Exhibit E - U.S. Patent No. 10,020,961 ("'961 Patent")

Accused Instrumentalities: Cisco products, systems, and software that implement Cisco ACI Multi-Site overlay functionality, including versions and variations thereof since the issuance of the '961 patent. This includes, without limitation:

- Cisco ACI Multi-Site deployments in which two or more ACI fabrics ("sites") are interconnected and managed as a single virtualization domain;
- Cisco Nexus Dashboard Orchestrator (NDO) and the underlying Nexus Dashboard management platform, which provide centralized control over Multi-Site policies, VNID and bridge-domain configuration, and intersite VXLAN connectivity;
- Cisco Application Policy Infrastructure Controller (APIC) software and controller clusters deployed at each ACI site and programmed by NDO to instantiate bridge domains, VNIDs, endpoint groups, and Multi-Site behavior;
- Cisco ACI fabric switches operating in Multi-Site mode, including Multi-Site–capable spine and leaf switches (e.g., Cisco Nexus 9000 series switches used as spines, leafs, and border/BL nodes) that perform VXLAN encapsulation/decapsulation, O-UTEP/O-MTEP tunnel-endpoint processing, and intersite BUM and unicast forwarding as described in the Cisco ACI Multi-Site documentation;
- Cisco ACI Multi-Site deployments in which a Cisco ACI Multi-Pod fabric is used as an ACI "site" within the Multi-Site architecture, including the Multi-Site–capable spine switches and border/BL nodes that participate in Multi-Site VXLAN/EVPN tunnels.

The Accused Instrumentalities further include any other Cisco hardware or software that, alone or in combination with the above, implements the Cisco ACI Multi-Site overlay control plane (MP-BGP EVPN, COOP, VNID/BD instantiation) and data plane (VXLAN encapsulation between O-UTEP/O-MTEP tunnel endpoints) as described in the Cisco ACI Multi-Site whitepaper and related documentation.

Issued Claim(s)	Public Documentation
[1 pre] A network virtualization apparatus comprising:	To the extent that the preamble is limiting, the Accused Instrumentalities include a network virtualization apparatus. For example, as evidenced below, the accused Cisco ACI Multi-Site products implement a VXLAN/EVPN-based overlay that virtualizes Layer 2 and Layer 3 networks across multiple independent fabrics (sites). Figure 6 depicts VTEP IPs, VNIDs, and Class IDs carried across sites under the MP-BGP EVPN control plane, with Nexus Dashboard Orchestrator coordinating policy across the virtualized network. Cisco's own materials also

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
001

	describe VXLAN as a "network virtualization technology" that creates virtual Layer 2 networks over a Layer 3 infrastructure. *See, e.g.*, **Note:** Starting from Nexus Dashboard Orchestrator Software Release 4.0(1), a new deployment model is supported, allowing you to deploy NDO to manage up to 100 "autonomous fabrics." In that specific use case, there is no VXLAN EVPN intersite connectivity between the fabrics that are part of the Multi-Site domain, and the Orchestrator essentially becomes a single pane of glass from where to provision configurations to all those sites. Layer-3-only communication is possible between the fabrics, leveraging the L3Out data path. For more information on the deployment of NDO with "autonomous fabrics," please refer to the "Layer-3-only connectivity across sites" section. The use of site-to-site VXLAN encapsulation greatly simplifies the configuration and functions required for the intersite IP network. It also allows network and policy information (metadata) to be carried across sites, as shown in Figure 6.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
002

