Data Center Fabric Evolution: FabricPath, 40G/100G, and Beyond

Marty Ma
Lead Technical Solution Architect
APJC Data Center and Virtualization Team
Agenda

- Trends in the Data Center
- Transform Network into Fabric
- Start the Journey with FabricPath
- Evolving Data Center Fabric
- Unification of Fabric and DC
- Key Takeaway
Trends in the Data Center
Transform the Network – Why Now?

Mega Trends Causing Ripples

Video
Mobility/BYOD
Security

Data Center Consolidation
Virtualization & Cloud
Business Continuity Disaster Recovery

The IT Challenge: “Is My Network Ready?”

CAPACITY
“Do I have the right performance to scale?”

COMPLEXITY
“How do I simplify deployments?”

COST
“How can I be operationally efficient?”
# Infrastructure Matters More Than Ever

<table>
<thead>
<tr>
<th>CLOUD REQUIREMENTS</th>
<th>IDEAL INFRASTRUCTURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elastic Resource Allocation</td>
<td>• Rapid scalability of virtual and physical</td>
</tr>
<tr>
<td></td>
<td>• Automated deployment</td>
</tr>
<tr>
<td>Infrastructure as a Service</td>
<td>• Integrated compute, network and storage resources</td>
</tr>
<tr>
<td>Low Cost Operations</td>
<td>• Automation; Simplified management</td>
</tr>
<tr>
<td></td>
<td>• Lower service provisioning cost</td>
</tr>
<tr>
<td>Multi-tenancy</td>
<td>• Secure separation at scale</td>
</tr>
<tr>
<td></td>
<td>• Integrated services per tenant</td>
</tr>
<tr>
<td>Self Service Consumption Model</td>
<td>• Physical resources deployed through API’s and policies</td>
</tr>
</tbody>
</table>

- Infrastructure Matters More Than Ever

- Multi-tenancy
- Low Cost Operations
- Infrastructure as a Service
- Elastic Resource Allocation
- Self Service Consumption Model
# Data Center Requirements

| Scale | Datacenters come in different sizes  
|       | Large number of servers  
|       | Significant VM scale  
|       | Multi-tenancy  
|       | High bisectional bandwidth within DC |
| Service | Network Virtualization  
|         | Connectivity and optimal forwarding at Layer-2 and Layer-3  
|         | Deterministic bandwidth, latency  
|         | Robust network availability and redundancy  
|         | Seamless integration with WAN, DCI |
| Flexibility | Add network capacity and load incrementally  
|             | Workload and VM mobility  
|             | Variety of server, access connectivity options, multi-homing  
|             | Uniform access to resources (storage, L4-L7 services)  
|             | Per-VM policies for traffic redirection |
| Manageability | Network configuration and operations at scale  
|                | Simplified network, service provisioning for tenants  
|                | Ease of data collection and troubleshooting  
|                | Support for OAM and proactive monitoring  
|                | Ability to extend and customize manageability |
Transform Network into Fabric
Data Center Operators Need to Broaden Mobility Domain with Network Intelligence
Cisco Innovations Offer Distinctive Solutions

Rack-Wide VM Mobility

DC-Wide VM Mobility

- Network protocols enable broader VM Mobility
- Implementable on Virtual and Physical
- Examples: FabricPath/TRILL, VXLAN, LISP
Advantages of Layer-2
“Plug-and-Play” and Mobility

- Practically “plug-and-play” – No user configuration is required to build forwarding database, unlike routing
- It makes it simple to support teaming or Layer-2 multicast for clusters
- Easy to segment traffic with VLANs
- Very fast movement of end station addresses (ability to update MAC address tables after a vMotion-type event)
Disadvantages of Layer-2
Availability and Scaling

- MAC address consumption
- BPDU generation is CPU intensive with increasing number of VLANs
- VLAN sprawl causes flooding and broadcasts to propagate even where they are not needed
- Half of the links in the topology are blocking
- Misconfigurations can cause Layer 2 loops which may make switches unmanageable
Advantages and Drawbacks of Layer 3

