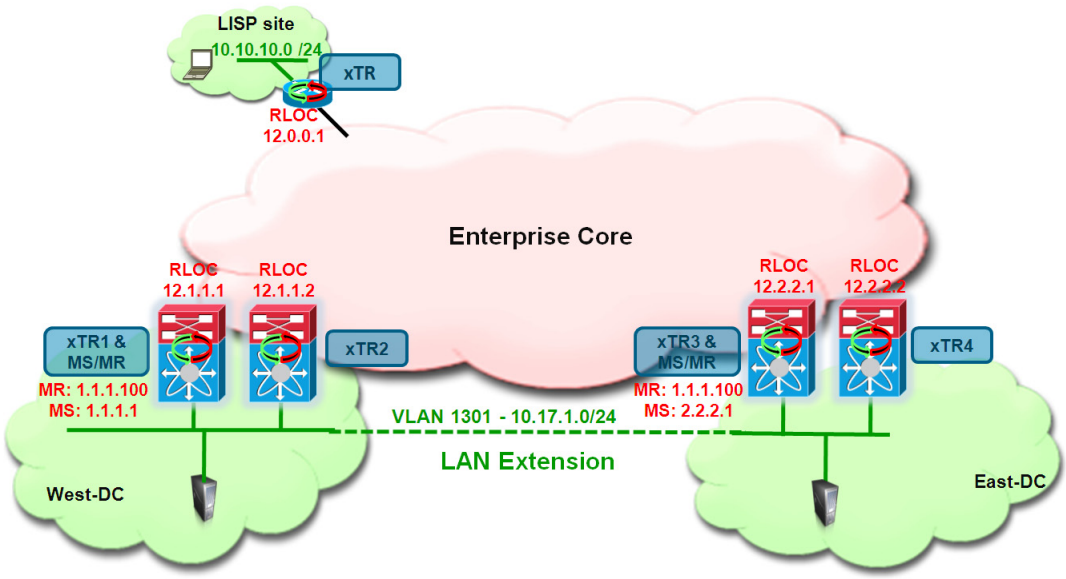
CHAPTER 4

Deploying LISP Host Mobility with an Extended Subnet

Figure 4-1 shows the Enterprise datacenter deployment topology where the 10.17.1.0/24 subnet in VLAN 1301 is extended between the West and East Data Centers using a Cisco LAN Extension solution (for example OTV or VPLS). Cisco Nexus 7000 data center switches are deployed as LISP xTRs, whereas for the remote site is positioned a Cisco IOS device hosting the 10.10.10.0/24 EID subnet. Redundant MS/MR systems co-located with the DC xTR devices are deployed in each DC (one per site); as previously mentioned, this is the typical MS/MR deployment model for an enterprise-scale LISP Host Mobility deployment.

Note The same considerations and configuration below would be valid in designs where a pair of dedicated NXOS/IOS devices were deployed as MS/MR.

Figure 4-1 LISP Host Mobility with Extended Subnet Topology



This section describes steps to configure these data center sites and remote IOS device sites as LISP sites with their corresponding EID spaces. It also highlights the required configuration to enable specific prefix hosts to move between data centers and describes how client-server and inter-DC traffic flows can be established.

LISP Host Mobility with an Extended Subnet Prerequisites

Before discussing the specific LISP Host Mobility functionality with an Extended Subnet Mode, is it important to highlight some important solution prerequisites:

- HSRP Hello messages should not be exchanged across the data center sites, allowing for the creation of an active-active HSRP setup. This is mainly done to provide an active default gateway in each physical Data Center location and avoid asymmetric traffic paths when optimizing ingress traffic with LISP (HSRP localization handles only the egress flows). Usually HSRP filtering is configured, leveraging ACLs to drop Hellos and prevent the exchange across the LAN extension connection. For more details on how to achieve HSRP localization when deploying OTV, please refer to paper below:

  http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/DCI/whitepaper/DCI3_OTV_Intro_WP.pdf

- The default gateway Virtual MAC (vMAC) and IP addresses (vIP) in both data centers should remain consistent, as the mobile workload would most likely continue to send packets to the same GW IP address after the live mobility event is completed. Virtual MAC consistency is usually achieved in Extended Subnet mode by configuring the same HSRP group associated to the same subnet in separate Data Center sites. In scenarios where the same HSRP group is not configured, an alternative approach consists of manually configuring a static vMAC as part of the HSRP configuration.

- OTV or any other deployed LAN extension solution must have multicast support over the extended L2 subnet for the proper operation of LISP Host Mobility in an extended subnet mode. This is because the LISP xTRs deployed across data center sites leverage the L2 logical connection to exchange multicast control plane messages carrying information about the different EIDs that were discovered in each site.

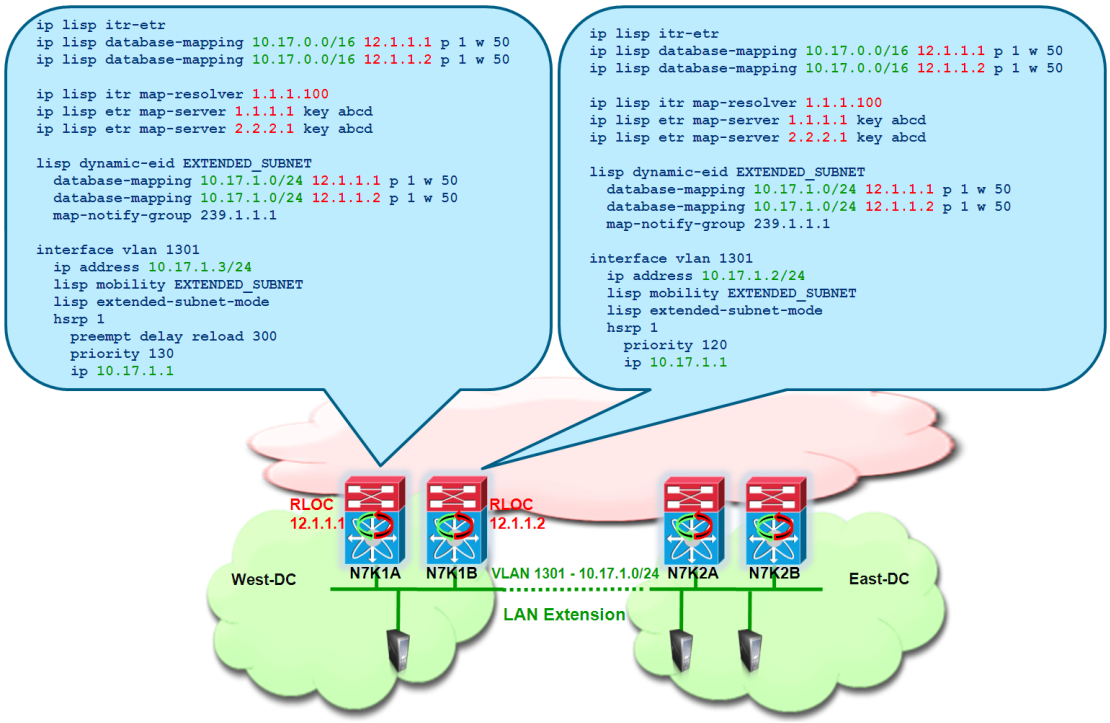
LISP Host Mobility with an Extended Subnet: Sample Config

With reference to Figure 4-1, the following are the basic configuration steps required on the various devices in the network to enable LISP Host Mobility.

Nexus 7000 N7K1A and N7K1B West DC-xTRs Configuration

The required configurations for the xTRs deployed in the West DC is shown in Figure 4-2.

Figure 4-2 LISP Host Mobility with Extended Subnet Configuration for xTRs in West DC



As it is easy to notice, the LISP configuration is pretty much identical across the two devices part of the same DC site. The following steps explanation of the various portions of the configuration.

- As first step, it is required to enable the LISP functionality on the Nexus devices and specify that they are going to perform the roles of LISP ETR (for decapsulating LISP traffic received from the L3 domain of the network) and ITR (for encapsulating LISP traffic destined to remote locations).

```
feature lisp
ip lisp itr-etr
```

- A global database mapping is then configured, including an aggregate prefix that ideally identifies all the IP subnets deployed in this specific Data Center site. Notice that this aggregate prefix may include both “mobile subnets” and “static subnets”. An additional piece of configuration is required to specifically identify these mobile subnets, as discussed later.

```
ip lisp database-mapping 10.17.0.0/16 12.1.1.1 priority 1 weight 50
ip lisp database-mapping 10.17.0.0/16 12.1.1.2 priority 1 weight 50
```

  The mapping above associates the aggregate prefix 10.17.0.0/16 to two RLOC IP addresses, which are the RLOCs identifying each xTR devices at this local DC. The recommendation is to define a loopback interface on each device as RLOC, so that communication to that IP address will remain successful as long as a valid L3 path connects the xTR to the L3 domain of the network.