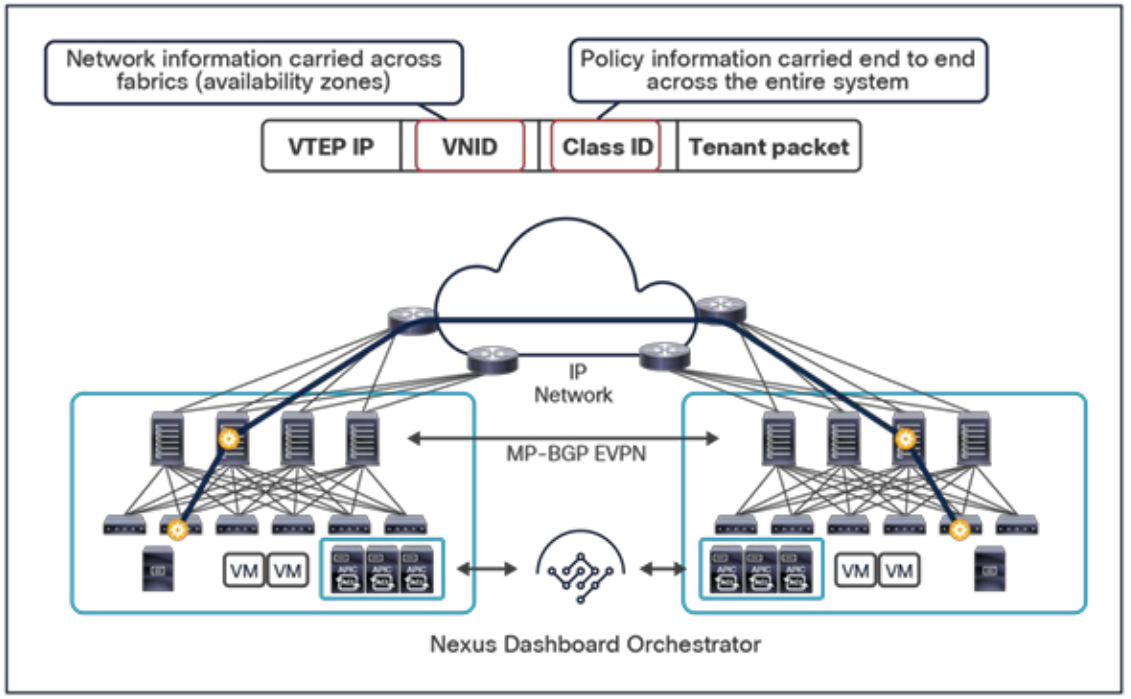


Figure 6.
Carrying network and policy information across sites

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
003

	• Intersite control plane: Endpoint reachability information is exchanged across sites using a Multiprotocol-BGP (MP-BGP) Ethernet VPN (EVPN) control plane. This approach allows the exchange of MAC and IP address information for the endpoints that communicate across sites. MP-BGP EVPN sessions are established between the spine nodes deployed in separate fabrics that are managed by the same instance of Cisco Nexus Dashboard Orchestrator, as discussed in more detail in the section "Cisco ACI Multi-Site overlay control plane." • Intersite data plane: All communication (Layer 2 or Layer 3) between endpoints connected to different sites is achieved by establishing site-to-site Virtual Extensible LAN (VXLAN) tunnels across a generic IP network that interconnects the various sites. As discussed in the section "Intersite connectivity deployment considerations", this IP network has no specific functional requirements other than the capability to support routing and increased Maximum Transmission Unit (MTU) size (given the overhead from the VXLAN encapsulation). https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
004

How It Works

How does virtualization work?

Virtualized services can run either within traditional virtual machines or in containers. A virtual machine has a guest OS and the service. A container is a complete standalone package, with the service and all its dependencies on the system.

What are network virtualization's benefits?

Virtualization enables new network services to be created and deployed in minutes, instead of in days. These services can be scaled up or down quickly to match demand.

Organizations of all sizes can benefit from automation. You can save time, eliminate manual work, and reduce OpEx. Virtualized services use hardware resources more efficiently and only as necessary, leading to savings in CapEx.

https://www.cisco.com/site/us/en/learn/topics/networking/what-is-virtualization.html.

Nested VXLAN overlays

A nested VXLAN overlay is a network virtualization mechanism that

- enables host-originated VXLAN tunnels (inner VXLAN) to be transported across existing VXLAN BGP EVPN data center fabrics (outer VXLAN)
- supports transparent integration between host overlays and base network overlays with automatic detection and handling of different VXLAN profiles, and
- provides attachment flexibility for both Layer 2 and Layer 3 interfaces, supporting untagged and tagged traffic across various port types.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
005

	https://www.cisco.com/c/en/us/td/docs/dcn/nx-os/nexus9000/105x/configuration/vxlan/cisco-nexus-9000-series-nx-os-vxlan-configuration-guide-release-105x/m_configuring_vxlan_93x.html. **VXLAN** VXLAN (Virtual Extensible LAN) is a network virtualization technology that enables the creation of virtualized Layer 2 networks over a Layer 3 infrastructure. It encapsulates Ethernet frames in UDP packets, allowing for the extension of Layer 2 segments across geographically dispersed data centers or cloud environments. It uses MAC Address-in-User Datagram Protocol (MAC-in-UDP) encapsulation to provide a means to extend Layer 2 segments across the core network. VXLAN is particularly useful in large-scale data center deployments, where it helps to overcome the limitations of traditional VLANs, such as the VLAN counts and scalability issues with Layer 2 protcols such as STP (Spanning Tree Protocol). Cisco provides a comprehensive solution for VXLAN management through the Nexus Dashboard management, which simplifies the deployment and operation of VXLAN in Cisco Nexus environments. The Nexus Dashboard offers a centralized platform for managing VXLAN configurations, monitoring network performance, and automating operational tasks. https://netascode.cisco.com/docs/guides/vxlan/overview/.
[1a] a tunnel manager configured to collect tunnel end point information within a network, and to connect each tunnel end point with a tunnel based on the tunnel end point information collected;	The Accused Instrumentalities include a tunnel manager configured to collect tunnel end point information within a network, and to connect each tunnel end point with a tunnel based on the tunnel end point information collected. For example, as evidenced below, Cisco ACI Multi-Site exchanges tunnel endpoint information — including each site's Overlay Unicast TEP (O-UTEP), Overlay Multicast TEP (O-MTEP), and the EVPN-RID — through MP-BGP EVPN

