtualization and Capacity
Virtual Port Channel (vPC)

- vPC is a Port-channeling concept extending link aggregation to two separate physical switches
- Allows the creation of resilient L2 topologies based on Link Aggregation.
  - Eliminates the need for STP in the access-distribution
- Provides increased bandwidth
  - All links are actively forwarding
- vPC maintains independent control planes
- vPC switches are joined together to form a “domain”
Evolving Data Centre Architecture

vPC—“Scaling Up” the Network Pod

- Nexus designs currently leveraging vPC to increase capacity and increase scale of layer 2 fabrics
  - Removing physical loops out of the layer 2 topology
  - Reducing the STP state on the access and aggregation layer

- Scaling the aggregate Bandwidth
  - Nexus 5000/5500 when combined with F1 line cards on Nexus 7000 can support port channels of up to 32 x 10G interfaces = 320 Gbps between access and aggregation
  - Nexus 5500/2000 virtualized access switch can support MCEC based port channels of up to 16 x 10G links = 160Gbps between server and virtualized access switch
Cisco FabricPath
NX-OS Innovation, Enhancing Layer 2 with Layer 3

More than 500 FabricPath customers on Cisco Nexus® 7000

Layer 2 Strengths
- Simple configuration
- Flexible provisioning
- Low cost

Layer 3 Strengths
- All links active
- Fast convergence
- Highly scalable

Simplicity  Flexibility  Bandwidth  Availability  Cost

Now Shipping on Nexus 5500 from Q4CY11

Enhanced Fabric Scale with FEX
Evolving Data Centre Architecture
FabricPath—“Scaling Out” the Workload/Fabric

- Scaling the Fabric to a much larger geographic domain
- Extending the tightly coupled workload domain

- Connect a group of switches using an arbitrary topology
- With a simple CLI, aggregate them into a Fabric:

  ```
  N7K(config)# interface ethernet 1/1
  N7K(config-if)# switchport mode fabricpath
  ```

- An open protocol based on L3 technology provides Fabric-wide intelligence and ties the elements together
Evolving Data Centre Architecture
FabricPath—The “BUS” for the Compute Pods

- FabricPath provides an extensible transport fabric that aids in decoupling the building blocks of the DC
- Provides a scalable communication “BUS” between design elements (Layer 3 Routing, Policy Services, Compute Pod)
Evolving Data Centre Architecture
Design Factor #2 to Re-Visit

- How is “striping of workload” across the physical Data Center accomplished (Rack, Grouping of Racks, Blade Chassis, …)?
- How is the increase in percentage of devices attached to SAN/NAS impacting the aggregated I/O and cabling density per “compute unit”?

**Goal:** Define the unit of Compute I/O and how it is managed (how does the cabling connect the compute to the network and fabric)
Nexus Virtualized Access Switch

Rack Mount

- Rack Mount Scale Unit
  - Rack Mount Servers are still predominantly 1Gbps (2 to 6 NIC’s common)
  - Migration to 2 x 10G NIC/CNA’s is occurring (10G cable requirements)
  - Server layout will primarily be driven by power density
  - Look to replicate Units of workload [Compute + I/O] to aid in providing the flexibility in mapping the physical layout to the logical design

20 to 30 Servers per Rack

10 to 15 Servers per Rack

8 to 12 Servers per Rack
Evolving Data Centre Architecture
FEX—Scaling Up and Distributing the Workload Domain

- De-Coupling of the Layer 1 and Layer 2/3 Topologies
- If the workload is defined to be local to the switch can provide an alternative method to geographic distribution
- Provides a per ‘rack’ based granularity view of the logical workload domain
Fabric Extender Technology Simplifies Data Center Infrastructure
Decouple Scale and Complexity ➔ Lower Costs

FABRIC EXTENDER TECHNOLOGY: Available in UCS and Nexus
Evolving Data Centre Architecture
Starting Point—Workload Virtualization

Partitioning
- Physical devices partitioned into virtual devices

Clustering
- Applications distributed across multiple servers
Architectural Goal is balanced between the need to scale workload and provide availability and manageability

Flexibility in the architecture of the Data Center Fabric

Where are the Tightly Coupled vs. Loosely Coupled Domains of Workload

---

Distribution of the Workload (Striping servers and VM’s amongst the rack, along the row, between the rows, … )

Scaling of the Workload Capabilities (Number of Servers and VM’s)
Server Virtualization Issues