  Notice also how a priority and a weight can be associated to each mapping statement: these values can be tuned to influence the inbound traffic, preferring for example the path through a specific xTR. In the configuration above the values are identical to ensure that inbound traffic can be load-balanced across both DC xTRs.

Note The definition of the global database-mapping statements is particularly important to enable routing between mobile and not mobile subnets, as will be explained in detail in the “East-West Traffic Flows Considerations” section on page 4-19.

- The next step consists then in defining the IP addresses of the Map-Servers and Map-Resolvers.

```
ip lisp itr map-resolver 1.1.1.100
ip lisp etr map-server 1.1.1.1 key abcd
ip lisp etr map-server 2.2.2.1 key abcd
```

  As already mentioned, in a typical Enterprise deployment, two devices perform the roles of MS/MR and work in a complete standalone fashion. As a consequence, on the xTRs we need to specify the IP addresses of the two Map-Servers (so that each xTR can register with both MS the EID prefixes) and the Anycast IP address of the Map-Resolver (so that the Map-Requests will be received by the MR that is closer from a routing perspective).

- A dynamic mapping is then required to identify the IP subnets to which the mobile workloads belong. When deploying LISP Host Mobility with Extended Subnet, these are the IP subnets/VLANs that are extended between DC sites.

```
lisp dynamic-eid EXTENDED_SUBNET
  database-mapping 10.17.1.0/24 12.1.1.1 priority 1 weight 50
  database-mapping 10.17.1.0/24 12.1.1.2 priority 1 weight 50
  map-notify-group 239.1.1.1
```

  In this example, the mobile subnet is a /24 prefix, which is associated to the same two RLOCs previously used for the global mapping. Priorities and weights are kept the same also in this case, to benefit of inbound load balancing for traffic destined to the mobile subnet. A multicast address (named “map-notify-group”) must also be associated to the dynamic-eid mapping. Its use will be clarified in the following sections of the document.

  Some additional considerations around the length of the network prefix specified in the dynamic-eid mapping:

  - If multiple mobile subnets are configured, it is required to define a different “lisp dynamic-eid” construct for each subnet and **not** to define a coarser prefix that includes all the mobile subnets. The multicast address of the map-notify-group can be the same across multiple constructs.
  - The mask associated to the dynamic-eid prefix should always be more specific that the one used in the global mapping statements.
  - The mask associated to the dynamic-eid prefix should match the length of the network mask of the interface (usually an SVI) where the mobility commands are configured. This is under the assumption (usually valid) that mobility should be enabled for the workloads belonging to the entire subnet.

- Finally, the LISP dynamic-eid policy configuration must be applied under the L3 interface connecting to the mobile subnet. Since the DC xTR is positioned at the aggregation layer, the L3 interface is usually a VLAN Interface (SVI). Notice how this is the only piece of configuration that is different between the two xTRs belonging to the same site (because of IP addressing and HSRP commands).

N7K1A

```
interface vlan 1301
  ip address 10.17.1.3/24
  lisp mobility EXTENDED_SUBNET
  lisp extended-subnet-mode
  hsrp 1
    preempt delay reload 300
    priority 130
    ip 10.17.1.1
```

N7K1B

```
interface vlan 1301
```

```
ip address 10.17.1.2/24
lisp mobility EXTENDED_SUBNET
lisp extended-subnet-mode
hsrp 1
  priority 120
  ip 10.17.1.1
```

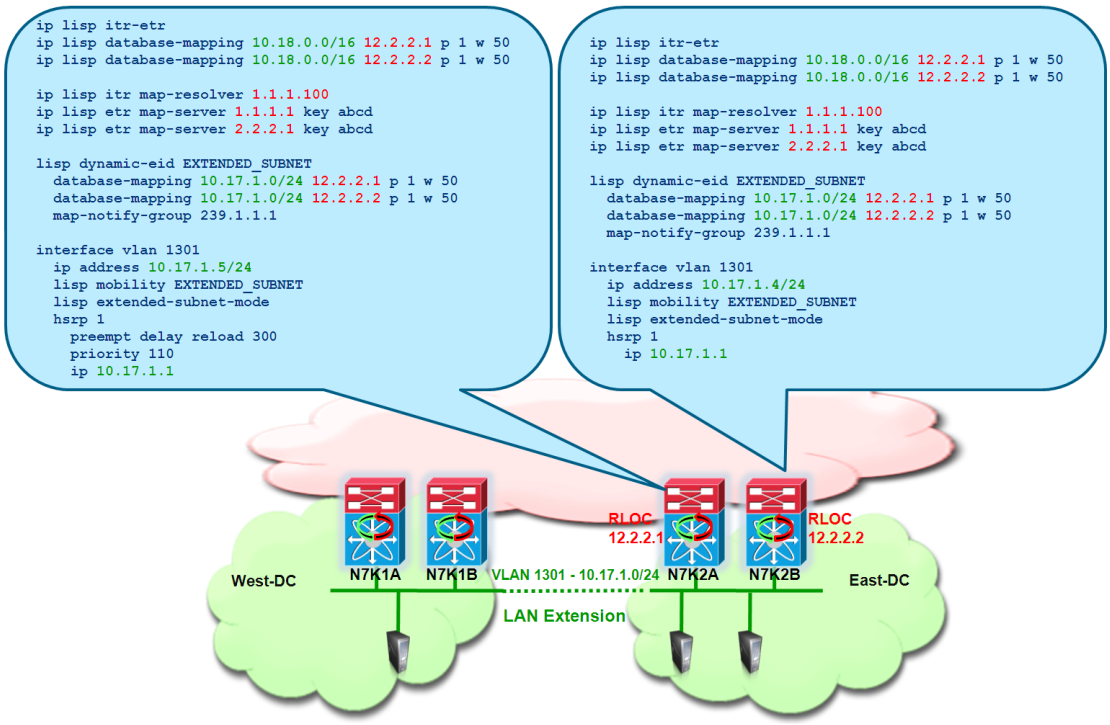
The “lisp mobility” command is used to attach the dynamic-eid construct to this interface, whereas “lisp extended-subnet-mode” is used to specify that the mobile subnet is extended across data center sites.

Note No specific multicast configuration (for example enabling PIM) is required under the SVI. Simply configuring LISP Host Mobility on the interface ensures that map-notify-group multicast frames can be sent and received successfully on the SVI interface (both for communication with the local peer xTR and for exchanging messages with the remote xTR devices across the logical LAN L2 extension between sites).

Nexus 7000 N7K2A and N7K2B East DC-xTRs Configuration

The required configuration for the xTRs deployed in the East DC is shown in Figure 4-3.

Figure 4-3 LISP Host Mobility with Extended Subnet Configuration for xTRs in East DC



Various configuration components have been explained in the previous section. The few things to notice when comparing it with the West DC xTRs are as follows:

- The global mapping is different from the one configured on the West DC xTRs. Again, the assumption here is that the IP subnets deployed in the East site can be aggregated by a unique IP prefix (10.18.0.0/16 in this case). Also, the RLOC associated to the global prefixes are now identifying the xTR devices in the East DC site.
- The prefix in the dynamic-eid mapping relative to the mobile subnet must be identical to the one defined on the West xTRs, since identifies the IP subnet extended across sites. This is the reason why 10.17.1.0/24 is specified also on the xTRs in the East DC. However, the RLOCs associated to the mobile subnet are now identifying the xTRs in the East site.