- Layer-3 topologies alleviate the consumption of forwarding tables via route summarization
- Layer 3 routed topologies provide for a degree of fault isolation
- “Routed Access” designs provide the logical extension of the design philosophy
- “Scaling Up” of the Access Switch via such mechanism as the FEX provide some degree of edge workload mobility
- Layer-2 domain extension of some form is required for most workload mobility requirements

Workload Domain for most Hypervisor and Clustering based solutions is restricted by the Layer 3 boundary
Data Centre Architecture - Trends

Virtualized Data Center
- SP and Enterprise
- Hypervisor Virtualization
- Shared infrastructure Heterogenous
- 1G Edge moving to 10G
- Nexus 1000v, 2000, 5500, 6000, 7000 & UCS

Warehouse Scale
- Layer 3 Edge (iBGP, ISIS)
- 1000’s of racks
- Homogeneous Environment
- No Hypervisor virtualization
- 1G edge moving to 10G
- Nexus 2000, 3548, 5500, 6000, 7000 & UCS

HPC/GRID
- Layer 3 & Layer 2
- No Virtualization
- iWARP & RCoE
- Nexus 2000, 3548, 5500, 6000, 7000 & UCS
- 10G moving to 40G

High Frequency Trading
- Layer 3 & Multicast
- No Virtualization
- Limited Physical Scale
- Nexus 3548 & UCS
- 10G edge moving to 40G
The traditional Layer-2 vs. Layer-3 debate has been based on a number of issues:

- Scalability, Topology Restrictions, Link Blocking
- Availability, Stability

The requirement for Data Center designs moving forward is a scalable, highly available switching fabric with the advantages of **both** Layer-2 and Layer-3
Evolution of Data Center

**Distributed**
- Manual Provisioning
- Limited scaling
- Rack-wide VM mobility

**Fabric Based**
- Policy-based Provisioning
- Scale Physical & Virtual/Cloud
- DC-wide/Cross-DC VM Mobility

**Application Driven**
- Service-centric Provisioning
- Flexible – Anywhere, Anytime
- Cross-cloud VM Mobility

- Monitoring Apps
- Provisioning Apps
- Networking Apps
- End-User Apps

Programmable
Provisionable

Integrated Fabric & Cloud
World of Many Clouds
Start the Journey with FabricPath
Cisco FabricPath
NX-OS Innovation Enhancing Layer-2 with Layer-3

FabricPath brings Layer 3 routing benefits to flexible Layer 2 bridged Ethernet networks

Switching
- Easy Configuration
- Plug & Play
- Provisioning Flexibility

Routing
- Multi-pathing (ECMP)
- Fast Convergence
- Highly Scalable
## FabricPath vs. TRILL

<table>
<thead>
<tr>
<th>Feature</th>
<th>FabricPath</th>
<th>TRILL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame Routing</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>vPC+(Active/Active Host connectivity)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>FHRP Active/Active</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Multiple Topologies</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Conversational Learning</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Inter-switch links</td>
<td>Point-to-point only</td>
<td>Point-to-point or Shared</td>
</tr>
</tbody>
</table>

- N7K/N6K/N5K hardware support TRILL
- Cisco will push FabricPath specific enhancements to TRILL
- TRILL standard is no where near completed and deployable (Base protocol published in July 2011, while there are still 30+ draft in progress to address potential deployment issues)
FabricPath and Design Flexibility
Reducing the traditional constraints of STP-centric models

FabricPath provides an extensible transport fabric that aids in decoupling the building blocks of the DC and reassembling them in different ways.

