Cisco Expo KSA

Cisco Data Center Interconnect

Faadil Adam
Systems Engineer – Service Provider

5th December 2011
Agenda

• Data Center Interconnect (DCI):
  Business Drivers
  DCI Solution Components
    SAN Extension
    LAN Extension
    Path Optimization

• LAN Extension Technologies
  Overview
  Ethernet Based
  Overlay Transport Virtualization
  A-VPLS
  VPLS

• Conclusion and Q&A
Data Centers Under Increasing Pressure

New Business Pressures

Satisfying Increasing Demand While Decreasing Cost

Operational Limitations

Collaborative Services

Visual Networking

User Expectations

Economy

Regulation

Power and Cooling

Asset Utilization

Provisioning

Security Threats

Bus. Continuity

© 2010 Cisco and/or its affiliates. All rights reserved.
Data Center Interconnect
Business Drivers

- Data Centers are extending beyond traditional boundaries
- Virtualization applications are driving DCI across PODs (aggregation blocks) and Data Centers

<table>
<thead>
<tr>
<th>Drivers</th>
<th>Business Solution</th>
<th>Constraints</th>
<th>IT Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Continuity</td>
<td>✅ Disaster Recovery</td>
<td>✅ Stateless</td>
<td>✅ GSLB</td>
</tr>
<tr>
<td></td>
<td>✅ HA Framework</td>
<td>✅ Network Service Sync</td>
<td>✅ Geo-clusters</td>
</tr>
<tr>
<td>Operation Cost Containment</td>
<td>✅ Data Center Maintenance / Migration / Consolidation</td>
<td>✅ VLAN Extension</td>
<td>✅ Distributed Virtual Data Center</td>
</tr>
<tr>
<td>Business Resource Optimization</td>
<td>✅ Disaster Avoidance</td>
<td>✅ Stateful</td>
<td>✅ VM Mobility</td>
</tr>
<tr>
<td></td>
<td>✅ Workload Mobility</td>
<td>✅ Bandwidth</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>✅ Latency</td>
<td></td>
</tr>
<tr>
<td>Cloud Services</td>
<td>✅ Inter-Cloud Networking</td>
<td>✅ Flexibility</td>
<td>✅ VM Mobility</td>
</tr>
<tr>
<td></td>
<td>✅ XaaS</td>
<td></td>
<td>✅ Automation</td>
</tr>
</tbody>
</table>
Data Center Interconnect
Solution Components

<table>
<thead>
<tr>
<th>DCI Function</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage Extension</td>
<td>Providing applications access to storage locally, as well as remotely with desirable storage attributes</td>
</tr>
<tr>
<td>LAN Extension</td>
<td>Extend same VLAN across Data Centers, to virtualize servers and applications</td>
</tr>
<tr>
<td>Path Optimization</td>
<td>Routing users to the data center where the application resides while keeping symmetrical routing in consideration for IP services (e.g. Firewall)</td>
</tr>
<tr>
<td>Inter-DC Routing</td>
<td>Provide routed connectivity between data centers (used for L3 segmentation/virtualization, etc.)</td>
</tr>
</tbody>
</table>
Data Center Interconnect
SAN Extension

- SAN Extension typically over DWDM/ Dark Fiber
- Synchronous/ Asynchronous Replication
- Storage Elasticity
Data Center Interconnect
LAN Extension

- STP Isolation is the key element
- Multipoint
- Loop avoidance + Storm-Control
  Unknown Unicast & Broadcast control
- Link sturdiness
- Scale & Convergence
Data Center Interconnect
Path Optimization

Options

- Egress
  - Addressed by FHRP Filtering
- Ingress:
  1. DNS redirection with ACE/GSS
  2. Route Injection
  3. LISP
Agenda

• Data Center Interconnect (DCI):
  Business Drivers
  DCI Solution Components
    SAN Extension
    LAN Extension
    Path Optimization

• LAN Extension Technologies
  Overview
  Ethernet Based
  Overlay Transport Virtualization
  A-VPLS
  VPLS

• Conclusion and Q&A
LAN Extension Overview
Data Center Interconnect
Technology Options

- **VSS & vPC, FabricPath/ TRILL**
  - Applies easily for dual site interconnection
  - Over dark fiber or protected DWDM
  - Easy crypto using end-to-end 802.1AE

- **OTV**
  - L2oL3 for link protection (Fast detection & convergence / Dampening)
  - Enterprise / DC focus
  - Easy integration over Core
  - Works over any transport
  - Innovative MAC routing