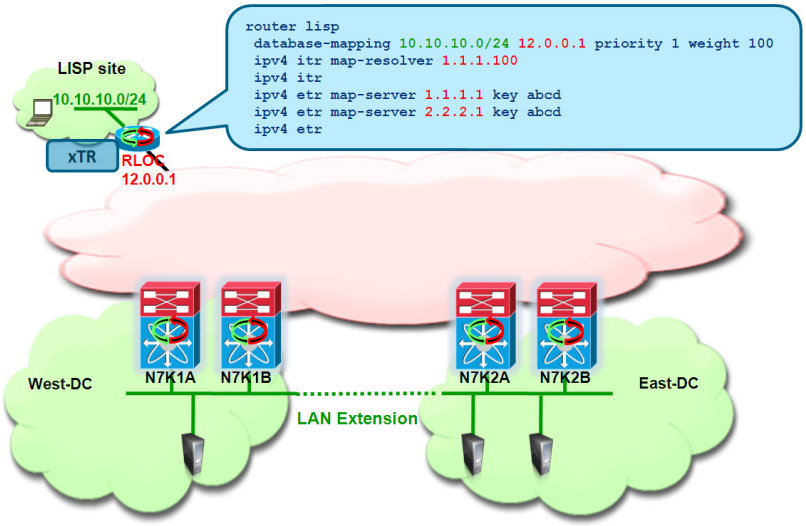
Note The same considerations would apply if we also had a part of the East DC address space that was configured as a mobile subnet extended across sites (for example 10.18.1.0/24).

- The map-notify-group associated to the dynamic-eid mapping must be identical to the one configured for the xTRs in the West site. This is because it will be used for control plane communication by all the xTRs connected to the extended subnet. The multicast communication will be established across the LAN extension logical connection (deployed with OTV, VPLS, etc.), so it is important to ensure that whatever technology is deployed to provide LAN extension services, multicast frames can be exchanged between sites.
- Notice how the HSRP VIP configured is the same (10.17.1.1) used on the West xTRs, as well as the HSRP group number ("hsrp 1") that creates the same vMAC. This is required to ensure that a workload moved to the East site can continue to communicate with the local default gateway without having to refresh its ARP cache. As previously mentioned, it is also required to filter HSRP hellos across the logical LAN extension connection to ensure that an Active HSRP device can be deployed in each data center site.

Remote Site Cisco IOS-xTR Configuration

The configuration of the branch xTR is shown in Figure 4-4.

Figure 4-4 Remote xTR IOS Configuration



Compared to the DC xTR, the configuration for the remote xTR is very simple, since there are no Host Mobility requirements for the EIDs belonging to remote locations. The explanation of the different commands is almost self-explanatory. Notice how IOS requires that the LISP configuration is added under a "router lisp" construct, in a similar fashion on how a routing protocol is usually enabled.

Step 1 Define the EID space where the clients that will communicate to the DC workloads are deployed.

```
database-mapping 10.10.10.0/24 12.0.0.1 priority 1 weight 100
```

The RLOC address associated to the EID prefix may be a loopback address (as recommended for the DC xTR devices) or, in scenarios where the remote xTR connects to separate SP connections, the IP address of the physical links toward the providers may be used as RLOCs. This last option is usually recommended when it is desirable to tune the priority and weight parameters associated to each RLOC, to influence the inbound traffic policies.

Step 2 Configure the Map-Servers and the Map-Resolver Anycast address.

```
ipv4 itr map-resolver 1.1.1.100
ipv4 etr map-server 1.1.1.1 key abcd
ipv4 etr map-server 2.2.2.1 key abcd
```

Note When deploying redundant xTR devices at the remote locations, multiple RLOCs are usually associated to the same EID subnet, similarly to that previously shown for the DC xTRs.

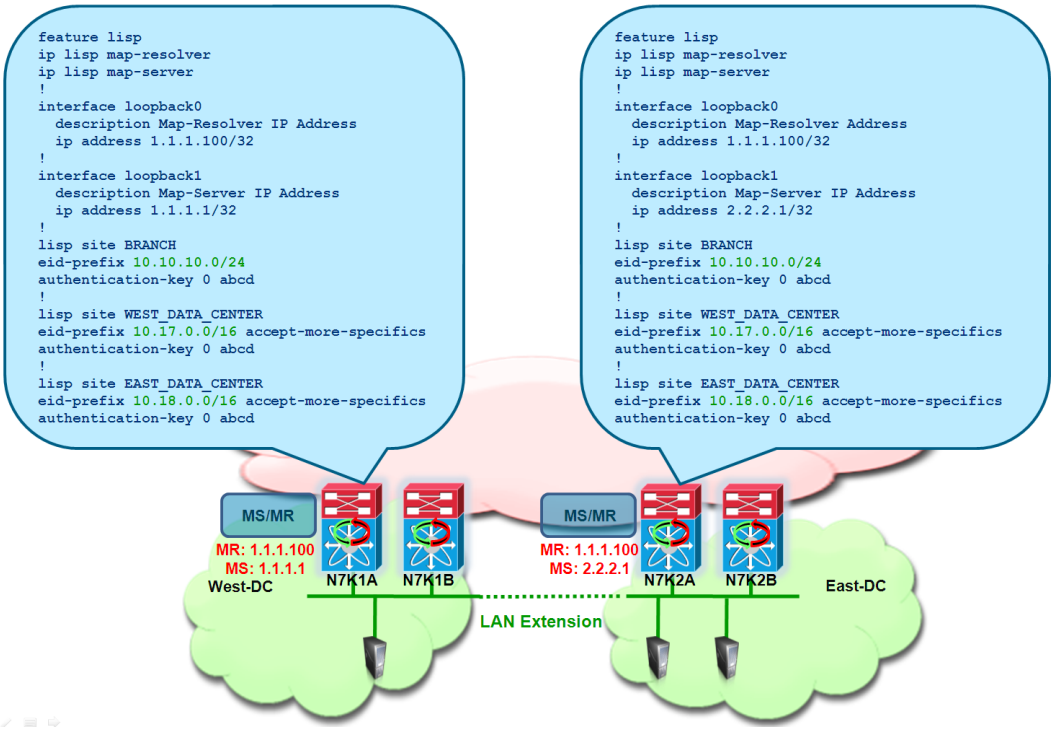
Step 3 Enable the ITR and ETR functionalities on the device.

```
ipv4 itr
ipv4 etr
```

NX-OS Map-Server and Map-Resolver Configuration

Considerations around the recommended MS/MR deployment options in LISP Enterprise deployments have been discussed in Map-Server and Map-Resolver Deployment Considerations, page 3-6. Figure 4-5 shows the configuration required when deploying the MS/MR on NX-OS platforms also configured as LISP DC xTR devices.

Figure 4-5 NX-OS and IOS MS/MR Configurations



Notice how the configuration on the two MS/MR devices is basically identical, with the only exception of the IP address used as Map-Server identifier. The different parts of the NX-OS configuration are explained below (the equivalent IOS configuration is also shown in scenarios where dedicated standalone MS/MR are deployed).

Step 1 Enable the MS and MR functionalities on the device.

NX-OS

```
feature lisp
ip lisp map-resolver
ip lisp map-server
```

IOS

```
router lisp
 ipv4 map-server
 ipv4 map-resolver
```

Step 2 Define the Loopback interfaces used as IP addresses for the Map-Resolver and Map-Server functions.

NX-OS

```
interface loopback0
  description Map-Resolver IP Address
  ip address 1.1.1.100/32
!
interface loopback1
  description Map-Server IP Address
  ip address 1.1.1.1/32
```

IOS

```
interface loopback0
  description Map-Resolver IP Address
  ip address 1.1.1.100 255.255.255.255
!
interface loopback1
  description Map-Server IP Address
  ip address 1.1.1.1 255.255.255.255
```

Both Map-Resolvers in Figure 4-5 are configured with the same IP address (Anycast IP address), so that Map-Requests originated from LISP ITR devices can be received on the MR device that is "closer" from a routing table point of view. A unique IP address is instead leveraged for the Map-Server, because the LISP ETRs must register their EID subnets with both standalone Map-Servers.

Step 3 Configure the remote branch site.

NX-OS

```
lisp site BRANCH
 eid-prefix 10.10.10.0/24
 authentication-key 0 abcd
```

IOS

```
router lisp
 site BRANCH
   authentication-key abcd
   eid-prefix 10.10.10.0/24
```

Step 4 Configure the West and East Data Center sites.

NX-OS

```
lisp site WEST_DATA_CENTER
eid-prefix 10.17.0.0/16 accept-more-specifics
authentication-key 0 abcd
!
lisp site EAST_DATA_CENTER
eid-prefix 10.18.0.0/16 accept-more-specifics
authentication-key 0 abcd
```

IOS

```
site WEST_DATA_CENTER
   authentication-key abcd
   eid-prefix 10.17.0.0/16 accept-more-specifics
 !
 site EAST_DATA_CENTER
   authentication-key abcd
   eid-prefix 10.18.0.0/16 accept-more-specifics
```

It is important to notice the "accept-more-specifics" keyword associated to the DC EID prefix. This must be configured to the sites where LISP Host Mobility is enabled, since specific /32 prefixes that are part of the larger aggregate prefix will be registered by the DC xTRs. The reasoning behind this behavior will be clarified in detail in the following section.

