Data Centre Interconnect with Overlay Transport Virtualisation

BRKDCT-2049

Mitch Roberts
Systems Engineer

#clmel
OTV – Overlay Transport Virtualisation
Simplifying Data Centre Interconnect

Any Workload

Anytime

Anywhere
Session Objectives

• The main goals of this session are:

• This session features a detailed analysis of the architectural aspects and deployment benefits behind OTV

• The attendees will learn how OTV is aimed at providing Layer 2 connectivity beyond the Layer 3 boundary while maintaining the failure containment and operational simplicity that the Layer 3 boundary provides

• The attendees will get a deep knowledge of how the OTV control-plane and data-plane work to provide the VLAN extension
Session Non-objectives

• This session does not include:
• In depth discussion of Path Optimisation technologies (DNS, LISP, etc.)
• Storage extension considerations associated to DCI deployments
• Workload mobility application specific deployment considerations
• In depth discussion Multicast

Please Visit – Great Content + Vods
• http://www.ciscolive.com/global/
## Related Cisco Live Events

<table>
<thead>
<tr>
<th>Session-ID</th>
<th>Session Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRKDCT-2334</td>
<td>Real World Data Centre Deployments and Best Practices</td>
</tr>
<tr>
<td>BRKDCT-2615</td>
<td>How to Achieve True Active-Active Data Centre Infrastructures</td>
</tr>
<tr>
<td>BRKDCT-2218</td>
<td>Small to Medium Data Centre Designs</td>
</tr>
<tr>
<td>BRKRST-1601</td>
<td>Fast Track to Fast-IT – A Practical Approach</td>
</tr>
</tbody>
</table>
Agenda

• Distributed Data Centres: Goals and Challenges
• OTV Architecture Principles
• OTV Design Considerations & New Features
Distributed Data Centres Goals

• Ensure business continuity
• Distributed applications
• Seamless workload mobility
• Maximise compute resources
Data Centre Interconnect

Traditional Layer 2 Extensions

- **Ethernet**
  - VSS & vPC or FabricPath
    - Applies easily for dual site interconnection
    - Over dark fibre or protected D-WDM
    - Easy crypto using end-to-end 802.1AE

- **IP**
  - OTV – Overlay Transport Virtualisation
    - MAC in IP

- **MPLS**
  - EoMPLS & VPLS & A-VPLS & H-VPLS
    - PE style
    - Multi-tenants
    - Most deployed today
Challenges in Traditional Layer 2 VPNs

Flooding Behaviour
- Unknown Unicast for MAC propagation
- Unicast Flooding reaches all sites

Pseudo-wire Maintenance
- Full mesh of Pseudo-wire is complex
- Head-End replication is a common problem

Multi-Homing
- Requires additional Protocols & extends STP
- Malfunctions impacts multiple sites
Technology Pillars
No Pseudo-Wire State Maintenance

Optimal Multicast Replication

Dynamic Encapsulation

Multipoint Connectivity

Point-to-Cloud Model
Preserve Failure Boundary  
Built-in Loop Prevention

Protocol Learning

Automated Multi-Homing  
Site Independence
OTV – Overlay Transport Virtualisation
Simplifying Data Centre Interconnect

Any Workload

Anytime

Anywhere

- **Nexus 7000** First platform to support OTV (since 5.0 NXOS Release)
- **ASR 1000** Now also supporting OTV (since 3.5 XE Release)
Agenda

• Distributed Data Centres: Goals and Challenges
• OTV Architecture
  – Terminology
  – Control Plane and Data Plane
  – Failure Isolation
  – Multi-homing
  – Mobility
  – L2 Multicast Forwarding
  – QoS and Scalability
  – Path Optimisation
• OTV Design Considerations & New Features
Terminology
OTV Devices and Interfaces

• Edge Device
  – Performs all OTV functionality
  – Usually located at the Aggregation Layer or at the Core Layer
  – Support for multiple OTV Edge Devices (multi-homing) in the same site

• Internal Interface
  – Site facing Interfaces of the Edge Devices
  – Carry VLANs extended through OTV
  – Regular Layer 2 interfaces
  – No OTV configuration required
  – Supports IPv4 & IPv6
**Terminology**