1. vMotion moves VMs across physical ports—the network policy must follow

2. Impossible to view or apply network policy to locally switched traffic

3. Need shared nomenclature and collaboration for security policies between network and server admin
Unified VM Awareness

**Nexus 1000V**
- Software Hypervisor Switching
- Tagless (802.1Q)
- Feature set Flexibility

**Cisco VM-FEX**
- External Hardware Switching
- Tag-based (Pre-standard 802.1BR)
- Performance Consolidation

---

**Policy-Based VM Connectivity**

**Mobility of Network and Security Properties**

**Non-Disruptive Operational Model**
Data Centre Architecture Evolution

Embedded 802.1Q Bridging — Nexus 1000v

- **Embedded Virtual Bridge - Nexus 1000V**
  802.1q standards based bridge
  Performs packet forwarding and applies advanced networking features
  Policy Based port profile applies port security, VLAN, and ACLs, policy maps for QoS treatment for all systems traffic including VM traffic, Console & Vmotion/Vmkernel
  Generic adapter on generic x86 server

- **Standard 802.1q based upstream switch**
  Leveraging standard switch to switch links (QoS, trunking, channeling, ..)
  Policy on upstream switch looks like standard ‘aggregation’ configuration
Cisco Nexus 1000V
Components

Virtual Ethernet Module (VEM)
- Replaces VMware’s virtual switch
- Data Plane for the N1KV
- Enables advanced switching capability on the hypervisor
- Provides each VM with dedicated "switch ports"

Virtual Supervisor Module (VSM)
- CLI interface into the Nexus 1000V
- Control Plane for the N1KV
- Leverages NX-OS
- Controls multiple VEMs as a single network device

Cisco VEM

Cisco VSMs

vCenter Server
Evolving Data Centre Architecture
Design Factor #4 to Re-Visit—Where Is the Edge?

- Hypervisor based compute virtualization moves the edge of the Fabric
- PCI-E bus and storage and network connectivity resources are virtualized
  - vSwitch
  - VMFS (VMWare)
  - NPV (provides FC SAN virtualization)
- With a shift in the edge of the fabric comes a change in the operational practices and fabric design requirements
Evolving Data Centre Architecture
Design Factor #4 to Re-Visit - Where Is the Edge?

- Unification of LAN and Storage fabrics
  FCoE, NFS, iSCSI, CIFS all provide the capability to ‘share’ the fabric

- 10G adapters provide rational to share IP storage and IP transactional

- Converged Network Adapters (CNAs) virtualize the PCI-E bus resources
  Two PCIe devices are seen by the OS/Hypervisor
  Ethernet NIC
  Fibre Channel HBA

- The edge of the SAN/Storage and LAN fabrics are merging
Evolving Data Centre Architecture
Design Factor #4 to Re-Visit - Where Is the Edge?

- The introduction of SR-IOV, Cisco Adapter-FEX and developing IEEE standards further extend the virtualization of the adapter and the first physical device port

- A single component is seen as many PCIe addresses on the bus and thus the OS talks to multiple devices but there is only one physical NIC, one physical wire and physical switch

- Extending the virtual port down from the physical fabric directly to the VM’s vNIC
Adapter FEX and VM-FEX
Expanding Cisco Fabric Extender Offerings

Benefits:

• Single point of management
• Increased 10 G bandwidth utilization – less power, fewer adaptors and cables with Adapter FEX
• Dynamic network & security policy mobility during VM migration with VM-FEX

Adapter FEX splits a physical NIC into multiple logical NICs

VM-FEX extends Adapter FEX technology to virtual machine

• IEEE 802.1BR
Scaling the Fabric: Physical to Virtual
Single Point of Policy and Management at Scale

**Physical**
Up to 1152 host ports

- Nexus 2248TP
- Nexus 2232PP
- Nexus 2232TM

**Logical**
Adapter FEX
Up to 1024 logical instances

- Logical NIC

**Virtual**
VM-FEX
Up to 4096 VMs

- VMs

Fabric Extensibility with Simplified Management
Evolving Data Centre Architecture
Design Factor #5 to Re-Visit—Where Are the Services?