Remote Clients Communicating to EIDs before a Mobility Event

Assuming the LISP configuration previously described has been applied to the various devices, let's now clarify how traffic flows to and from the mobile workloads can be established. The first things to verify is that the DC and remote site xTRs are successfully registering their EID subnets with the MS. Leveraging the following commands does this:

NX-OS

```
NXOS-MS# sh lisp site
LISP Site Registration Information for VRF "default"
* = truncated IPv6 address, -x = more-specifics count

Site Name      Last        Actively    Who last        EID-prefix
               Registered  Registered  Registered
BRANCH         00:00:08    yes         12.4.3.2        10.10.10.0/24
WEST_DATA_CENT 00:00:46    yes         12.1.1.2        10.17.0.0/16-0
EAST_DATA_CENT 00:00:35    yes         12.2.2.1        10.18.0.0/16-0
```

IOS

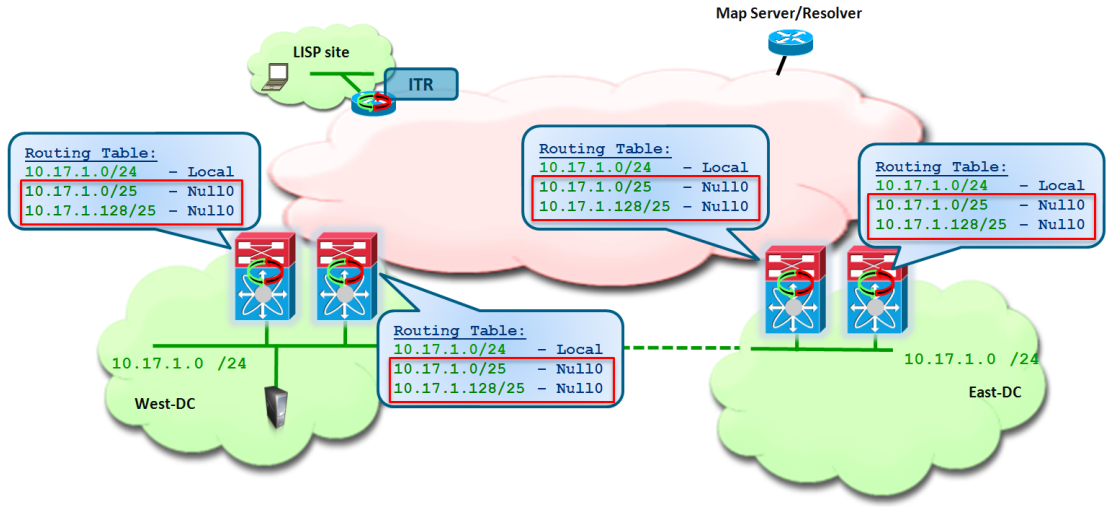
```
IOS-MS#sh lisp site
LISP Site Registration Information

Site Name      Last      Up   Who Last             Inst     EID Prefix
               Register       Registered           ID
BRANCH         00:00:08  yes  12.4.3.2                      10.10.10.0/24
WEST_DATA_CENT 00:00:14  yes  12.1.1.1    10.17.0.0/16
EAST_DATA_CENT 00:00:40  yes  12.2.2.1                      10.18.0.0/16
```

The global prefixes are registered every 60 seconds by each xTR. This implies that the timer in the "Last Registered" column should never have a value above 60 seconds. Since both xTRs send Map-Register messages independently, the IP address in the "Who last Registered" column will continuously change, identifying the xTR that sent the last map-registry meesage to the MS.

Figure 4-6 highlights the initial content of the routing tables on the DC xTRs.

Note In all the following network diagrams, the MS/MR is generically represented as a standalone device connected to the core, to ease the explanation of the discussed functionality.

Figure 4-6 Routing Table Content in DC xTRs before Workload Migration



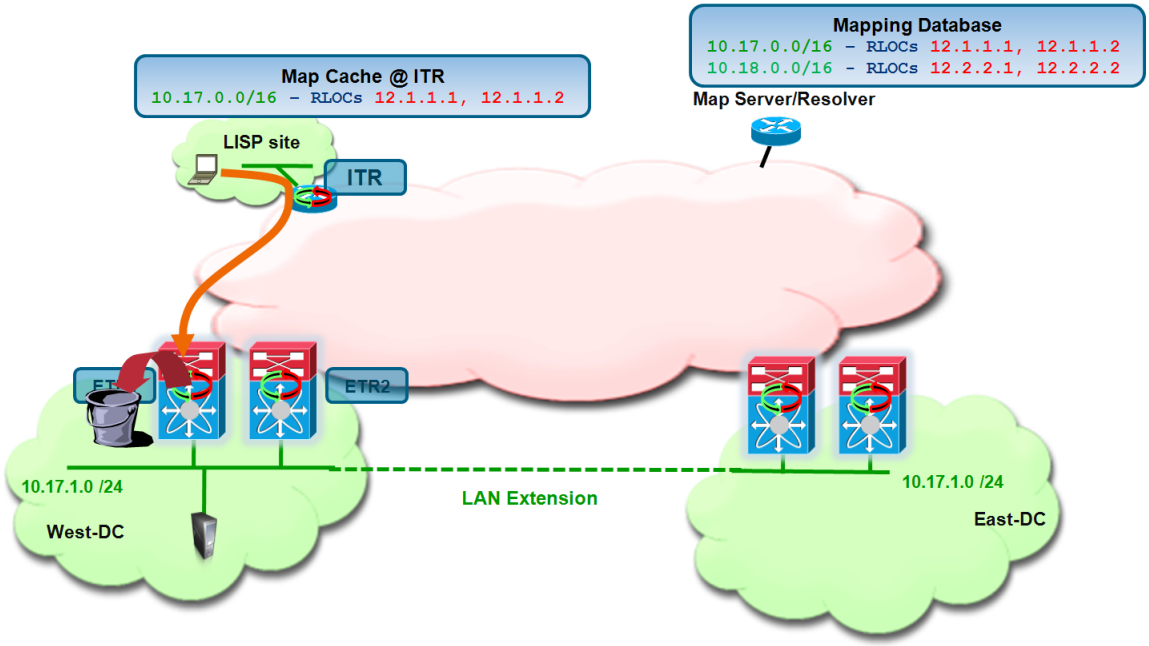
The two /25 Null0 routes are automatically installed by (and owned by) the LISP process, as soon as the "lisp extended-subnet-mode" command is configured for the VLAN 1301. These routes are needed to allow the xTR devices to punt to the CPU the first packet received from a mobile workload, so that the xTR can dynamically discover it and register the specific /32 prefix with the Map-Server. In order for the punting to CPU to happen, a sort of Unicast Reverse Path Forwarding (URPF) logic is enabled automatically on an interface configured for LISP Host Mobility. This means that the following sequence of events happens:

1. The workload generates an IP packet that is received by the xTR on the interface enabled for LISP Host Mobility (SVI 1301).
2. The LISP logic checks the IP address of the received packet. In order for the check to be successful, the packet must be received on the interface that the router would use to forward the return packet. This is not the case in this example because of the existance of /25 Null0 routes that are more specific than the /24 subnet directly connected via the SVI.
3. Because the packet matches one of the /25 Null0 routes owned by LISP and that are more specific of the /24 subnet directly connected via the SVI, the packet is punted to the CPU triggering the EID dynamic discovery.

The two /25 routes in this example are created to cover the entire /24 subnet. This is because with the configuration previously discussed the desire is to enable mobility for the hosts belonging to the full /24 mobile subnet.