**OTV Devices and Interfaces**

- **Join Interface**
  - One of the uplink of the Edge Device
  - Point-to-point routed interface (physical interface, sub-interface or port-channel supported)
  - Used to physically “join” the Overlay network
  - No OTV specific configuration required
  - IPv4 only

- **Overlay Interface**
  - **Virtual** interface with most of the OTV configuration
  - Logical multi-access multicast-capable interface
  - Encapsulates Layer 2 frames in IP unicast or multicast
Agenda

• Distributed Data Centres: Goals and Challenges

• OTV Architecture Principles
  – Terminology
  – Control Plane and Data Plane
  – Failure Isolation
  – Multi-homing
  – Mobility
  – L2 Multicast Forwarding
  – QoS and Scalability
  – Path Optimisation

• OTV Design Considerations & New Features
OTV Control Plane

Building the MAC Tables

- No unknown unicast flooding (selective unicast flooding in 6.2)
- Control Plane Learning with proactive MAC advertisement
- Background process with no specific configuration
- IS-IS used between OTV Edge Devices
OTV Control Plane
Neighbour Discovery and Adjacency Formation

- Before any MAC address can be advertised the OTV Edge Devices must:
  - Discover each other
  - Build a neighbour relationship with each other

- Neighbour Relationship built over a transport infrastructure:
  - Multicast-enabled (all shipping releases)
  - Unicast-only (from NX-OS release 5.2 & IOS-XE 3.9)
OTV Control Plane

Neighbour Discovery (over Multicast Transport)

**Mechanism**
- Edge Devices (EDs) join an multicast group in the transport, as they were hosts (no PIM on EDs)
- OTV hellos and updates are encapsulated in the multicast group

**End Result**
- Adjacencies are maintained over the multicast group
- A single update reaches all neighbours
OTV Control Plane (Multicast Transport)

1. All edge devices join OTV control-group G
2. Multicast state for group G established throughout transport
3. Neighbour IP Addr
4. OTV Hello
5. OTV Hello
6. OTV Hello
7. OTV Hello

1. West
2. IP A
3. IP B
4. IP C
5. IP D
6. IP E
7. IP F

OTV Control Plane (Multicast Transport)

- Transport natively replicates multicast to all OIFs
- All edge devices join OTV control-group G
- Multicast state for group G established throughout transport
- Neighbour IP Addr
- OTV Hello

West

Multicast-enabled Transport

Decap

Encap

IGMP Join G

IGMP Join G

IGMP Join G

OTV Control Plane

OTV Control Plane

Neighbour West

IP Addr IP A

Neighbour West

IP Addr IP A
OTV Control Plane (Multicast Transport)

- **The South Site creates its hello with West’s address in the TLV.**

- **Bidirectional adjacency formed**

- **Multicast-enabled Transport**

- **Decap**

- **Encap**

- **IP A**

- **IP B**

- **IP C**

- **West**

- **South**

- **East**

- **Neighbour IP Addr**

- **Neighbour IP Addr**

- **OTV Hello**

- **OTV Control Plane**
OTV Control Plane

MAC Advertisements (over Multicast Transport)

New MACs learned in VLANs that are OTV extended

MAC Table

VLAN  | MAC   | IF
---|---|---
100  | MAC A | e1/1
101  | MAC B | e1/1
102  | MAC C | e1/1

1. New MACs learned in VLANs that are OTV extended
2. Craft OTV update with new MACs
3. Multicast-enabled Transport
4. OTV Control Plane
5. MAC Table
6. Add MACs learned through OTV
7. Add MACs learned through OTV

IP A ➔ G
Decap
Encap

MAC Table

VLAN  | MAC   | IF
---|---|---
100  | MAC A | IP A
101  | MAC B | IP A
102  | MAC C | IP A
Multicast Transport
OTV Control and Data Plane over Multicast Transport

• Use a High-Available Multicast Rendez-Vous Point (RP) configuration
  – PIM Anycast (RFC4610) or MSDP (Multicast Source Discovery Protocol)