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
006

sessions between the spine nodes of the participating fabrics. The Cisco whitepaper explains that "endpoint reachability information is exchanged across sites using a Multiprotocol-BGP (MP-BGP) Ethernet VPN (EVPN) control plane," and that O-UTEP and O-MTEP are the anycast tunnel-endpoint addresses used by the fabrics. Once this tunnel-endpoint information is exchanged, "all communication…between endpoints connected to different sites is achieved by establishing site-to-site VXLAN tunnels" that use the collected O-UTEP/O-MTEP addresses as tunnel endpoints. Accordingly, the accused instrumentalities include a "tunnel manager" that collects tunnel endpoint information and creates VXLAN tunnels based on that information. *See, e.g.*,

- Intersite control plane: Endpoint reachability information is exchanged across sites using a Multiprotocol-BGP (MP-BGP) Ethernet VPN (EVPN) control plane. This approach allows the exchange of MAC and IP address information for the endpoints that communicate across sites. MP-BGP EVPN sessions are established between the spine nodes deployed in separate fabrics that are managed by the same instance of Cisco Nexus Dashboard Orchestrator, as discussed in more detail in the section "Cisco ACI Multi-Site overlay control plane."
- Intersite data plane: All communication (Layer 2 or Layer 3) between endpoints connected to different sites is achieved by establishing site-to-site Virtual Extensible LAN (VXLAN) tunnels across a generic IP network that interconnects the various sites. As discussed in the section "Intersite connectivity deployment considerations", this IP network has no specific functional requirements other than the capability to support routing and increased Maximum Transmission Unit (MTU) size (given the overhead from the VXLAN encapsulation).

- Overlay Unicast TEP (O-UTEP): This common anycast address is shared by all the spine nodes in the same pod and is used to source and receive unicast VXLAN data-plane traffic. Each pod is characterized by an O-UTEP address that, essentially, uniquely identifies the site (for single pod fabric deployments).

  Note: When deploying a Multi-Pod fabric, each pod gets assigned a unique O-UTEP address. For more information please refer to the "Integration of Cisco ACI Multi-Pod and Multi-Site" section.

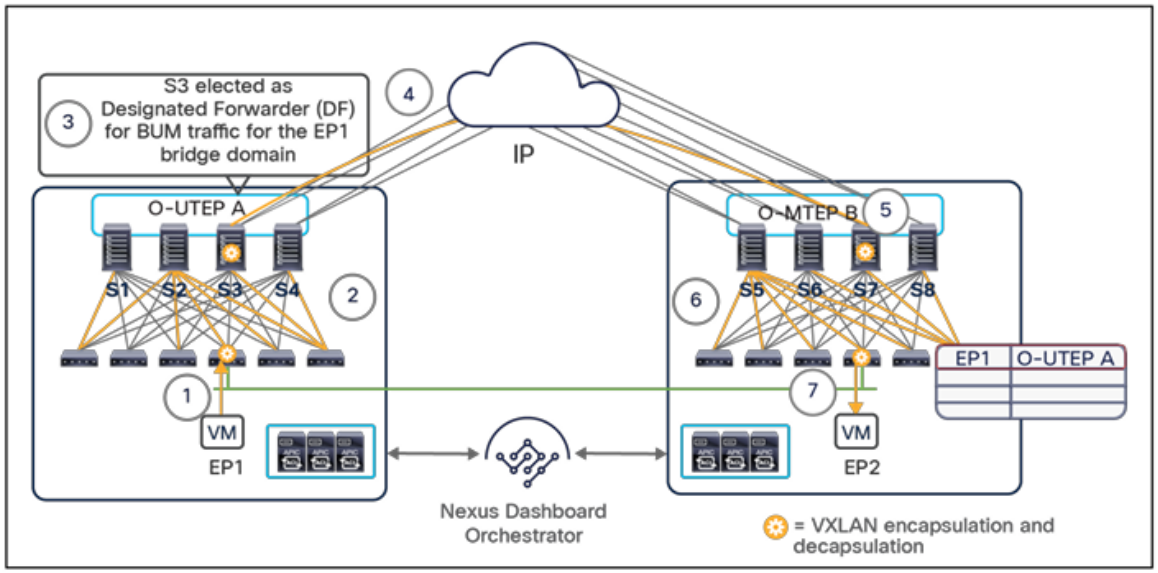
- Overlay Multicast TEP (O-MTEP): This common anycast address is shared by all the spine nodes in the same site and is used to perform head-end replication for BUM traffic. BUM traffic is sourced from the O-UTEP address defined on the local spine nodes and destined for the O-MTEP of remote sites to which the given bridge domain is being stretched.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
007

