APPENDIX A

Best Practices/Caveats

DRaaS solution architecture is integration of different technologies and partnership with different vendors. Caveats, workarounds, recommendations and issues identified in proof of concept (POC) testing have been documented in this chapter.

Key Findings

The following key OTV and LISP findings are identified for consideration.

- OTV, page A-1
- LISP, page A-2

OTV

- VMXNET3 is recommended driver for CSR vNICs
- Cisco IOS XE 3.10S does not support BDI MAC address learning over OTV. DDTS CSCuj59314 has been logged to track support of BDI over OTV.
- With vCenter vswitch, the security setting for promiscuous mode needs to be changed to Accept.
- When a vSwitch has more than one pNIC in it, the second pNIC (even if standby in an active/passive fail over) replicates back the ARP requests, causing the Linux bridge to incorrectly update its MAC table. The workarounds are as follows:
  - Remove the second pNIC from the vSwitch; of course compromising redundancy.
  - Replace the built in vSwitch with a Nexus 1000v
- LSP-MTU is link state MTU. Default MTU is 1392. This MTU needs to be lower than overlay interface MTU for bridge-domain LSP to exchange over OTV.
- Throughput depends on packet size and OTV transport MTU depends on path size MTU. It supports an MTU range from 1,500 to 9,216 bytes. However, this configured MTU on Cisco CSR 1000V should not exceed the maximum MTU value supported on the hypervisor.
- Without IPsec, the maximum MTU supported with 1500 bytes OTV path MTU is 1472 bytes. Packets greater than that gets fragmented and the maximum fragment size on CSR is 1480 bytes. This fragment was found to be further fragmented in the network (OTV Path). This lowers the throughput significantly. CSCul56068 has been logged to support fragment size configuration CSR.

- With IPsec, the maximum MTU supported with 1500 bytes OTV path MTU is 1372 bytes. Since default LSP-MTU is 1392, this MTU needs to be lowered for LSP to exchange in bridge-domain. Default MTU of overlay interface is 1400 bytes.
- When the source VM is on the same ESXi host as the CSR OTV, throughput is extremely low. The workaround is to create VMware anti-affinity rules such that CSR will always be placed on a separate ESXi host as the VM hosts or disable TSO in the guest OS as a workaround. CSCuj55254 has been logged track this issue.
- Default MAC age out timer is 1800 seconds and ARP age out timer is 240 minutes on CSR1000v. Configure ARP age out timer to 1500 seconds so MAC addresses do not age out in bridge-domain. This can be configured on CSR interfaces using the **arp timeout 1500** command.
- AES IPsec is recommended encryption on CSR. 3DES is more CPU-intensive resulting in lower throughput.
- Premium license is required for OTV and LISP feature on CSR 1000v. Throughput depends on vCPU and RAM. For 250M throughput, it is recommended to use 4vCPU and 4G RAM.

LISP

- The following should be considered before using CSR1000v with VMware vDS. If the CSR1000v has either an interface in bridge mode or an interface in routed mode with HSRP, the vDS port-groups associated with these interface types must be configured as promiscuous. However, doing this causes packet flooding in these port-groups. This may adversely impact intra-VLAN traffic because traffic between non-routed EIDs will be flooded to the local CSR1000V, which, in turn, will redirect those packets back onto the VLAN. The net impact to local intra-VLAN L2 traffic is additional packets on the wire due to the router sending IP redirects or retransmitting L2 packets.
- This issue only applies to a vMotion or a DRaaS operation where the source and target VMs use the same MAC address. When a vMotion or DRaaS operation triggers a LISP VM Mobility event, the MAC address of the target VM in the secondary data center is learned via the Nexus 1000v on two different Veth ports, locally as a static entry and as a dynamic entry over OTV, which was the original location of the source VM prior to doing a DRaaS recovery operation. CSCul95338 could occur when the secondary data center is a VMware cluster. If the target VM and CSR1000v end up on different ESXi hosts in the VMware cluster after the DRaaS recovery operation, ARP broadcast packets destined to the CSR1000v default gateway will be dropped by the Nexus 1000v. The workaround is to either wait 3 minutes for the OTV entry to expire (the MAC aging time is 3 minutes in the Nexus 1000v) or use the **clear mac address-table dynamic vlan <VLAN ID>** command to clear the VLAN dynamic MAC address table.
- In a Windows environment where NetBIOS over IP is enabled on servers residing in the data center, an inbound L3 ACL should be applied to the CSR server-facing L3 interfaces to drop NetBIOS over IP packets. Adding an L3 ACL to drop NetBIOS over IP packets will help speed up the time it takes LISP to converge following the VM move. Refer to LISP Dynamic EID Detection, page 5-22 for further details.
- When deploying LISP VM Mobility ESM, configure a L2 MAC ACL on the Overlay to prevent the HSRP mac address from being learned at the remote data center via OTV. Refer to LISP VM Mobility ESM Prerequisites, page 5-21 for additional details on FHRP isolation.
- Dynamic EID detection in IOS-XE release 3.10S is data plane only. However, control plane dynamic EID detection may be available in a future releases.