• Requirements to Control Plane
  – PIM Any-Source-Multicast (ASM) Sparse-Mode

• Requirements to Data Plane
  – PIM Source-Specific-Multicast (SSM) or BiDir

Example:
Multicast for OTV on Nexus 7000

```
feature pim
!
interface loopback 0
  ip pim spare-mode
  ip address 192.168.1.100/32
!
interface loopback 1
  ip pim sparse-mode
  ip address 10.254.254.n1-x/32
!
ip pim rp-address 192.168.1.100 group-list 239.1.1.1
ip pim anycast-rp 192.168.1.100 10.254.254.n1
ip pim anycast-rp 192.168.1.100 10.254.254.n2
ip pim ssm range 232.239.1.0/24
!
interface port-channel1
  # This Interface peers with the OTV Join Interface
  ip igmp version3
```

*“n” in the last Octet reflects a unique IP address per Router joining the PIM Anycast Group*
OTV Control Plane

Neighbour Discovery (Unicast-only Transport)

• Ideal for connecting a small number of sites
• With a higher number of sites a multicast transport is the best choice

Mechanism
• Edge Devices (EDs) register with an “Adjacency Server”
• EDs receive a full list of Neighbours (oNL) from the AS
• OTV hellos and updates are encapsulated in IP and *unicast* to each neighbour

End Result
• Neighbour Discovery is automated by the “Adjacency Server”
• All signalling must be replicated for each neighbour
• Data traffic must also be replicated at the head-end

Release 5.2 and above
OTV Control Plane

CLI Verification

- Establishment of control plane adjacencies between OTV Edge Devices (multicast or unicast transport):

  ```
dc1-agg-7k1# show otv adjacency

Overlay Adjacency database
Overlays Interface Overlay100:

<table>
<thead>
<tr>
<th>Hostname</th>
<th>System-ID</th>
<th>Dest Addr</th>
<th>Up Time</th>
<th>Adj-State</th>
</tr>
</thead>
<tbody>
<tr>
<td>dc2-agg-7k1</td>
<td>001b.54c2.efc2</td>
<td>20.11.23.2</td>
<td>15:08:53</td>
<td>UP</td>
</tr>
<tr>
<td>dc1-agg-7k2</td>
<td>001b.54c2.e1c3</td>
<td>20.12.23.2</td>
<td>15:43:27</td>
<td>UP</td>
</tr>
<tr>
<td>dc2-agg-7k2</td>
<td>001b.54c2.e142</td>
<td>20.22.23.2</td>
<td>14:49:11</td>
<td>UP</td>
</tr>
</tbody>
</table>
```

- Unicast MAC reachability information:

  ```
dc1-agg-7k1# show otv route

OTV Unicast MAC Routing Table For Overlay100

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC-Address</th>
<th>Metric</th>
<th>Uptime</th>
<th>Owner</th>
<th>Next-hop(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>0000.0c07.ac01</td>
<td>1</td>
<td>3d15h</td>
<td>site</td>
<td>Ethernet1/1</td>
</tr>
<tr>
<td>2001</td>
<td>0000.1641.d70e</td>
<td>1</td>
<td>3d15h</td>
<td>site</td>
<td>Ethernet1/2</td>
</tr>
<tr>
<td>2001</td>
<td>0000.49f3.88ff</td>
<td>42</td>
<td>2d22h</td>
<td>overlay</td>
<td>dc2-agg-7k1</td>
</tr>
<tr>
<td>2001</td>
<td>0000.49f3.8900</td>
<td>42</td>
<td>2d22h</td>
<td>overlay</td>
<td>dc2-agg-7k2</td>
</tr>
</tbody>
</table>
```
OTV Data Plane

Inter-Site Packet Flow
OTV Data Plane

Encapsulation

- **42 Bytes** overhead to the packet IP MTU size (IPv4 packet)
  - Outer IP + OTV Shim - Original L2 Header (w/out the .1Q header)
- 802.1Q header is **removed** and the VLAN field copied over to the OTV shim header
- Outer OTV shim header contains VLAN, overlay number, etc.
- Consider Jumbo MTU Sizing

```
Encapsulation

DMAC  SMAC  Ether Type  IP Header  OTV Shim  L2 Header  Payload  CRC
6B     6B     2B        20B        8B       14B*       4B
```