- **EoMPLS & VPLS & A-VPLS**
  - L2oL3 for link protection (Fast detection & convergence / Dampening)
  - PE style
  - Large scale
  - Multi-tenants
  - Works over GRE
  - Most deployed today
Ethernet
VSS & vPC
Both VSS-MEC and vPC are 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 dependence on STP in the L2 access-distribution Layer. Scale Available Layer 2 Bandwidth. Simplify Network Design.
Dual Sites Interconnection
Leveraging MECs Between Sites

At DCI point:
- STP Isolation (BPDU Filtering)
- Broadcast Storm Control
- FHRP Isolation

- Link utilization with MEC
- New Links for DCI
  - DCI port-channel
    - 2 with VSS
    - 4 with vPC (pre 4.2(6) release)
- Requires protected DWDM or Direct fibers
Dual Sites Use Case Summary
Cisco Validated Design on CCO

Overlay Transport Virtualization (OTV)
Overlay Transport Virtualization
Technology Pillars

OTV is a “MAC in IP” technique to extend Layer 2 domains OVER ANY TRANSPORT

Dynamic Encapsulation
- No Pseudo-Wire State Maintenance
- Optimal Multicast Replication
- Multipoint Connectivity
- Point-to-Cloud Model

Nexus 7000
First platform to support OTV

Protocol Learning
- Preserve Failure Boundary
- Built-in Loop Prevention
- Automated Multi-homing
- Site Independence
OTV Data Plane
Inter-Sites Packet Flow

MAC TABLE
<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>MAC 1</td>
<td>Eth 2</td>
</tr>
<tr>
<td>100</td>
<td>MAC 2</td>
<td>Eth 1</td>
</tr>
<tr>
<td>100</td>
<td>MAC 3</td>
<td>IP B</td>
</tr>
<tr>
<td>100</td>
<td>MAC 4</td>
<td>IP B</td>
</tr>
</tbody>
</table>

MAC TABLE
<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>MAC 1</td>
<td>IP A</td>
</tr>
<tr>
<td>100</td>
<td>MAC 2</td>
<td>IP A</td>
</tr>
<tr>
<td>100</td>
<td>MAC 3</td>
<td>Eth 3</td>
</tr>
<tr>
<td>100</td>
<td>MAC 4</td>
<td>Eth 4</td>
</tr>
</tbody>
</table>

Layer 2 Lookup
2

Encap
3

Decap
5

Layer 2 Lookup
6

MAC 1 ➔ MAC 3

West Site

Server 1

East Site

Server 3

OTV Data Plane
Inter-Sites Packet Flow
OTV Data Plane
Encapsulation

- OTV encapsulation adds 42 Bytes to the packet IP MTU size
  Outer IP Header and OTV Shim Header in addition to original L2 Header
- The outer OTV shim header contains information about the overlay (VLAN, overlay number)
- The 802.1Q header is removed from the original frame and the VLAN field copied over into the OTV shim header

42 Byte total overhead

*4B of original 802.1Q header removed
**OTV Terminology**

- **Edge Device**: Responsible for performing all the OTV functionality.
- **Internal Interface**: Regular L2 interface with typical L2 functions (local switching, STP, learning, flooding).
- **Join Interface**: Join OTV overlay, Discover & Adjacencies with other OTV Edge devices.
- **Overlay Interface**: Configuration applied, OTV encapsulation performed.

---

**Transport Infrastructure**

- **OTV Edge Device**
- **Overlay Interface**
- **OTV Join Interface**
- **OTV Internal Interfaces**
OTV Control Plane
Neighbor Discovery (Multicast/ Unicast)

**Mechanism**
- Edge Devices (EDs) join an multicast group in the transport, as they were hosts (ASM)
- OTV hellos and updates are encapsulated in the multicast group

**End Result**
- Adjacencies are maintained over the multicast group
- A single update reaches all neighbors
OTV Configuration

OTV over a Multicast Transport

- Minimal configuration required to get OTV up and running

```plaintext
feature otv
otv site-id 1*
otv site-vlan 99
interface Overlay100
  otv join-interface e1/1
  otv control-group 239.1.1.1
  otv data-group 232.192.1.0/24
  otv extend-vlan 100-150

feature otv
otv site-id 3*
otv site-vlan 99
interface Overlay100
  otv join-interface e1/1.10
  otv control-group 239.1.1.1
  otv data-group 232.192.1.0/24
  otv extend-vlan 100-150

feature otv
otv site-id 2*
otv site-vlan 99
interface Overlay100
  otv join-interface Po16
  otv control-group 239.1.1.1
  otv data-group 232.192.1.0/24
  otv extend-vlan 100-150
```