	As shown in Figure 60, the EVPN-RID, O-UTEP, and O-MTEP addresses are the only prefixes that must be exchanged across sites to enable the intersite EVPN control plane and the VXLAN data plane. Consequently, they are the only prefixes that should be learned in the ISN routing domain. This implies that those IP addresses must be globally routable across the ISN, which should normally not be a problem, because they are independent of the original TEP pools associated to each fabric and assigned separately on Cisco Nexus Dashboard Orchestrator at the time of Multi-Site deployment. https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html.
[1b] a tunnel packet end point configured to receive a tunnel packet and process the packet such that an L2 switching may be performed, and to transmit the packet to a domain VSI (virtual switching instance);	The Accused Instrumentalities include a tunnel packet end point configured to receive a tunnel packet and process the packet such that an L2 switching may be performed, and to transmit the packet to a domain VSI (virtual switching instance). For example, as evidenced below, Cisco ACI Multi-Site spine nodes act as VXLAN tunnel endpoints that receive and process VXLAN-encapsulated tunnel packets. The Cisco whitepaper explains that "one of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain," and Figure 69 depicts VXLAN decapsulation occurring at these spine nodes. Cisco further states that "the VXLAN Network Identifier (VNID) identifies the bridge domain (BD) (for Layer 2 communication)," confirming that VNID translation prepares the packet for Layer 2 switching inside the BD. After processing, the spine "sends the traffic" into the bridge domain, which serves as the L2 switching construct. Accordingly, the accused products include a tunnel packet end point that receives VXLAN tunnel packets, processes them so that L2 switching may be performed, and forwards the processed packet into the domain's Layer 2 switching instance. *See, e.g.*,

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
008

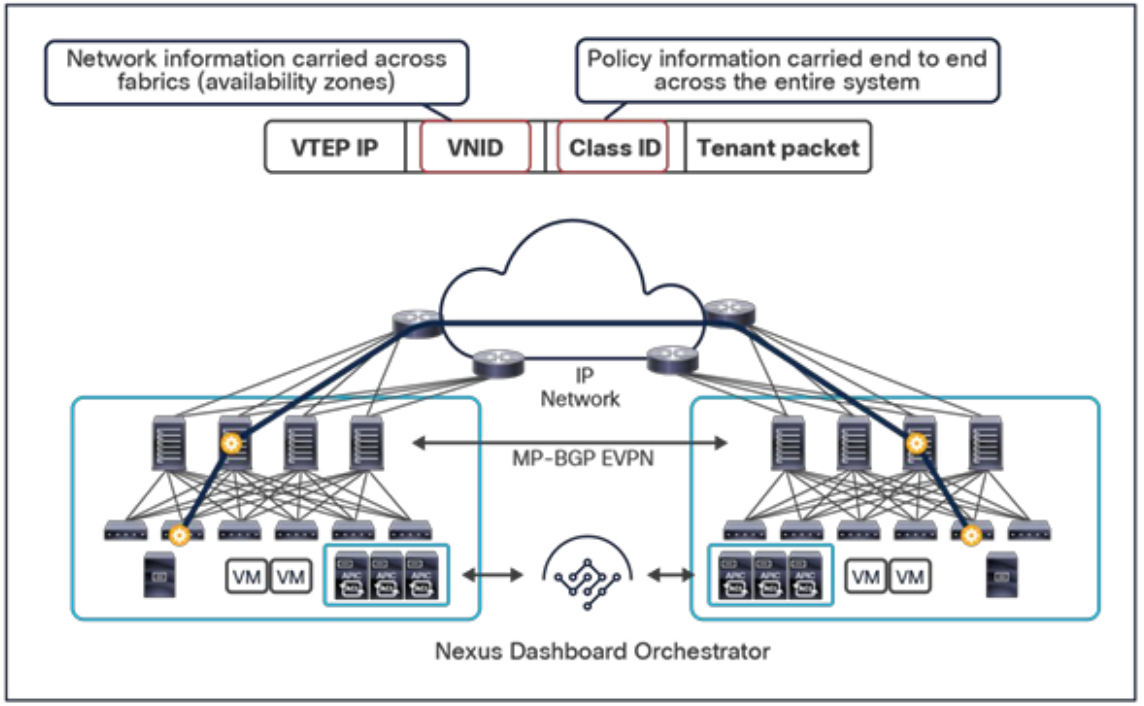
Figure 69 shows the sequence of events required to send a Layer 2 BUM frame across sites.



Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
009

	Figure 69. Layer 2 BUM traffic across sites 1. EP1, belonging to a specific bridge domain, generates a Layer 2 BUM frame. 2. Depending on the type of frame and the corresponding bridge domain settings (as previously clarified above), the leaf may need to flood the traffic in that specific Layer 2 domain. As a consequence, the frame is VXLAN-encapsulated and sent to the specific multicast group (called GIPo) associated with the bridge domain within the fabric along one of the specific multi-destination trees associated to that GIPo, so it can reach all the other leaf and spine nodes. 3. One of the spine nodes connected to the external intersite network is elected as the designated forwarder for that specific bridge domain (this election is held between the spine nodes using IS-IS protocol exchanges). The designated forwarder is responsible for replicating each BUM frame for that bridge domain to all the remote sites with the same stretched bridge domain. 4. The designated forwarder makes copies of the BUM frame and sends it to the remote sites. The destination IP address used when VXLAN encapsulating the packet is the special IP address (O-MTEP) identifying each remote site and is used specifically for the transmission of BUM traffic across sites. The O-MTEP is another anycast IP address defined on all the remote spine nodes that are connected to the intersite network (each site uses a unique O-MTEP address). The source IP address for the VXLAN-encapsulated packet is instead the anycast O-UTEP address deployed on all the local spine nodes connected to the intersite network. **Note:** The O-MTEP (referred to as "Overlay Multicast TEP" in the Cisco Nexus Dashboard Orchestrator GUI) is yet another IP address that must be sent to the Layer 3 network connecting the fabrics. 5. One of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain, and sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain (using the multicast group locally assigned to the BD as destination in the VXLAN header). https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html.
[1c] the domain VSI configured to perform the L2 switching on the processed packet, and to	The Accused Instrumentalities include the domain VSI configured to perform the L2 switching on the processed packet, and to transmit the packet to a tunnel packet generator. For example, As evidenced below, Cisco states that "the VXLAN Network Identifier (VNID)

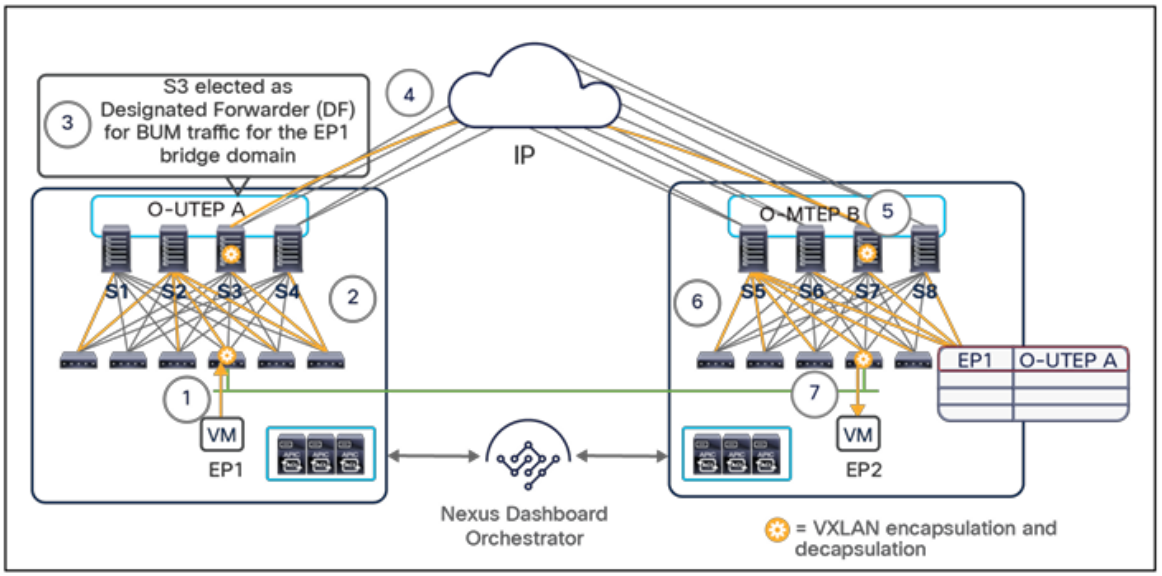
Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
010

transmit the packet to a tunnel packet generator; and	identifies the bridge domain (BD) (for Layer 2 communication),” demonstrating that the BD is the Layer 2 switching construct in ACI. Figure 69 further shows that, after VNID translation, incoming VXLAN packets are forwarded along “the local multi-destination trees defined for the bridge domain,” which reflects BD-controlled Layer 2 switching behavior. The associated Cisco text explains that these forwarding trees are determined within the BD. In addition, the Cisco whitepaper describes that intersite communication is performed over VXLAN tunnels, meaning that once the BD determines the destination is in a remote site, the packet is handed off to the VXLAN encapsulation logic on the leaf or spine. Accordingly, the BD performs the Layer 2 switching on processed packets and then transmits those packets to the tunnel packet generator for VXLAN encapsulation, satisfying Claim 1(c). *See, e.g.*,  **Figure 6.** Carrying network and policy information across sites