* The 4 Bytes of .1Q header have already been removed

20B + 8B + 14B* = 42 Bytes of total overhead
Agenda

• Distributed Data Centres: Goals and Challenges

• OTV Architecture Principles
  – Terminology
  – Control Plane and Data Plane
  – Failure Isolation
  – Multi-homing
  – Mobility
  – L2 Multicast Forwarding
  – QoS and Scalability
  – Path Optimisation

• OTV Design Considerations & New Features
Spanning-Tree and OTV

Site Independence

- Site transparency: no changes to the STP topology
- Total isolation of the STP domain
- **Default behaviour**: no configuration is required
- BPDUs sent and received ONLY on Internal Interfaces
Unknown Unicast and OTV
No Longer Unknown Unicast Storms Across the DCI

- No requirements to forward unknown unicast frames
- Assumption: end-host are not silent or uni-directional
- Default behaviour: no configuration is required

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>Eth1</td>
</tr>
<tr>
<td>100</td>
<td>MAC 2</td>
<td>IP B</td>
</tr>
</tbody>
</table>

No MAC 3 in the MAC Table
Unknown Unicast and OTV
Selective Unicast Flooding

- Some Application requirement to forward unknown unicast frames
- Selective Unicast Flooding can be enabled per mac address
- Default behaviour: no unknown unicast forwarding

```
OTV-a # conf
Enter configuration commands, one per line. End with CNTL/Z
OTV-a(config)# otv flood mac 0000.2102.1111 vlan 172
```
Controlling ARP Traffic

ARP Neighbour-Discovery (ND) Cache

- ARP cache maintained in Edge Device by snooping ARP replies
- First ARP request is broadcasted to all sites. Subsequent ARP requests are replied by local Edge Device
- Timeout can be adjusted (as per NX-OS 6.1(1))
- Drastic reduction of ARP traffic on DCI
- ARP spoofing can be disabled
- IPv4 only feature
- Default behaviour: no configuration is required

New: Release 6.1

```
OTV-a(config)# interface overlay 1
OTV-a(config-if-overlay)# no otv suppress-arp-nd

# Allows ARP requests over an overlay network and disables ARP caching on edge devices. This command does not support IPv6.
```

```
OTV-a(config)# interface overlay 1
OTV-a(config-if-overlay)# otv arp-nd timeout 70

# Configures the time, in seconds, that an entry remains in the ARP-ND cache.
The time is in seconds varying from 60 to 86400. The default timeout value is 480 seconds.
```
Agenda

• Distributed Data Centres: Goals and Challenges

• OTV Architecture Principles
  – Terminology
  – Control Plane and Data Plane
  – Failure Isolation
  – Multi-homing
  – Mobility
  – L2 Multicast Forwarding
  – QoS and Scalability
  – Path Optimisation

• OTV Design Considerations & New Features
OTV Multi-homing

Fully Automated Multi-homing

- No additional protocols required (i.e. BGP)
- OTV site-vlan used to discover OTV neighbour in the same site
- Authoritative Edge Device (AED) Election takes place
- Extended VLANs are split across the AEDs
- The AED is responsible for:
  - MAC address advertisement for its VLANs
  - Forwarding its VLANs’ traffic inside and outside the site
Hardened Multi-homing

Introducing OTV Site-identifier

• Same site devices must use common site-identifier
• Site-id information is included in the control plane
• Makes OTV multi-homing more robust and resilient
  – Site Adjacency and Overlay Adjacency are now both leveraged for AED election
• An overlay will not come up until a site-id is configured
  – Site and Overlay Adjacency are both leveraged for AED election

feature otv
otv site-identifier 0x1
otv site-vlan 99
OTV Multi-homing
VLANs Split across AEDs

- Automated and deterministic algorithm
- In a dual-homed site:
  - Lower IS-IS System-ID (Ordinal 0) = EVEN VLANs
  - Higher IS-IS System-ID (Ordinal 1) = ODD VLANs

```
OTV-a# show otv vlan
OTV Extended VLANs and Edge Device State Information (* - AED)
VLAN  Auth. Edge Device  Vlan State         Overlay
----  -------------------  ------------------  -------
 100   East-b              inactive(Non AED) Overlay100
 101*  East-a              active             Overlay100
 102   East-b              inactive(Non AED) Overlay100

OTV-b# show otv vlan
OTV Extended VLANs and Edge Device State Information (* - AED)
VLAN  Auth. Edge Device  Vlan State         Overlay
----  -------------------  ------------------  -------
 100*  East-b              active             Overlay100
 101   East-a              inactive(Non AED) Overlay100
 102*  East-b              active             Overlay100
```
OTV Multi-homing
AED and Broadcast Handling