Provides a flexible, scalable communications “bus” between design elements (Layer 3 Routing functions, data center services, compute pods).
Spine-Leaf Design
Changes to the Approach to Structured High Availability
Evolution of FabricPath Architecture

Anycast HSRP Gateway

Today (vPC+ with HSRP)
- A pair of FP node running vPC+ support HSRP Active/Active
- Limit to 2 FP nodes as gateway for a given VLAN
- Can be placed at spine or have dedicated leaf to provide routing

FP Anycast  HSRP Gateway
- More than two  FP nodes as default gateway
- N-way ECMP for Inter-VLAN traffic
- Can be placed at spine or routed leaf
- No need for vPC+ related configuration
Evolution of FabricPath Architecture
Distributed Routing Increases Scalability (2HCY13)

FP Anycast HSRP Gateway
- All HSRP nodes still need to maintain ARP entries for all hosts/VMs
- Scalability of fabric is limited by the ARP table size
- All HSRP nodes need to process ARP for all hosts

FP with Distributed Routing
- All leaf nodes as gateway for all VLANs
- Great VM mobility with any subnet anywhere
- Each leaf node only need to process ARP for locally attached hosts/VMs
- No ARP or unknown unicast flooding within FabricPath network
Evolution of FabricPath Architecture
Multi-Topology

- Topology consists of FP nodes and interface
- Each VLAN is mapped to one topology
- Traffic localization. Same VLAN is used in multiple places
- Traffic isolation for security reason
- Traffic engineering or steering
Evolution of FabricPath Architecture
Multi-tenant with FabricPath Segment ID (2HCY13)

- 24-bit segment ID is embedded in FabricPath frame. Extend “VLAN” space to 16 million
- Provide great scale to support large number of tenants in the fabric

Ingress leaf switch maps local VLAN to 24-bit segment ID

Traffic from VM with local VLAN

Egress leaf switch maps segment ID back to local VLAN

L3 Cloud
Evolving Data Center Fabric
Introduce Cisco Data Center Fabric Architecture
Leading The Data Center Flexibility And Innovation With Nexus

Simplified Management
Extensible Network Topologies
Distributed Control Plane
Integrated L2/L3 Across All Nodes

Enhanced Forwarding

Massive Scale On Demand
Secure Multi-Tenancy
Automation and Simplicity

Network Based Next-Generation Cloud Services
Cisco’s Fabric Approach Differentiation
Enabling Business Transformation

Massive Scale On Demand
- Fine grained redundancy
- Non-blocking Fabric 240Tb+, Scale of 50k+ ports
- Integrated L2/L3 services at the leaf
- Optimal traffic forwarding

Secure Multi-Tenancy
- Large scale multi-tenancy – 64K+ segments
- VM aware network – Seamless workload mobility
- Dynamic service and policy provisioning

Automation and Simplicity
- Zero touch deployment
- Automated network configuration & operations at scale
- Single point of management
- Open APIs – Seamless mgmt applications integration

Delivering To Your Data Center Needs
DC Networking Strategy

CPU
Memory
NIC
Comm BUS

File IO
Block IO
DAS, NAS, SAN

Encap/Decap
NV Gateway/Router
Entropy
Single Control Plane
1st Hop Protocol

Virtual Switch
Hypervisor Agnostic
vService Integration
Service Path

SinglePoint of MGMT
Visibility
Programmability
APIs: north-south
Agents/Controllers

Fundamental Services: bridging, routing, Security
Advanced Services: Proxy GW, Rack failure domain
Fabric Attributes: Scale, Latency, Non-blocking, Resilient
Fabric Topologies: Flexible, adjustable per workload
Custom Silicon Investment: matching requirements & TTM
Addressing The Fabric Strategy

Extensible & Flexible Topologies
- Classic, Clos, Rings
- Overlays
- 40G & 100G ready

- High Performance
- Highly available
- Low Latency
- Scalable

Simplified and Programmable
- Programmatic APIs
- Unified Fabric Management
- Scripting Capabilities
Need a 10G or 40G Fabric?