*Introduced from release 5.2
OTV Layer 2 Fault Isolation

- **STP isolation** – No configuration required
  - No BPDUs forwarded across the overlay
  - STP remains local to each site
  - Edge device internal interfaces behave as any other switchport

- **Unknown unicast isolation** – No configuration required
  - No unknown unicast frames flooded onto the overlay
  - Assumption is that end stations are not silent
  - Option for selective unknown unicast flooding (for certain applications)

- **Proxy ARP cache for remote-site hosts** – On by default
  - On ARP request for remote host, request forwarded through OTV and initial ARP reply generated by that host
  - OTV edge device snoops ARP replies and caches data
  - Subsequent ARP replies proxied by local OTV edge device using ARP cache
OTV Scalability
Current and Future Supported Values

The following values have been tested and verified:

- 6 Sites
- 256 OTV extended VLANs
- 16K MAC Addresses across all the extended VLANs
- 3000 Sites’ Multicast Data Groups
OTV
Summary

- Extensions over any transport (IP, MPLS)
- Failure boundary preservation
- Site independence
- Optimal BW utilization with multicast enabled transport infrastructure (no head-end replication)
- Automated Built-in Multihoming
- End-to-End loop prevention
- Scalability
  - Sites, VLANs, MACs
- Operations simplicity
MPLS
EoMPLS, A-VPLS, VPLS
MPLS for DCI
Efficient WAN Technology

- Scalable L3 Segmentation
- FastReRoute for sub-50ms convergence
- Traffic-Engineering for SLA control and path diversity

MPLS
- EoMPLS
- P2P
- A-VPLS
- H-VPLS

- L2oL3 + DCI
- Multi-tenants
- Mature L2 transport
- Tight Core SLA control

BGP VRF
EoMPLS Usage with DCI
End-to-End Loop Avoidance Using Edge to Edge LACP

- BPDU Filtering to maintain STP domains isolation
- Storm-control for data-plane protection
- Configuration applied at aggregation layer on the logical port-channel interface

```plaintext
interface port-channel70
  description L2 PortChannel to DC 2
  spanning-tree port type edge trunk
  spanning-tree bpdufilter enable
  storm-control broadcast level 1*
  storm-control multicast level x
```

*Value to be tuned, min is 0.3
Multi-Point Topologies
What Is VPLS?

One extended bridge-domain built using:
- VFI = Virtual Forwarding Instance
  ( VSI = Virtual Switch Instance)
- PW = Pseudo-Wire
- SVI = Switch Virtual Interface
- xconnect

Mac address table population ➔ is pure Learning-Bridge
VPLS
L2 Signalling and Forwarding (aka Transparent-Bridging)

A VSI/VFI operates like a conventional L2 switch!
VPLS

Split-Horizon for Loop Avoidance

- A packet will never be bridged from a PW to an other PW in the VFI
- Assuming PW full-mesh in a VFI:
  - Full reachability
  - Core link back-up
  - No core L2 loop
  - No need for a loop prevention core STP

Important remark:
Split-Horizon protects against core loop, but Split-Horizon does not protect against global loops due to dual homing of edge devices to PE
VPLS
Applying Key DCI Functions with VPLS

- BPDU are not transmitted by default
- Storm-control is on ingress link
- FHRP isolation to allow active/active default gateway + localization
High Scale VPLS DCI

The Main Model

Main functional elements:

- Standard MPLS core
- Standard VPLS
- Multi-Chassis Etherchannel access circuit

Classic VPLS full mesh, MAC bridging
Mature, proven, Cisco “CVD” solution

Simplify the VPLS dual-homing with MC-LAG, ~2 second convergence time, eliminate STP
High-Scale DCI Key Requirements
Building the Data Center Cloud

- Resilient
  - <50ms~1sec service convergence
  - Multi-Homing
  - Load balancing: per VLAN/per flow
  - L2 domain isolation, Storm control

- Scale
  - VLAN, MAC, VPLS PW Scale
  - FIB, ARP Scale

- SLA
  - Guaranteed QoS
  - Security: OAM, etc.