1. Broadcast reaches all the Edge Devices within the site
2. Only the AED forwards the traffic to the Overlay
3. All the Edge Devices at the other sites receive the broadcast
4. At the remote sites only the AEDs forward it into the site
Agenda

• Distributed Data Centres: Goals and Challenges

• OTV Architecture Principles
  – Terminology
  – Control Plane and Data Plane
  – Failure Isolation
  – Multi-homing
  – Mobility
  – L2 Multicast Forwarding
  – QoS and Scalability
  – Path Optimisation

• OTV Design Considerations & New Features
OTV and MAC Mobility
MAC Moving and OTV Updates (1)

1. Workload moved between Data Centre sites
OTV and MAC Mobility

MAC Moving and OTV Updates (2)

1. Workload moved between Data Centre sites
2. Workload is detected in East DC and OTV control plane is triggered

2.1) Server originates a Gratuitous ARP (GARP) frame
2.2) AED detects MAC X is now local
2.3) AED advertises MAC X with a metric of zero
2.4) EDs in site West see MAC X advertisement with a better metric from site East and change them to remote MAC address.
OTV and MAC Mobility
MAC Moving and OTV Updates (3)

1. Workload moved between Data Centre sites
2. Workload is detected in East DC and OTV control plane is triggered
3. East to West OTV data plane traffic allows to update the MAC tables of the L2 devices in West Site

Note: GARP is used as example traffic, same behaviour is achieved with any other L2 broadcast frames exchanged
Agenda

• Distributed Data Centres: Goals and Challenges

• OTV Architecture Principles
  – Terminology
  – Control Plane and Data Plane
  – Failure Isolation
  – Multi-homing
  – Mobility
  – L2 Multicast Forwarding
  – QoS and Scalability
  – Path Optimisation

• OTV Design Considerations & New Features
L2 Multicast Traffic Between Sites
Multicast Enabled Transport

• OTV can leverage the multicast support available in the transport network to **optimise** the delivery of the multicast traffic for the VLANs stretched across sites

• Three steps:
  1. **Automated mapping** of the sites’ multicast groups to a **range of multicast groups** in the transport network
  2. Creation of the Multicast state information at the OTV Edge Devices
  3. Sites’ Multicast traffic delivered over the Overlay
L2 Multicast with Multicast Transport

Step 1 – Mapping of the Site Multicast Group

- The site multicast groups are mapped to a SSM group range in the core
- Each (S1,Gs1) maps to a different SSM group in round-robin fashion

1) The Mcast source starts sending traffic to the group Gs1
2) The West ED maps (S1,Gs1) to a delivery group Gd1
3) The West ED communicates the mapping information (including the source VLAN) to the other EDs
4) Same process happens once source S2 is enabled (sending to a different group Gs2)
L2 Multicast with Multicast Transport

Step 2 – Multicast State Creation

It is important to clarify that the edge devices join the core multicast groups as hosts, not as routers!
L2 Multicast with Multicast Transport

Step 3 – Multicast Packet Flow

1. **Lookup**

<table>
<thead>
<tr>
<th>OIF-List</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
</tr>
<tr>
<td>Gs1</td>
</tr>
</tbody>
</table>

2. **Encap**

3. **Multicast Transport Replication**

4. **Decap**

5. **Decap**
L2 Multicast with Multicast Transport

Multicast Groups in the Core

OTV can leverage the benefits of a multicast-enabled transport for both control and data planes. The following summarises the requirements for a multicast transport:

• **Control group** – Single PIM-SM or PIM-Bidir group used to form adjacencies and exchange MAC reachability information

• **Data groups** – Range of SSM groups used to carry multicast data traffic generated by the sites

```
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
```

The right number of SSM groups to be used depends on a tradeoff between the amount of multicast state to be maintained in the core and the optimisation of Layer 2 multicast traffic delivery.
Agenda