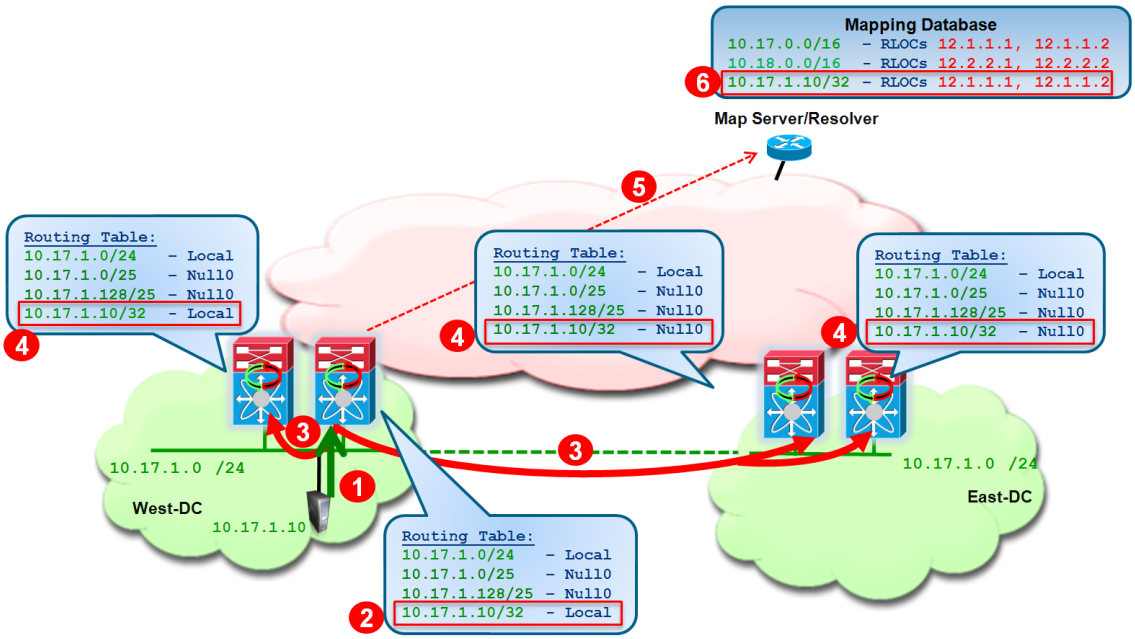
The existence of these /25 Null0 routes has also another interesting side-effect: since they are more specific than the local /24 subnet, traffic originated from the Layer 3 domain (for example from a remote client) will never be routed to the mobile subnet, but it will be dropped first because of the /25 Null0 routes. In the current LISP Host Mobility with Extended Subnet implementation this is a basic principle to keep in mind: communication to a workload belonging to a LISP mobile subnet can only be established after that workload has been dynamically discovered by the xTR, independently from which DC site the workload belongs to (Figure 4-7).

Figure 4-7 Dropping Client Initiated Traffic Flows



Hence, to establish a successful client-server communication, we first need to discover the EID in the West DC location, following the sequence of events shown in Figure 4-8.

Figure 4-8 Initial Discovery of an EID in West DC



1. The workload sends out an IP packet that is intercepted by one of the xTRs and punted to the CPU, triggering a data-plan driven EID discovery event. Notice that for this to happen, the packet must be an IP packet containing the source IP address of the EID. Once the EID is discovered, it is added to the dynamic-eid table of the discovering xTR, as shown below.

```
N7K1A# show lisp dyn summary
```

```
LISP Dynamic EID Summary for VRF "default"
* = Dyn-EID learned by site-based Map-Notify
Dyn-EID Name   Dynamic-EID      Interface     Uptime    Last      Pending
                                                       Packet    Ping Count
EXTENDED_SUBNE 10.17.1.10       Vlan1301      00:01:54  00:01:54  0
```

2. LISP installs in the routing table of the discovering xTR a local /32 route associated to the EID.

```
N7K1A# show ip route 10.17.1.10
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]

10.17.1.10/32, ubest/mbest: 1/0, attached
    *via 10.17.1.10, Vlan1301, [250/0], 00:04:00, am
     via 10.17.1.10, Vlan1301, [251/0], 00:04:00, lisp, dyn-eid
```

Notice that in the sample above there are actually two /32 entries associated to the EID address. The highlighted one is installed by LISP after the EID discovery, whereas the first is added by the Adjacency Manager (AM) once the IP address of the EID is successfully added to the local ARP table.

3. The discovering xTR sends out a map-notify-group multicast message (using the previously configured 239.1.1.1 multicast group) that reaches the other local xTR and via the LAN extension connection also the two remote xTRs.

4. After reception of the map-notify-group message, two different things happen:

   – On the other local xTR in the West site, a local /32 route is added to the routing table for the discover EID. This is required because inbound traffic from a remote xTR may be delivered to either DC devices, so it is important that both have local routing information to allow traffic to the destination EID (avoiding the dropping of traffic caused by the /25 Null0 entries shown in Figure 4-8).

```
N7K1B# show lisp dyn summary
LISP Dynamic EID Summary for VRF "default"
* = Dyn-EID learned by site-based Map-Notify
Dyn-EID Name   Dynamic-EID      Interface     Uptime    Last      Pending
                                                       Packet    Ping Count
EXTENDED_SUBNE*10.17.1.10       Vlan1301      00:07:00  00:00:52  0

N7K1B# show ip route 10.17.1.10
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]

10.17.1.10/32, ubest/mbest: 1/0, attached
    *via 10.17.1.10, Vlan1301, [251/0], 00:03:40, lisp, dyn-eid
```

   Notice how the entry in the dynamic-eid table is marked with an *, meaning that it was added to the table after the reception of a Map-Notify message.

Note When deploying vPC to connect the local xTRs to the access layer devices, the discovery of the EIDs can be performed by either LISP devices, depending on how the traffic is hashed by the access layer switch on the port-channel links. However, at steady state (i.e. after few seconds from the actual EID discovery) all the entries in the dynamic-eid table should be marked with the * on the HSRP Standby device and without the * on the HSRP Active device. This also means that it is the HSRP Active node that is responsible for registering periodically these entries in the Map-Server database.

- On the xTRs in the East site a /32 Null0 routes is installed for the EID since the workload has been discovered in a different DC site. The use of these /32 Null0 entries will be clarified in the following section.

```
N7K2A# show ip route 10.17.1.10
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]

10.17.1.10/32, ubest/mbest: 1/0, attached
    *via Null0, [252/0], 00:27:33, lisp, dyn-eid
```

5. The discovering xTR registers the /32 prefix with the Map-Server by sending a Map-Registry control plane message.
6. The Map-Server adds the specific EID information to the database. The specific EID information is displayed differently depending if the MS is deployed on NX-OS or IOS systems.

NX-OS

```
NXOS-MS# show lisp site
LISP Site Registration Information for VRF "default"
* = truncated IPv6 address, -x = more-specifics count

Site Name      Last        Actively    Who last        EID-prefix
               Registered  Registered  Registered
BRANCH         00:00:29    yes         12.4.3.2        10.10.10.0/24
WEST_DATA_CENT 00:00:35    yes         12.1.1.1        10.17.0.0/16-1
EAST_DATA_CENT 00:00:23    yes         12.2.2.1        10.18.0.0/16-0

NXOS-MS# show lisp site WEST_DATA_CENTER detail
LISP Site Registration Information for VRF "default"
* = truncated IPv6 address, -x = more-specifics count

Site name:   "WEST_DATA_CENTER"
Description: none configured
Allowed configured locators: any
  Configured EID-prefix: 10.17.0.0/16, instance-id: 0
    More-specifics registered: 1
    Currently registered:      yes
    First registered:          00:22:36
    Last registered:           00:00:04
    Who last registered:       12.1.1.1
    Routing table tag:         0
    Proxy Replying:            no
    Wants Map-Notifications:   yes
    Registered TTL:            1440 minutes
    Registered locators:
      12.1.1.1 (up), priority: 1, weight: 25
      12.1.1.2 (up), priority: 1, weight: 25
    Registration errors:
```

```
    Authentication failures: 0
    Allowed locators mismatch: 0
  More-specific EID-prefix: 10.17.1.10/32, instance-id: 0
    Currently registered:      yes
    First registered:          00:20:29
    Last registered:           00:00:04
    Who last registered:       12.1.1.1
    Routing table tag:         0
    Proxy Replying:            no
    Wants Map-Notifications:   yes
    Registered TTL:            1440 minutes
    Registered locators:
      12.1.1.1 (up), priority: 1, weight: 25
      12.1.1.2 (up), priority: 1, weight: 25
    Registration errors:
      Authentication failures: 0
      Allowed locators mismatch: 0
```

The symbol "-1" next to the 10.17.0.0/16 EID subnet is a counter highlighting how many specific /32 EID prefixes have been discovered so far. Leveraging "**show lisp site detail**" is then possible to verify more detailed information related to the discovered EIDs.