Most Important Factors…

- **10G Flows or 40G flows**
  - 10G Fabric: support only 10G flows
  - 40G Fabric: supports both 10G and 40G flows
  - Mixed Fabric: supports only 10G flows

- **Switching mode**
  - Store & Forward or Cut-Through Switching
  - Effect on Latency and Jitter

- **Spine Factors**
  - Fine vs Coarse Grain redundancy
  - Buffer Elasticity

- **Scale**
  - Vertical scale (bigger box) or horizontal scale (more boxes)
Data Center Architecture Evolution

Overlay Networking: applicable to any fabric

L2 Fabrics

STP ➔ VPC ➔ FP ➔ Trill NG

L3/L2 Fabrics

FP+L3 ➔ FP+ Distributed routing ➔ Future

L3 Fabrics

OSPF – BGP – ISIS ➔ Higher Scale + Features
Overlay Networking Strategy

Fabric, NIC and Host Integrated Functions
- Encap/Decap
- Gateway
- Router
- Entropy
- SegmentID Mapping

Normalization: Technology Convergence
- Various Hypervisors
- Different Encapsulation Types: Agnostic
- Various Control Plane Protocols: common Prot
- Mobility over L2 and L3 Fabrics
- Integration of Physical and virtual
- Avoid Mcast and flooding
- > 4k Segments
- Provide visibility and scale

Nexus Modular
- Encap/Decap
- Entropy
- Gateway Router
- Service Path

Nexus ToR
- Encap/Decap
- Entropy
- 1st Hop Protocol
- Gateway & Router
- Service Path

Nexus 1KV
- Encap/Decap
- Gateway
- 1st Hop Protocol
- Service Path (vPath)
Cisco Open Network Environment (ONE)

Industry’s Most Comprehensive Networking Portfolio

Hardware + Software
Physical + Virtual
Network + Compute

Applications

1. Platform and Network Protocol APIs
   - One Platform Kit (onePK)
   - Network Programmatic APIs
   - NetOS: NX-OS, IOS, IOS-XR

2. Controllers and Agents
   - SDN:
     - ONE Controller (OpenFlow, onePK)
     - OpenFlow 1.x support

3. Overlays and Fabrics
   - Multi-hypervisor
   - Multi-service
   - Multi-cloud
   - VXLAN/OTV/LISP
   - Openstack
   - Converged Fabric

www.cisco.com/go/one
Unification of Fabric and Data Center
Cisco Unified Fabric
Delivering Architectural Flexibility for All Data Centers

DELIVERING TO YOUR DATA CENTER NEEDS

Resilient, High Performance, Scalable Fabric
Workload Mobility Within/Across DCs
Secure Separation/multitenancy
LAN+SAN Convergence
Operational Efficiency / Consistency– P-V-
Cisco Unified Fabric
Delivering Architectural Flexibility for All Data Centers

**SCALE**
- Resilient, High Performance
- Revolutionary Scale
- Geographic Span

**CONVERGENCE**
- Wire Once for LAN/SAN
- Single Point of Management for LAN/SAN
- Device Consolidation

**INTELLIGENCE**
- Seamless VM Networking
- Workload Mobility

When the Network is UNIFIED
You Get CONSISTENCY Across Physical, Virtual and Cloud
When the Network is UNIFIED You Get SOLUTIONS WHEN YOU NEED

LAN/SAN Convergence

Data Center Consolidation

Virtualization/Journey to Private Cloud

Business Continuity, Disaster Recovery

Desktop Virtualization

High-Performance Computing and High-Frequency Trading
Key Takeaways

- Next generation Data Center Architecture needs to take a Fabric-based approach to address different needs from diverse application and service model
- FabricPath is central to Cisco DC Fabric strategy and has a long life ahead
- Lots of new functionalities and platform support planned for FabricPath
- Important to start thinking about the paradigm shift of removing L3 functionalities from the FP spine devices
  - Deployment of L2/L3 boundary on Border Leaf Devices
  - Allows for a smoother migration/coexistence with Cisco Next Generation Fabric offering
- Fabric selection should take speed and media into consideration
Q & A
Thank you.