• Distributed Data Centres: Goals and Challenges

• OTV Architecture Principles
  – Terminology
  – Control Plane and Data Plane
  – Failure Isolation
  – Multi-homing
  – Mobility
  – L2 Multicast Forwarding
  – QoS and Scalability
  – Path Optimisation

• OTV Design Considerations & New Features
QoS and OTV
Marking on Encapsulation

• On Encapsulation
  – CoS bits (802.1p) copied to the OTV shim header
  – If IP traffic: The original (inner) DSCP value is also copied to “outer” DSCP
QoS and OTV
Marking on De-capsulation

• On De-capsulation
  – CoS value is recovered from the OTV shim and added to the 802.1Q header

• Original CoS and DSCP are both preserved

• OTV Control Traffic is statically marked at CoS = 6/DSCP = 48
OTV Scalability
Current and Future Supported Values

NX-OS 6.2
- Sites: 8*
- OTV extended VLANs: 1500
- MAC addresses across all the extended VLANs: 32k
- Multicast Data Groups: 4000

NX-OS 5.2
- Sites: 6*
- OTV extended VLANs: 256
- MAC addresses across all the extended VLANs: 16k
- Multicast Data Groups: 2000

* two ED per Site
Agenda

• Distributed Data Centres: Goals and Challenges

• OTV Architecture Principles
  – Terminology
  – Control Plane and Data Plane
  – Failure Isolation
  – Multi-homing
  – Mobility
  – L2 Multicast Forwarding
  – QoS and Scalability
  – Path Optimisation

• OTV Design Considerations & New Features
Path Optimisation
Egress Routing Optimisation

Hot Potato Routing
Path Optimisation
Egress Routing with LAN Extension

• Extended VLANs typically have associated HSRP groups
• By default, only one HSRP router elected active, with all servers pointing to HSRP VIP as default gateway
• **Result:** sub-optimal routing
Egress Routing Localisation

FHRP Filtering Solution

• Filter FHRP with combination of VACL and MAC route filter

• **Result:** Still have one HSRP group with one VIP, but now have active router at each site for optimal first-hop routing
Path Optimisation
Optimal Routing Challenges

• Layer 2 extensions represent a challenge for optimal routing
• Challenging placement of gateway and advertisement of routing prefix/subnet
Path Optimisation
Is it relevant to my Data Centre model?

• Logical Data Centre or Physical Data Centre?
• High Availability or Disaster Recovery?

Is this ONE Logical Data Centre?
(Or do I have TWO Physical & Logical... separated Data Centre?)
Specific Use-Case
IPv6 and OTV

- IPv6 Unicast Forwarding and Multicast Flooding supported across OTV
  - Requires to disable optimised multicast forwarding (OMF) in IGMP snooping on OTV ED
- IPv6 Transport Network (Join Interface & Source Interface, not yet supported)

Global (all VLAN):
no ip igmp snooping optimise-multicast-flood

Per VLAN with IPv6 Traffic
vlan vlan-id
vlan configuration
no ip igmp snooping optimise-multicast-flood
Ingress Routing Localisation

Possible Solutions

**Challenge**
- Subnets are spread across locations
- Subnet information in the routing tables is not specific enough
- Routing doesn’t know if a server has moved between locations
- Traffic may be sent to the location where the application is not available

**Options**
- DNS Based
- Route Injection
- LISP – Locator/ID Separation Protocol

For more details on LISP and OTV Deployment see: BRKDCT-2131
OTV – Overlay Transport Virtualisation
Simplifying Data Centre Interconnect

Any Workload

Anytime

Anywhere
Agenda

• Distributed Data Centres: Goals and Challenges
• OTV Architecture Principles
• OTV Design Considerations & New Features
OTV Support

ASR1000

• OTV has been introduced in IOS XE 3.5 (Nov 2011)

• To use OTV on ASR1000, you require:
  – Advance Enterprise Image or Advance IP Service + OTV feature license

• ASR1k <-> N7k Inter-Site Interoperability has been tested
  – No ASR1k <-> N7k Multihoming Support (Intra-Site Interoperability)

• OTV on ASR1000 Use Cases are:
  – Legacy Deployments – where DC may still be Catalyst based
  – New Small Data Centre and/or Disaster Recovery Sites – where Main DC is equipped with Nexus 7000
  – OTV with Layer-3 Encryption – where MACSec is no option for Inter-DC Encryption
OTV Support