Similar information can be retrieved on the IOS Map-Server:

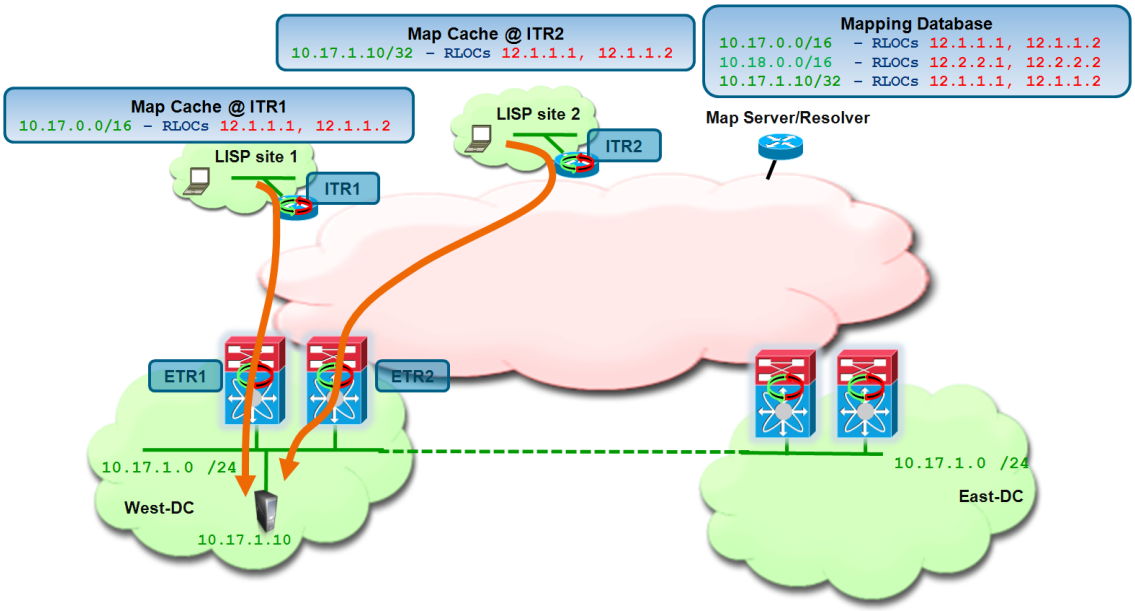
IOS

```
IOS-MS#sh lisp site
LISP Site Registration Information

Site Name      Last      Up   Who Last             Inst     EID Prefix
               Register       Registered           ID
BRANCH         00:00:25  yes  12.4.3.2                      10.10.10.0/24
WEST_DATA_CENT 00:00:31  yes  12.1.1.2                      10.17.0.0/16
               00:00:21  yes  12.1.1.1                      10.17.1.10/32
EAST_DATA_CENT 00:00:40  yes  12.2.2.1                      10.18.0.0/16

IOS-MS#sh lisp site name WEST_DATA_CENTER
Site name: WEST_DATA_CENTER
Allowed configured locators: any
Allowed EID-prefixes:
  EID-prefix: 10.17.0.0/16
    First registered:     00:47:44
    Routing table tag:    0
    Origin:               Configuration, accepting more specifics
    Merge active:         No
    Proxy reply:          No
    TTL:                  1d00h
    State:                complete
    Registration errors:
      Authentication failures:   51
      Allowed locators mismatch: 0
    ETR 12.1.1.2, last registered 00:00:56, no proxy-reply, map-notify
                  TTL 1d00h, no merge, nonce 0x00000000-0x00000000
                  state complete
      Locator   Local  State      Pri/Wgt
      12.1.1.1  yes    up           1/25
      12.1.1.2  no     up           1/25
  EID-prefix: 10.17.1.10/32
    First registered:     00:27:28
    Routing table tag:    0
    Origin:               Dynamic, more specific of 10.17.0.0/16
    Merge active:         No
    Proxy reply:          No
    TTL:                  1d00h
```

```
State:                complete
Registration errors:
  Authentication failures:   0
  Allowed locators mismatch: 0
ETR 12.1.1.1, last registered 00:00:56, no proxy-reply, map-notify
              TTL 1d00h, no merge, nonce 0x00000000-0x00000000
              state complete
  Locator   Local  State      Pri/Wgt
  12.1.1.1  yes    up           1/25
  12.1.1.2  no     up           1/25
```

At this point, a client situated in a remote location is able to successfully communicate to the EID. As shown in Figure 4-9, traffic destined to other less specific /16 prefix hosts is steered to the West DC based on ITR1 having that map-cache entry installed, while traffic destined to the more-specific /32 reaches the West DC based on that map-cache entry, as shown for ITR2. Notice that since the map-cache entries on the ITRs list both West DC RLOCs for each prefix, traffic inbound to the West DC will be load-balanced across both ETRs (the hashing value is calculated based on L3 and L4 flow parameters).

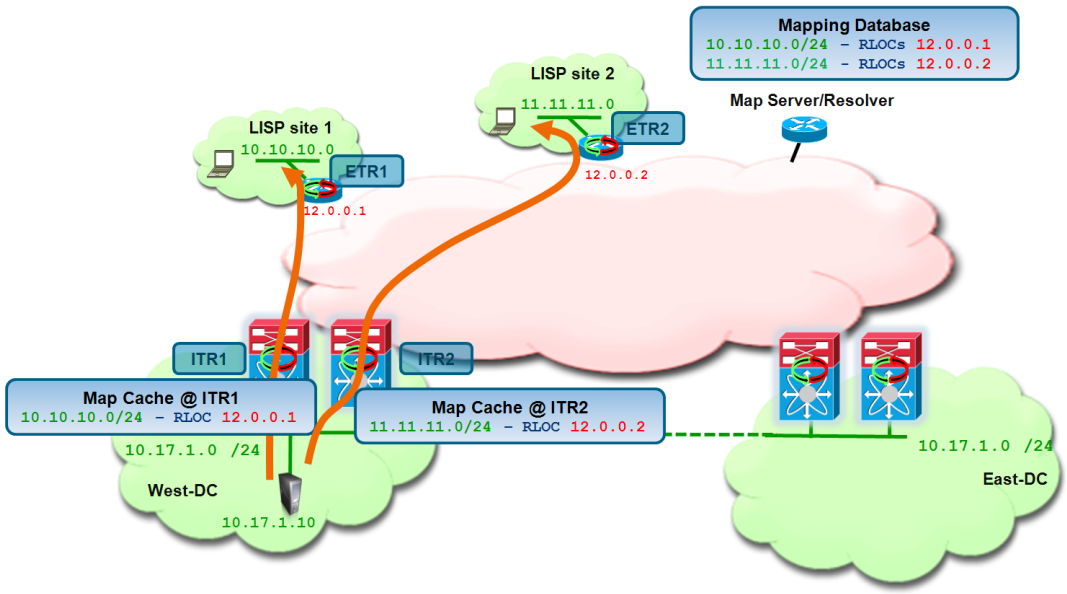
Figure 4-9 Successful Establishment of Client-Server Traffic Flows



The example above focuses only on the client-to-server traffic flows (prior to a host move). For return traffic concerns, two different scenarios are possible:

1. The branch subnets where the clients are deployed are injected in the core of the network. In this case, the DC xTRs will receive routing information about the branch subnets and as a consequence traffic will be natively forwarded (not LISP encapsulated).\
2. The branch subnets are EIDs as well (hence not injected in the routing protocol running in the core of the network). In this case, communication between the DC EIDs and the clients must happen through LISP. The mechanism is similar to the one discussed above, with the only difference that now the remote LISP devices become ETRs and the DC devices play the role of ITRs. In this case, the DC xTRs have to populate their map-cache tables to be able to encapsulate traffic to the remote locations (Figure 4-10).