- Perf
  - Dense 10G Line Rate Forwarding
  - 100GE roadmap
  - Multicast replication efficiency

Cloud Dispersed Data Centers to Empower:
- Distributed applications
- Pool and maximize global resources
- Business continuity
Advanced VPLS (A-VPLS)

- Easy-to-use CLI for VLAN extension
- Reduced complexity!
- VSS single chassis redundancy
- Fast sub-second failover
- Ethernet LAN extension over MPLS or IP:
  - “Any flow Any Link” load-balancing
  - Multipoint loop-free connectivity
- A-VPLS
- Inter-operates with Standard VPLS Without flow based LB
- DC edge to aggregation
- DC edge to WAN
- WAN core

Cisco Enhancements to VPLS Standard
A-VPLS Simplified Configuration

Traditional VPLS Config
(Repeat per VLAN)

l2 vfi for_10 manual
  vpn id 10
  neighbor 2.2.2.2 encap mpls
  neighbor 3.3.3.3 encap mpls
!
l2 vfi for_20 manual
  vpn id 20
  neighbor 2.2.2.2 encap mpls
  neighbor 3.3.3.3 encap mpls
!
.............
!
interface Vlan10
  xconnect vfi for_10
!
interface Vlan20
  xconnect vfi for_20
.............

A-VPLS Simplified Config
(Configure once)

pseudowire-class cl1
  encap mpls

! enable ML PW (ECMP LB)
load-balance flow

! enable FAT PW
flow-label enable
!
interface virtual-ethernet 1
  ! transport configuration
transport vpls mesh
  neighbor 2.2.2.2 pw-class cl1
  neighbor 3.3.3.3 pw-class cl1

! service configuration
switchport
switchport mode trunk
switchport trunk allowed vlan
range 10 to 2000

PE1 (1.1.1.1)

IP/MPLS

PE2 (2.2.2.2)

PE3 (3.3.3.3)

Traditional VPLS Config
(Repeat per VLAN)

A-VPLS Simplified Config
(Configure once)
Simplified Redundancy

Single Virtual Ethernet Interconnect across Multiple Interfaces

Single control plane and management

Pseudowire
Agenda

• Data Center Interconnect (DCI):
  Business Drivers
  DCI Solution Components
    SAN Extension
    LAN Extension
    Path Optimization

• LAN Extension Technologies
  Overview
  Ethernet Based
  Overlay Transport Virtualization
  A-VPLS
  VPLS

• Conclusion and Q&A
Choosing an Appropriate DCI Design

- Number of VLANs that need to be extended across data centers
- Number of data centers that need to interconnected
- Total amount of intra-data center bandwidth required
- Convergence and recovery times in failure scenarios
- Number of servers (MAC address scalability)
- Platform capable of providing existing data center features
- Possibility of leveraging existing network equipment
# LAN Extension

Solution to Product Portfolio Table

<table>
<thead>
<tr>
<th></th>
<th>ASR 1000</th>
<th>Cat 6500</th>
<th>Nexus 7000</th>
<th>ASR 9000</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ethernet Based</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vPC</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>VSS</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td><strong>MPLS Based</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EoMPLS</td>
<td>✓</td>
<td>✓</td>
<td>Q4 2012</td>
<td>✓</td>
</tr>
<tr>
<td>EoMPLSoGRE</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(requires SIP card support)</td>
<td></td>
</tr>
<tr>
<td>VPLS</td>
<td></td>
<td>✓</td>
<td>Q4 2012</td>
<td>✓</td>
</tr>
<tr>
<td><strong>IP Based</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTV</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Q4 2011: EoMPLS; Q4 2012: VPLS
Data Center Interconnect
Where to Go for More Information

Extend Layer 2 Applications
Overlay Transport Protocol (OTV) technology supports transparent workload mobility within and across dispersed data centers. (3.54 min)

Get Instant Workload Mobility Among Data Centers
Cisco Data Center Interconnect (DCI) solutions can help your IT organization meet business continuity and corporate compliance objectives.

- Reduce the business impact of disaster events and help ensure business continuity
- Improve productivity through enhanced application and data availability
- Meet corporate and regulatory compliance needs and improve data security

These solutions transparently extend LAN and SAN connectivity and provide accelerated, highly secure data replication, server clustering, and workload mobility between geographically dispersed data centers. This enhances business resilience, and helps enable application and data mobility between data centers, while maintaining operational consistency.

Featured Products

- **Cisco Nexus 7000 Series LAN Extension**
  Simplify Layer 2 applications across distributed data centers.

- **Cisco Catalyst 6500 Series LAN Extension**
  Deliver high-performance, scalable Layer 2 extension with subsecond convergence.

- **Cisco MDS 9500 Series SAN Extension Over IP**
  Gain an integrated, cost-effective, reliable business continuance solution.

http://www.cisco.com/go/dci
Thank you.