ASR 1000

• New Features for IOS-XE 3.9
  – OTV Adjacency Server (unicast)
  – OTV with LISP ESM
  – RPVST STP Support

• New Features for IOS-XE 3.10
  – Portchannel for join interface
  – VRF Aware
  – Subinterface for join interface
  – Layer 2 portchannel
Specific Use-Case
Transparent Firewall and extended Inside & Outside VLANs

- Transparent/Bridged Firewall is separating OTV extended VLANs
- OTV is sharing the same MAC address per Edge-Device
Specific Use-Case
Transparent Firewall and extended Inside & Outside VLANs

- OTV is sending PIM hellos with source of 0.0.0.0 destination 224.0.0.13
- Hello is sourced from OTV Edge Device (VDC) MAC Address

<table>
<thead>
<tr>
<th>Firewall MAC Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAC</td>
</tr>
<tr>
<td>OTV ED</td>
</tr>
<tr>
<td>OTV ED</td>
</tr>
</tbody>
</table>

PIM Hellos sourced from 0.0.0.0 (ED MAC Address) on VLAN 150 to VLAN 100
PIM Hellos sourced from 0.0.0.0 (ED MAC Address) on VLAN 100 to VLAN 150
OTV compared to FabricPath
Is FabricPath a valid Solution to replace OTV

• OTV is purpose build for Data Centre Interconnects
  – Cisco Validated Designs (CVDs)
  – Specific Data Centre Interconnect features

• On Data Centre Interconnect, FabricPath is NOT so Plug and Play
  – No specific DCI functions
  – Designs gotchas but do not impact all customers
  – Multidestination Trees capacity planning is key

• FabricPath can be a valid Data Centre Interconnect solution when:
  – Short distances between Data Centres
  – Multicast is not massively used
  – If you know and accept where your Traffic Flows (Multidestination Trees)
OTV compared to FabricPath

- Yes, but Data Centre Interconnect is NOT LAN Switching
- Customer’s constraints/needs are unique
- Scoping is based on
  - Application Involved
  - Number of DC sites, meshing, distances, bandwidth requirements
  - Customer Perception
  - Traffic Flows (Unicast, Multicast & Flooding)

<table>
<thead>
<tr>
<th></th>
<th>Operations Simplicity</th>
<th>Failure Isolation</th>
<th>Transport Failure Detection</th>
<th>3+ Sites Optimisation</th>
<th>High Availability</th>
<th>L2 Functions</th>
<th>L3 Unicast Functions</th>
<th>Multicast Functions</th>
<th>Scalability</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTV</td>
<td>✓</td>
<td>✓✓</td>
<td>✓</td>
<td>✓</td>
<td>✓✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>FabricPath</td>
<td>✓✓</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Stacking</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>VSS</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>vPC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
New Feature for OTV in NX-OS 6.2
Nexus 7000 Hardware Support

- F3 Support for OTV in 6.2(6)
  - Enable OTV on Nexus 7700 Series
  - No Tunnel Depolarisation or VLAN Translation in 6.2(6) on F3

<table>
<thead>
<tr>
<th>Internal Interface</th>
<th>M1</th>
<th>M2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>F1</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>F2e</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

- F1 and F2e support for OTV internal Interface
  - F1 and F2e linecards have the ability to be internal interfaces
New Features for OTV
Tunnel Depolarisation & Secondary IP

• Secondary IP command introduced
  – Configured within interface, not OTV interface
• Introduction of multiple IPs results in tunnel depolarisation

```bash
OTV-a(config-if)# ip address 2.100.11.1/24 secondary
Disabling IP Redirects on port-channel11 :secondary address configured.
OTV-a(config-if)# sh run int po11
  Command: show running-config interface port-channel11
  version 6.2(2)
  interface port-channel11
    no ip redirects
    ip address 2.100.11.100/24
    ip address 2.100.11.1/24 secondary
    ip ospf network point-to-point
    ip router ospf 1 area 0.0.0.0
    ip igmp version 3
OTV-a (config-if)# sh otv
  OTV Overlay Information
  Site Identifier 0000.0000.0011
  Overlay interface Overlay1
    VPN name: Overlay1
    VPN state: UP
    Extended vlans: 25-50 72-227 (Total:182)
    Control group: 224.1.1.0
    Data group range(s): 232.1.0.0/24
    Broadcast group: 224.1.1.0
    Join interface(s): Po11 (2.100.11.100)
    Secondary IP Addresses: : 2.100.11.1
    Site vlan: 1 (up)
    AED-Capable: Yes1
    Capability: Multicast-Reachable
```
New Features for OTV