Figure 4-10 Establishing of Server to Clients Communication

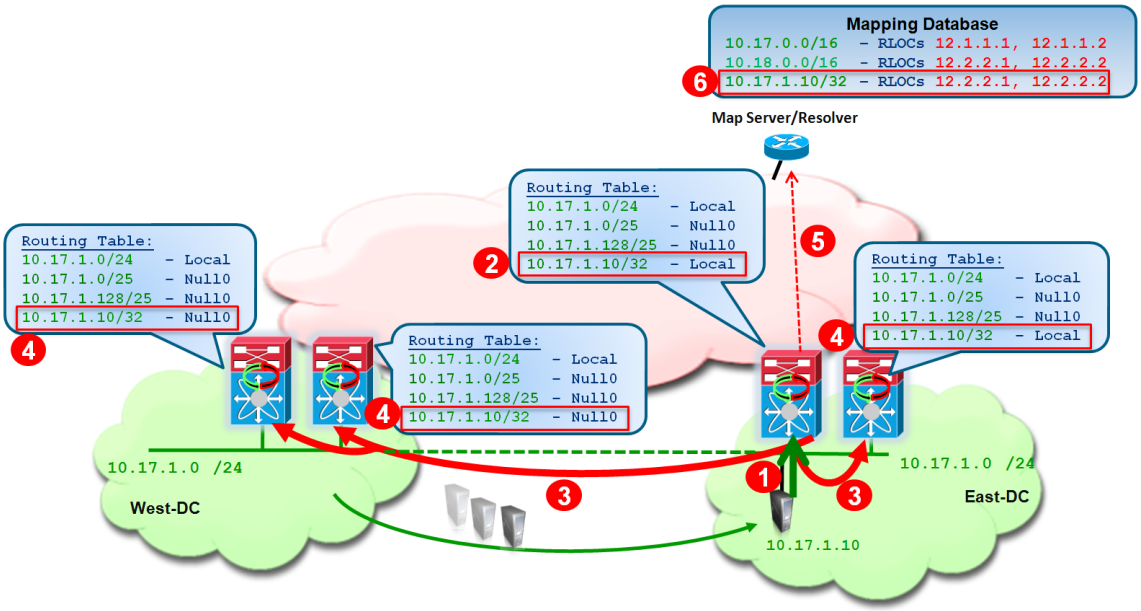


Note Server-client traffic flows will also be balanced across both of DC xTRs, assuming they are connected to the edge of the network leveraging an active/active mechanism like vPC. In a STP based POD topology, all the traffic flows will be instead handled by the same xTR (the HSRP active one).

Remote Clients Communicating to EIDs after a Mobility Event

Figure 4-11 highlights the sequence of events to move a workload from the West DC to the East DC and the consequent updates of the LISP data structures on the MS and on the DC xTRs.

Figure 4-11 Update of LISP Data Structures after Workload Migration

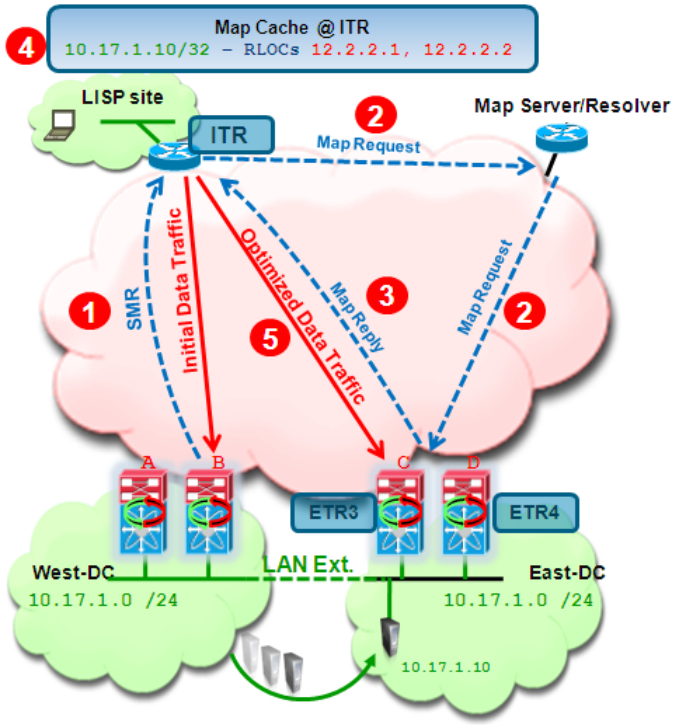


1. The workload is migrated from the West DC LAN extension to the East DC LAN extension. The workload VM retains its IP address and MAC address. The workload sources an IP packet that reaches one of the two DC xTR devices in the new (East) DC. This triggers the match against the /32 Null0 route (which is owned by the LISP process and was installed when the EID was initially discovered in the West site), resulting in the first packet failure event, as previously described, and causing the discovery of the dynamic-EID.
2. The xTR that discovered the EID replace the /32 Null0 route associated to the EID with a valid local route. This is important to allow exchange of traffic (like ARP) to happen with the workload.
3. The discovering xTR sends out a Map-Notify-Group message on the interface (SVI 1301) where the discovery of the EID happened.
4. The multicast message reaches the peer xTR in the East DC (which also install a valid /32 route for the EID), and to the two xTRs in West DC, causing them to add local /32 Null0 routes (since the EID is now available in a different DC site).
5. The discovering xTR sends a Map-Register messages for the /32 EID address to the Map-Server.
6. The Map-Server updates the entry for the specific EID, replacing the original RLOCs (12.1.1.1 and 12.1.1.2) associated to the xTRs in the West DC with the new RLOCs (12.2.2.1 and 12.2.2.2) assigned to the xTRs in the East DC.

The procedure above updates the information in the DC xTR devices and in the mapping database. To establish successful communication between the remote client and the migrated workload is necessary to complete a further step: updating the map-cache information in the map-cache of the remote xTR devices. New connection establishements will receive the correct (new) map-cache entries, However, for existing flows, even after the move of the workload, the map-cache of the remote xTR may still hold old mapping information, associating the EID with the RLOCs of the xTRs in the West DC site.

The consequence is that data packet destined to the EID, will be sent toward the West DC. Once one of the xTRs in the West DC site receives the first LISP encapsulated packet after the move and decapsulates it, it will perform a routing lookup and find that the destination address is associated to the Null0 route installed at step 4 above. Because of this, the packet will be dropped and punted to the CPU to be handled by the LISP process. This will allow the steps shown in Figure 4-12 to follow:

Figure 4-12 Updating the Remote xTR Map-Cache Entry



1. The LISP process on the xTR receiving the first data packet creates a control plane message (called Solicit-Map-Request – SMR) and sends it to the remote ITR that generated the packet. This is to inform the remote ITR that there is a need to refresh its map-cache information because the destination workload has been moved to a different location. It is important to notice that the SMR message will be created only because there is a /32 Null0 entry associated to the EID (i.e. hitting the more generic /25 Null0 won't cause that). Hence, it is critical to verify that the xTRs in the original site have the entry populated once the EID is discovered in the East DC.
2. The remote ITR receives the SMR and send a new Map-Request for the desired destination (10.17.1.10) to the Map-Server. The Map-Request is forwarded by the Map-Server to the DC xTR in the East site that registered last the /32 EID address.
3. The DC xTR in the East DC replies with updated mapping information to the remote ITR.
4. The remote ITR updates the information in its map-cache, replacing the old RLOCs with the RLOCs of the xTRs deployed in the East site (12.2.2.1 and 12.2.2.2).
5. Traffic is now optimally steered toward the East DC site.

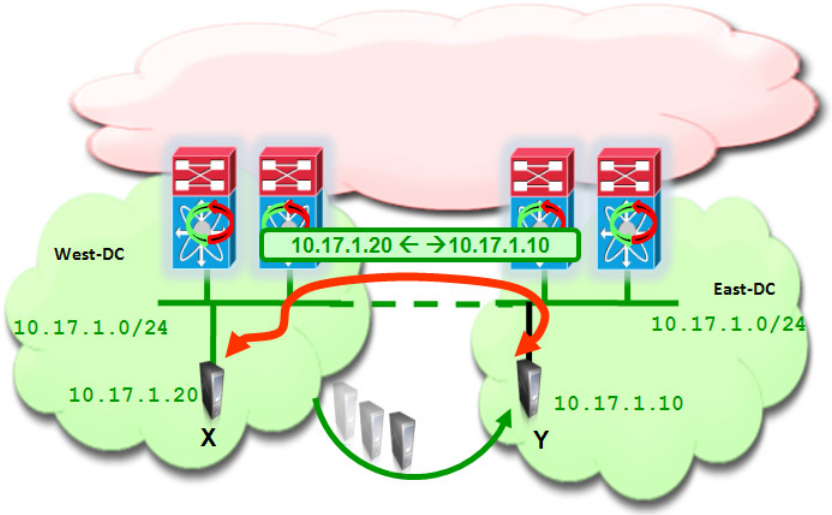
East-West Traffic Flows Considerations

The discussion in the previous sections focused on the establishment of traffic flows between clients situated in remote locations (behind an ITR) and a workload in the DC (before and after a workload migration). The next step is clarifying how communication can instead happen between a workload migrated to the East DC and resources still deployed in the original West site. From that point of view, there are two types of traffic flows to consider: intra-subnet and inter-subnets.