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
011

The VXLAN Network Identifier (VNID) identifies the bridge domain (BD) (for Layer 2 communication) or the Virtual Routing and Forwarding (VRF) instance (for Layer 3 traffic) of the endpoint sourcing the traffic (for intra-VRF communication). The class ID is the unique identifier of the source Endpoint Group (EPG). However, these values are locally significant within a fabric. Because a completely separate and independent APIC domain and fabric are deployed at the destination site, a translation function (also referred to as "name-space normalization") must be applied before the traffic is forwarded inside the receiving site, to help ensure that locally significant values identifying that same source EPG, bridge domain, and VRF instance are used.

Figure 69 shows the sequence of events required to send a Layer 2 BUM frame across sites.



Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
012

	Figure 69. Layer 2 BUM traffic across sites 1. EP1, belonging to a specific bridge domain, generates a Layer 2 BUM frame. 2. Depending on the type of frame and the corresponding bridge domain settings (as previously clarified above), the leaf may need to flood the traffic in that specific Layer 2 domain. As a consequence, the frame is VXLAN-encapsulated and sent to the specific multicast group (called GIPo) associated with the bridge domain within the fabric along one of the specific multi-destination trees associated to that GIPo, so it can reach all the other leaf and spine nodes. 3. One of the spine nodes connected to the external intersite network is elected as the designated forwarder for that specific bridge domain (this election is held between the spine nodes using IS-IS protocol exchanges). The designated forwarder is responsible for replicating each BUM frame for that bridge domain to all the remote sites with the same stretched bridge domain. 4. The designated forwarder makes copies of the BUM frame and sends it to the remote sites. The destination IP address used when VXLAN encapsulating the packet is the special IP address (O-MTEP) identifying each remote site and is used specifically for the transmission of BUM traffic across sites. The O-MTEP is another anycast IP address defined on all the remote spine nodes that are connected to the intersite network (each site uses a unique O-MTEP address). The source IP address for the VXLAN-encapsulated packet is instead the anycast O-UTEP address deployed on all the local spine nodes connected to the intersite network. **Note:** The O-MTEP (referred to as "Overlay Multicast TEP" in the Cisco Nexus Dashboard Orchestrator GUI) is yet another IP address that must be sent to the Layer 3 network connecting the fabrics. 5. One of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain, and sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain (using the multicast group locally assigned to the BD as destination in the VXLAN header). https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html.
[1d] the tunnel packet generator configured to add a tunnel header to the packet for which the L2 switching has been performed to generate a	The Accused Instrumentalities include the tunnel packet generator configured to add a tunnel header to the packet for which the L2 switching has been performed to generate a new tunnel packet, and to transmit the new tunnel packet. For example, as evidenced below, Cisco

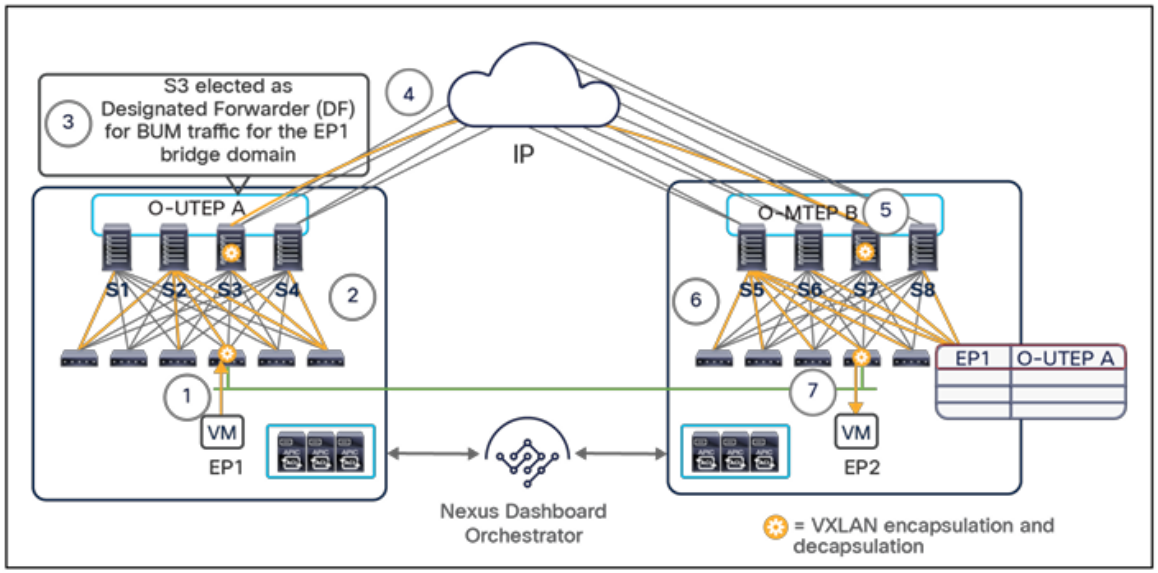
Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
013