Troubleshooting General Issues

Refer to the following topics to address general troubleshooting issues.

Network Connectivity

- Verify that there is an active and unexpired license installed on the CSR VM using 'show license.'
- Verify that the vNIC for the CSR VMs are connected to the correct physical NIC, or to the proper vswitch.
- Verify that the vNICs are configured using a supported network driver VMXNET3.

OTV

- Verify license. OTV is supported with premium license.
- Tunnel Verification - Verify OTV tunnel events using the **show otv internal event-history debug** and **show tunnel internal implicit otv brief** commands.
- Verify OTV VLAN is part of same bridge-domain on all edge devices.
- If OTV adjacency is up and bridge-domain database is not being updated, check LSP-MTU and overlay interface MTU. LSP MTU should be lower than overlay interface MTU for exchange to happen.

Packet Drops

- Verify packet drops on CSR using the **show platform hardware qfp active statistics drop** command.
- Verify packet drops on interface using the **show platform hardware qfp active interface if-name <interface> statistics** command.

IPSec

- Verify the IPSec tunnel is up with the **show crypto ipsec sa** command.

- Verify path OTV path MTU over join interface.
- Verify fragmentation bit configuration. By default, DF bit is set to 1.
- Verify access-list/interesting traffic for IPSec is defined based on OTV tunnel join interface and not based on host address.

Commands

OTV Show Commands

- show otv—Check OTV status and parameters
- show otv vlan—Check vlan involving OTV
- show otv adjacency detail—Check adjacency with end devices in OTV domain
- show otv route—Check MAC address entries for unicast routing over OTV
- show otv arp-nd-cache—Check OTV ARP entries cached on CSR
- show otv isis—Check ISIS status and configuration
- show otv isis database detail—Check OTV IS-IS internal database
- show bridge-domain—Check MAC addresses learned in Bridge-domain
- show platform software status control-processor—Check control processor status (CPU)
- show platform hardware qfp active datapath utilization—Check quantum flow processor active datapath utilization
- show platform hardware qfp active statistics drop—Check quantum flow processor active global drop statistics
- show platform hardware qfp active feature firewall drop—Check quantum flow processor active firewall drop counts

OTV Clear Commands

- clear otv arp-nd—Clear OTV arp entries
- clear otv isis adjacency * —Clear OTV adjacencies
- show platform hardware qfp active statistics clear—Clear all packet drops on CSR

OTV Debug Commands

- show otv log event/error—Check OTV logs on CSR
- debug otv overlay—Debug overlay interface activities
- debug otv adjacency—Debug otv adjacency
- debug platform hardware qfp act fea firewall datapath global all detail

Packet Capture

1. Capture Packet on any CSR interface
   - Configure capture filter on CSR in exec mode using ‘monitor capture <filter-name> match ipv4 host <ip address> any interface <interface-id> in
   - To start capture—**monitor capture <filter-name> start**