VLAN Translation: Translation through transit VLAN

- When a different VLAN is used at multiple sites
- Usually for 3 or more sites
New Features for OTV

VLAN Translation: Translation through transit VLAN

OTV-a(config)# int overlay1
OTV-a(config-if-overlay)# otv vlan mapping 100 to 400

OTV-a(config-if-overlay)# sh run int overlay1
!Command: show running-config interface Overlay1
!Time: Fri Mar 29 19:01:04 2013
version 6.2(2)

interface Overlay1
  otv isis hello-multiplier 9
  otv join-interface port-channel11
  otv control-group 224.1.1.0
  otv data-group 232.1.0.0/24
  otv extend-vlan 25-50, 72-497
  otv vlan mapping 100 to 400
  no shutdown

OTV-a(config-if-overlay)# sh otv vlan-mapping
Original VLAN -> Translated VLAN
--------------------------------- 100 -> 400

OTV-B(config)# int overlay1
OTV-B(config-if-overlay)# otv vlan mapping 200 to 400
OTV-B(config-if-overlay)# sh run int overlay1
!Command: show running-config interface Overlay1
!Time: Fri Mar 29 19:02:29 2013
version 6.2(2)

interface Overlay1
  otv isis hello-multiplier 9
  otv join-interface port-channel21
  otv control-group 224.1.1.0
  otv data-group 232.1.0.0/24
  otv extend-vlan 25-50, 72-497
  otv vlan mapping 200 to 400
  no shutdown

OTV-B(config-if-overlay)# sh otv vlan-mapping
Original VLAN -> Translated VLAN
--------------------------------- 200 -> 400
OTV Convergence
Small and Large Scale Targets (Extreme Failures)

Large Scale
- <30sec
- <10sec

Small Scale
- <10sec
- <5sec
Challenges in Traditional Layer 2 VPNs
Solved by OTV

Flooding Behaviour
- Unicast Flooding reaches all sites
- Dynamic Encapsulation is complex
- Head-End replication is a common problem
- Native Automated Multi-Homing

Pseudo-wire Maintenance
- Control Plane Based Learning
- Unicast Flooding reaches all sites
- Dynamic Encapsulation is complex
- Head-End replication is a common problem
- Native Automated Multi-Homing

Multi-Homing
- Control Plane Based Learning
- Unicast Flooding reaches all sites
- Dynamic Encapsulation is complex
- Head-End replication is a common problem
- Native Automated Multi-Homing
- Malfunctions impacts multiple sites
Agenda

• Distributed Data Centres: Goals and Challenges
• OTV Architecture Principles
• OTV Design Considerations & New Features
OTV – Overlay Transport Virtualisation
Simplifying Data Centre Interconnect

Any Workload

Anytime

Anywhere
Where can OTV help YOU simplify Data Centre Interconnects?

http://www.cisco.com/go/DCI
Participate in the “My Favorite Speaker” Contest
Promote Your Favorite Speaker and You Could be a Winner

- Promote your favorite speaker through Twitter and you could win $200 of Cisco Press products (@CiscoPress)

- You can submit an entry for more than one of your “favorite” speakers
- Don’t forget to follow @CiscoLive and @CiscoPress
Continue Your Education

- Demos in the Cisco Campus
- Walk-in Self-Paced Labs
- Meet the Expert 1:1 meetings
Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2015 T-Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site http://showcase.genie-connect.com/clmelbourne2015
- Visit any Cisco Live Internet Station located throughout the venue

T-Shirts can be collected in the World of Solutions on Friday 20 March 12:00pm - 2:00pm

Learn online with Cisco Live!
Visit us online after the conference for full access to session videos and presentations. www.CiscoLiveAPAC.com
Thank you.