new tunnel packet, and to transmit the new tunnel packet.	explains that intersite communication in ACI Multi-Site always occurs over VXLAN tunnels, stating that "all communication (Layer 2 or Layer 3) between endpoints connected to different sites is achieved by establishing site-to-site Virtual Extensible LAN (VXLAN) tunnels." Cisco further describes that the spines "rewrite the source IP address of the VXLAN-encapsulated packet" to use the local O-UTEP address as the tunnel source and use the remote site's O-UTEP or O-MTEP address as the tunnel destination. Figure 69 likewise depicts the designated forwarder spine "VXLAN encapsulating" outgoing traffic and sending the resulting VXLAN packet to the remote site. Accordingly, after L2 switching has been performed in the bridge domain, the accused products add the required VXLAN tunnel header to generate a new tunnel packet and transmit that packet across the intersite network, satisfying Claim 1(d). *See, e.g.*, • Intersite control plane: Endpoint reachability information is exchanged across sites using a Multiprotocol-BGP (MP-BGP) Ethernet VPN (EVPN) control plane. This approach allows the exchange of MAC and IP address information for the endpoints that communicate across sites. MP-BGP EVPN sessions are established between the spine nodes deployed in separate fabrics that are managed by the same instance of Cisco Nexus Dashboard Orchestrator, as discussed in more detail in the section "Cisco ACI Multi-Site overlay control plane." • Intersite data plane: All communication (Layer 2 or Layer 3) between endpoints connected to different sites is achieved by establishing site-to-site Virtual Extensible LAN (VXLAN) tunnels across a generic IP network that interconnects the various sites. As discussed in the section "Intersite connectivity deployment considerations", this IP network has no specific functional requirements other than the capability to support routing and increased Maximum Transmission Unit (MTU) size (given the overhead from the VXLAN encapsulation).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
014

The definition of the O-UTEP and O-MTEP ensures that all the resources connected to a specific site are always seen as reachable from the remote sites through one of those two IP addresses (depending on if it is unicast traffic or BUM/multicast communication). This implies that when observing east-west VXLAN traffic exchanged between two sites in the ISN, all the packets will always be sourced from the O-UTEP address identifying the source site and destined to the O-UTEP (or O-MTEP) of the destination site. Notice that this does not prevent taking advantage of multiple ECMP links that may interconnect the two sites: when building the VXLAN encapsulated packet, a hashing of the L2/L3/L4 headers for the original frame generated by the endpoint is used as a UDP source port in the external header, and this allows you to build "entropy" in the packet. Different application flows between the same sites (and even between the same pair of endpoints) would cause the creation of different UDP source port values: as long as the network devices inside the ISN consider the L4 port information to choose the link to forward traffic, it will hence be possible to load-balance VXLAN packets across multiple paths.

Figure 69 shows the sequence of events required to send a Layer 2 BUM frame across sites.



Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
015

	Figure 69. Layer 2 BUM traffic across sites 1. EP1, belonging to a specific bridge domain, generates a Layer 2 BUM frame. 2. Depending on the type of frame and the corresponding bridge domain settings (as previously clarified above), the leaf may need to flood the traffic in that specific Layer 2 domain. As a consequence, the frame is VXLAN-encapsulated and sent to the specific multicast group (called GIPo) associated with the bridge domain within the fabric along one of the specific multi-destination trees associated to that GIPo, so it can reach all the other leaf and spine nodes. 3. One of the spine nodes connected to the external intersite network is elected as the designated forwarder for that specific bridge domain (this election is held between the spine nodes using IS-IS protocol exchanges). The designated forwarder is responsible for replicating each BUM frame for that bridge domain to all the remote sites with the same stretched bridge domain. 4. The designated forwarder makes copies of the BUM frame and sends it to the remote sites. The destination IP address used when VXLAN encapsulating the packet is the special IP address (O-MTEP) identifying each remote site and is used specifically for the transmission of BUM traffic across sites. The O-MTEP is another anycast IP address defined on all the remote spine nodes that are connected to the intersite network (each site uses a unique O-MTEP address). The source IP address for the VXLAN-encapsulated packet is instead the anycast O-UTEP address deployed on all the local spine nodes connected to the intersite network. **Note:** The O-MTEP (referred to as "Overlay Multicast TEP" in the Cisco Nexus Dashboard Orchestrator GUI) is yet another IP address that must be sent to the Layer 3 network connecting the fabrics. 5. One of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain, and sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain (using the multicast group locally assigned to the BD as destination in the VXLAN header). https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html.
[2] The apparatus according to claim 1, wherein the domain VSI, when perform the L2 switching, obtains L2 SA (source address) information for	The Accused Instrumentalities include the apparatus according to claim 1, wherein the domain VSI, when perform the L2 switching, obtains L2 SA (source address) information for the processed packet, determines whether or not L2 DA (destination address) information of the