- In the non-virtualized model services are inserted into the Data Path at ‘choke points’

- Logical Topology matches the Physical

- Virtualized workload may require a re-evaluation of where the services are applied and how they are scaled

- Virtualized Services associated with the Virtual Machine (Nexus 1000v & vPath)
What Is vPATH?
Services Interception in Nexus 1000v

- Intelligence build into Virtual Ethernet Module (VEM) of Cisco Nexus 1000V virtual switch (version 1.4 and above);

- vPATH has the following main functions:
  1. Intelligent Traffic interception for Virtual Service Nodes (VSN): vWAAS & VSG;
  2. Offload the processing of Pass-through traffic (from vWAAS, for instance);
  3. ARP based health check;

- vPATH is Multitenant Aware

- Leveraging vPATH can enhance the service performance by moving the processing to hypervisor;
Scalable Pod Deployment with VXLAN within a Data Center

Logical Network Spanning Across Layer 3

Utilize All Links in Port Channel w/ UDP

Add More Pods to Scale
Cisco Nexus 1000 Portfolio

Virtual Appliance
- Virtual ASA
- vWAAS
- VSG
- VSM

Nexus 1010
- Primary
- Secondary
- Virtual Supervisor Module (VSM)
- Network Analysis Module (NAM)
- Virtual Security Gateway (VSG)
- Data Center Network Manager (DCNM)*

Virtual Blades
- Virtual Supervisor Module (VSM)
- Network Analysis Module (NAM)
- Virtual Security Gateway (VSG)
- Data Center Network Manager (DCNM)*

VSM: Virtual Supervisor Module
VEM: Virtual Ethernet Module
vPath: Virtual Service Data-path
VXLAN: Scalable Segmentation
VSG: Virtual Security Gateway
vWAAS: Virtual WAAS
Virtual ASA: Tenant-edge security

vPath
- Service Binding (Traffic Steering)
- Fast-Path Offload

VXLAN
- 16M address space for LAN segments
- Network Virtualization (Mac-over-UDP)

* Est. 4Q CY11 **With Microsoft Window 8 in 2012
Evolution of the Data Center Architecture
Loosely Coupled Workload—Burst and Disaster Recovery

- Burst workload (adding temporary processing capacity) and Disaster Recovery leverage out of region facilities
- Loosely coupled workload domain has a different set of network, storage, virtualization and services requirements
Data Center Interconnect
OTV Layer 2 Mobility

- Ethernet traffic between sites is encapsulated in IP: “MAC in IP”
- Dynamic encapsulation based on MAC routing table
- No Pseudo-Wire or Tunnel state maintained

Communication between MAC1 (site 1) and MAC2 (site 2)
LISP Next Gen Routing Architecture
Locator-ID Separation Protocol

- Decouples the *Identity* of a resource (IP address) from its *Location*
- Keeps ID-to-Location mapping in a Directory
- IP addresses are resolved to a location by consulting the directory
  Similar to DNS name-to-IP resolution, but this is IP-to-location resolution
- The directory is a distributed hierarchical database
- Identifiers can move in the ‘Cloud’
  - Mobility for Virtual Machines
  - Scalability
  - IETF Draft Standard

Please see BRKVIR-2931 “End-to-End Data Center Virtualization” for more information on DCI
Bringing It All together: NX-OS
Any Application. Any Location. Any Scale.

1. Compute Fabric: Converged, Scalable and Stateless
   - UCS, FabricPath, FCOE, Unified Ports

2. Extensible Fabric
   - OTV, LISP

3. VM-Aware Fabric
   - VM-FEX, VSG

4. Policy Based Management
   - Service/Port/Security Profiles and Open API’s

Workload Mobility
Raising The Bar With New Cisco Unified Fabric Innovations

Cisco NX-OS: One OS from the Hypervisor to the Data Center Core

SCALABILITY

SAN
- MDS 9500
- MDS 9200
- MDS 9100

LAN
- Nexus 3000
- Nexus 2000
- Nexus 1010
- Nexus 1000V

LAN/SAN
- Nexus 4000
- Nexus 5000
- Nexus 7000

CONVERGENCE

VM-AWARE NETWORKING

10 GbE SWITCHING

FABRIC EXTENSIBILITY

CLOUD MOBILITY
Conclusion

- You should now have a thorough understanding of the Nexus Data Center features that enable a more flexible switching fabric supporting faster provisioning models, virtual machine mobility while maintaining the security isolation and availability you have to come to rely on in your current DC 2.0 designs

- Any questions?
Complete Your Session Evaluation

- Please give us your feedback!!
  Complete the evaluation form you were given when you entered the room
- This is session 3.1 (Data Center & Virtualization)

Don’t forget to complete the overall event evaluation form included in your registration kit

YOUR FEEDBACK IS VERY IMPORTANT FOR US!!! THANKS