Intra-Subnet Traffic Flows

The establishment of intra-subnet communication between two workloads part of the same IP subnet but connected in separate DC sites is shown in Figure 4-13.

Figure 4-13 Intra-Subnet Communication Between DC Sites



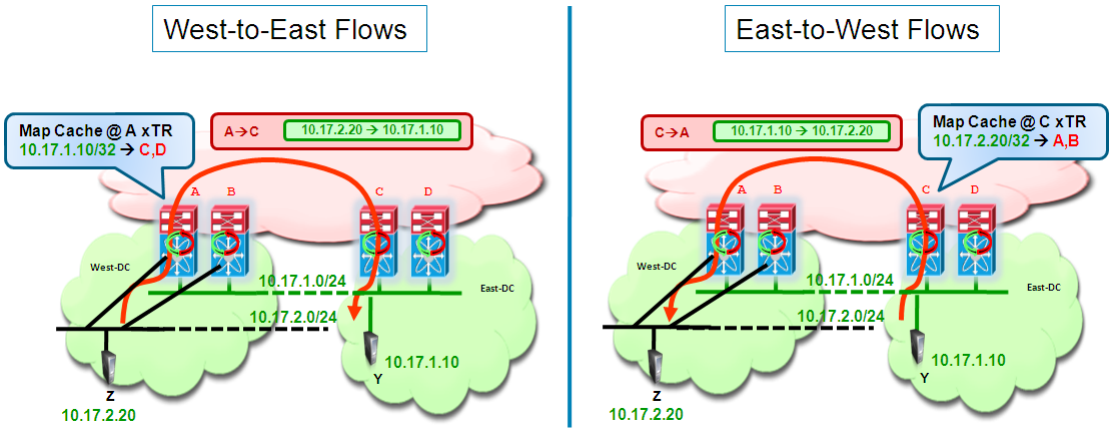
Since the IP subnet is extended between the West and East DC sites, the intra-subnet communication happens at L2 across that logical connection, without requiring any LISP functionality. This means that intra-subnet traffic flows can be established even before the EIDs are discovered by the LISP xTRs, and as a consequence these packets will never trigger a dynamic discovery event.

Inter-Subnets Traffic Flows

Let's assume the desire is to establish inter-subnet communication between a host Z in subnet 10.17.2.0/24 in the West site and the previously discussed workload Y (10.17.1.10) that was migrated to the East DC. There are two scenarios to consider:

1. 10.17.2.0 is also a mobile subnet, which means it is extended between the DC sites. This scenario is highlighted in Figure 4-14.

Figure 4-14 Traffic Flows between Extended Subnets



In this case, when Z wants to send a packet to Y, it first sends it to its local default gateway, positioned on one of the DC xTR in the West site. The xTR tries to route the packet to Y, but hits the Null0 route installed when Y was discovered in the East site. This punts the packet to the CPU and allows triggering a LISP Map-Request to the Map-Server. Once the xTR receives valid information for Y, it will start encapsulating traffic to it. This means that traffic in the Z-to-Y direction will flow LISP encapsulated across the L3 infrastructure.

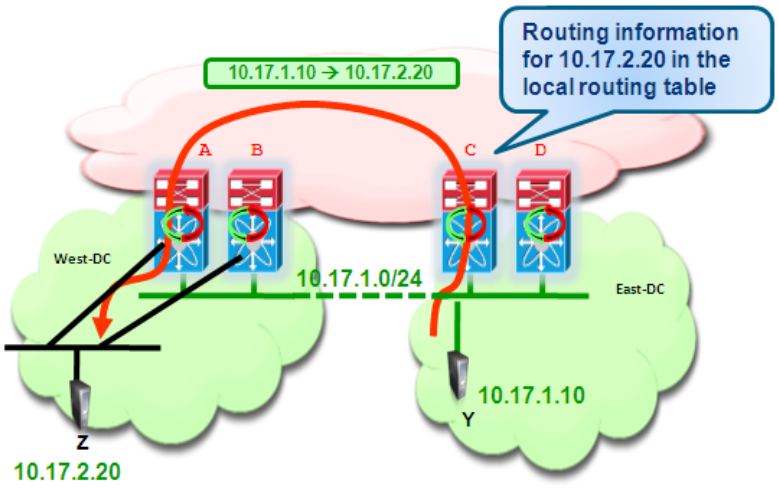
Similarly, when Y tries to send packets back to Z, the LISP control plan is triggered (a Null0 is installed for Z in the East DC xTRs) and data packet are then LISP encapsulated and sent across the L3 infrastructure.

2. 10.17.2.0 is not a mobile subnet, which means it is only defined in the West DC and no LISP mobility commands are configured under the SVI associated to it. In this case, the communication in the Z-to-Y direction happens identically to the scenario above. However, in the current implementation, an xTR performs a check on the source IP address before triggering the LISP control plane. This means that Z must be an EID to be able to communicate to Y via LISP. Since the subnet to which Z belongs is not a mobile subnet, this essentially means that Z needs to be part of the global mapping defined with the "ip lisp database-mapping" command. This is indeed an additional reason why the global mapping should always be defined, covering all the non mobile subnets deployed in a specific DC site.

   For what concerns the communication in the Y-to-Z direction, two scenarios are possible:

   1. Z subnet is advertised in the protocol used to exchange routing information between the West and East DC sites. In this case, the xTR in the East site will have valid routing information for the subnet Z belongs to and traffic will hence be natively routed back (Figure 4-15).

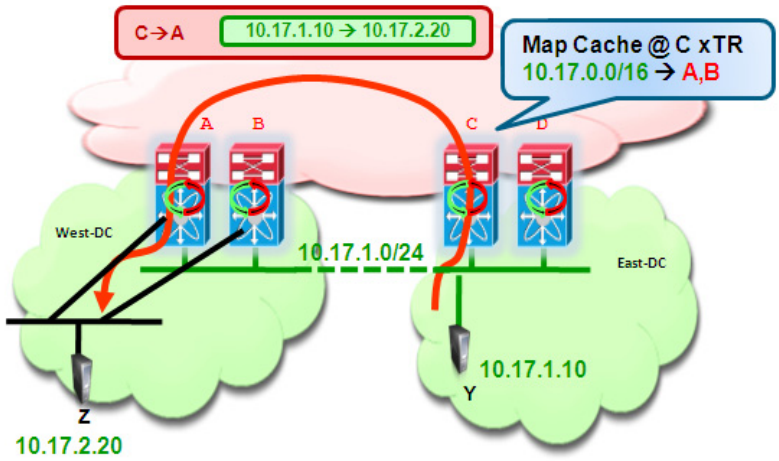
Figure 4-15 Natively Routing Traffic from East to West DC



2. Z subnet is not injected in the routing protocol, because the goal is to make it only reachable via LISP encapsulation. In this case, when the packet from Y reaches the East xTR, a Map-Request for Z will be sent out and the reply from the Map-Server will be the global prefix covering Z subnet (10.17.0.0/16 in our configuration example), associated to the RLOCs of the DC xTRs belonging to the West DC site. Traffic will then be LISP encapsulated and sent toward that site across the L3 network infrastructure and eventually routed to Z (Figure 4-16).

Note /25 Null0 routes (10.17.2.0/25 and 10.17.2.128/25) are not installed in this case, since the subnet is not configured as mobile. This means that traffic can be successfully decapsulated and routed to Z.

Figure 4-16 LISP Encapsulating Traffic from East to West DC



Summary

In summary, the LISP Host Mobility solution with Extended Subnet provides the following key benefits:

- Automated move detection for workloads migrated between DC sites.
- Dynamic-EID discovery in multi-homed data centers.
- Inbound Path Optimization functionality, which provides automatic traffic redirection and avoids the traffic triangulation problem.
- Connections maintained across move: this makes this deployment suitable to solve "hot" migration use cases (like vMotion).
- No routing re-convergence.
- No DNS updates required, so the traffic optimization can happen transparently to the clients.