- To stop capture—**monitor capture <filter-name> stop**
- To view capture—**show monitor capture <filter-name> buffer brief**

2. If there are tail drops on CSR, check throughput and license.
3. Use the **show platform hardware qfp active statistics drop** command to view drops on CSR.

```
cvf6-t19-csr1#sh platform hardware qfp active statistics drop
--------------------------------------------------------------------------
Global Drop Stats                          Packets                  Octets
--------------------------------------------------------------------------
Disabled                                        23                    1572
Ipv4NoRoute                                      2                     118
ReassDrop                                  1109547               774759798
ReassNoFragInfo                             579637               128712054
ReassOverlap                                    36                   20760
ReassTimeout                                579870                  214988
```

LISP Commands

The following LISP commands were used and are provided for clarification.

Map Server

Check Map Server database to determine which RLOC has registered the EID-prefixes. This command list the registered dynamic EID prefixes and who last registered each prefix. In this example, West-DC xTR registered both dynamic EIDs.

```
MS-MR#show lisp site
LISP Site Registration Information

Site Name      Last      Up   Who Last             Inst     EID Prefix
               Register       Registered           ID
EastWestDC     00:00:10  yes  11.1.5.1                      8.24.0.0/16
               00:00:10  yes  11.1.5.1                      8.24.81.40/32
               00:00:10  yes  11.1.5.1                      8.24.82.40/32
```

xTRUse the following command to display the configured EID-prefix blocks, dynamic EIDs, and their associated locator-sets (RLOC). In this example, we can see that the xTR in the West-DC detected dynamic EID prefixes for both 8.24.81.40/32 and 8.24.82.40/32.

```
West-DC#show ip lisp database
LISP ETR IPv4 Mapping Database for EID-table default (IID 0), LSBs: 0x1, 3 entries

8.24.0.0/16, locator-set West-DC
  Locator   Pri/Wgt  Source     State
  11.1.5.1    1/100  cfg-addr   site-self, reachable
8.24.81.40/32, dynamic-eid vlan2481, locator-set West-DC
  Locator   Pri/Wgt  Source     State
  11.1.5.1    1/100  cfg-addr   site-self, reachable
8.24.82.40/32, dynamic-eid vlan2482, locator-set West-DC
  Locator   Pri/Wgt  Source     State
  11.1.5.1    1/100  cfg-addr   site-self, reachable
```

The routing table will display the route as having been learned via LISP.

```
West-DC#sh ip route 8.24.81.40
Routing entry for 8.24.81.40/32
  Known via "lisp", distance 10, metric 1, type intra area
  Last update from 8.24.81.40 on GigabitEthernet3, 2d03h ago
  Routing Descriptor Blocks:
  * 8.24.81.40, from 0.0.0.0, 2d03h ago, via GigabitEthernet3
      Route metric is 1, traffic share count is 1
```

PxTR

Ping the dynamic EID IP address from a device located behind the PxTR. Verify that an EID to RLOC entry exist in the local PxTR map-cache for the dynamic EID prefix.

```
pxtr#sh ip lisp map-cache
LISP IPv4 Mapping Cache for EID-table default (IID 0), 3 entries

8.24.0.0/16, uptime: 5d20h, expires: never, via static send map-request
  Negative cache entry, action: send-map-request
8.24.81.40/32, uptime: 23:00:27, expires: 00:59:33, via map-reply, complete
  Locator   Uptime    State      Pri/Wgt
  11.1.5.1  23:00:27  up           1/100
8.24.82.40/32, uptime: 00:00:01, expires: 23:59:58, via map-reply, complete
  Locator   Uptime    State      Pri/Wgt
  11.1.5.1  00:00:01  up           1/100
```

Packets from non-LISP sites to LISP EIDs will be LISP encapsulated. The CEF next hop should be the virtual interface LISP0 which is where LISP encapsulation happens on the PxTR.

```
pxtr#sh ip cef 8.24.81.40
8.24.81.40/32
  nexthop 11.1.5.1 LISP0
```