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
016

the processed packet, determines whether or not L2 DA (destination address) information of the processed packet is a unicast address already learned, and performs a flooding based on a type of an input virtual port, when the L2 DA information is not already learned unicast address, or is multicast or broadcast address; wherein the flooding is performed on only domain virtual ports within the domain VSI, when the type of the input virtual port being an edge virtual port, and the flooding is performed on only edge virtual ports within the domain VSI when the type of the input virtual port being a domain virtual port.	processed packet is a unicast address already learned, and performs a flooding based on a type of an input virtual port, when the L2 DA information is not already learned unicast address, or is multicast or broadcast address; wherein the flooding is performed on only domain virtual ports within the domain VSI, when the type of the input virtual port being an edge virtual port, and the flooding is performed on only edge virtual ports within the domain VSI when the type of the input virtual port being a domain virtual port. For example, as evidenced bekow, Cisco's ACI Multi-Site documentation describes the Layer 2 forwarding behavior performed by the Bridge Domain (BD), which is the Layer 2 switching construct in ACI. Figures 54, 55, 69, and 70 collectively show that the BD learns the source L2 information when frames arrive on leaf-facing interfaces, determines whether the destination MAC is already known in the BD, and forwards known-unicast frames accordingly. When the destination is not learned, or when the frame is multicast or broadcast, Cisco explains that the BD floods the traffic only along the BD's multi-destination trees (GIPO groups), which correspond to the subset of ports designated for that BD. These figures further show that the flooding behavior depends on the type of ingress port. Leaf-originated traffic (edge-facing) is flooded only toward the spine nodes participating in the intersite BD (domain-facing), while spine-originated intersite traffic is forwarded only to the local leaf nodes within the BD and is not propagated back across intersite interfaces. Accordingly, the accused products include a domain VSI that (1) obtains L2 source address information, (2) determines whether the L2 destination is a known unicast or requires flooding, and (3) performs directional flooding based on the ingress port type, satisfying Claim 2. *See, e.g.*,

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
017

Layer 2 connectivity across sites with flooding

In the next use case, the Cisco ACI Multi-Site design provides traditional Layer 2 stretching of bridge domains across sites, including the capability to flood Layer 2 BUM frames (Figure 54).

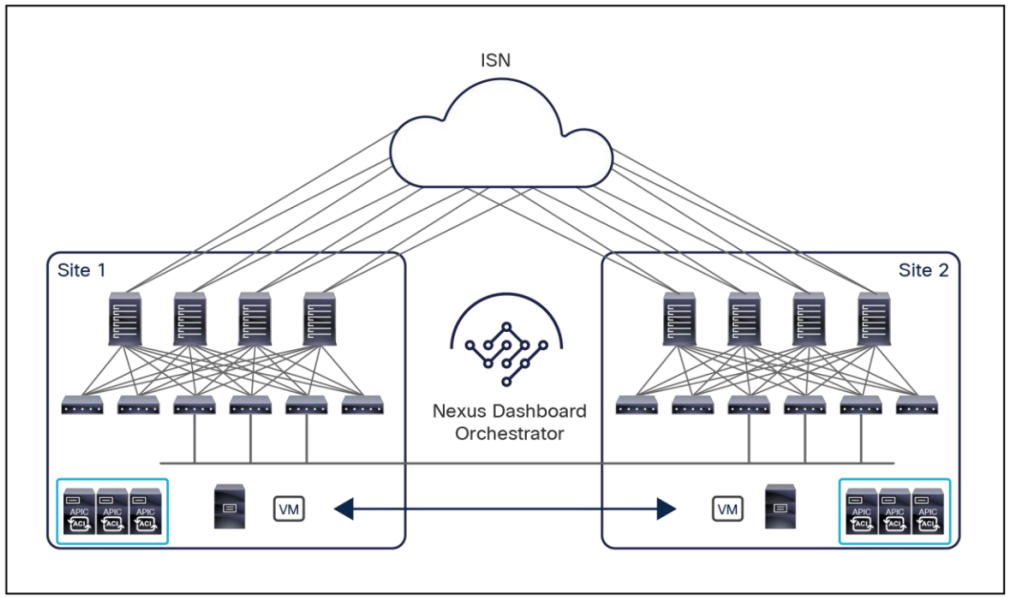


Figure 54.
Layer 2 connectivity across sites with flooding

The need to flood BUM traffic is driven by specific requirements, such as application clustering (which traditionally calls for the use of Layer 2 multicast communication between different application cluster nodes).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
018

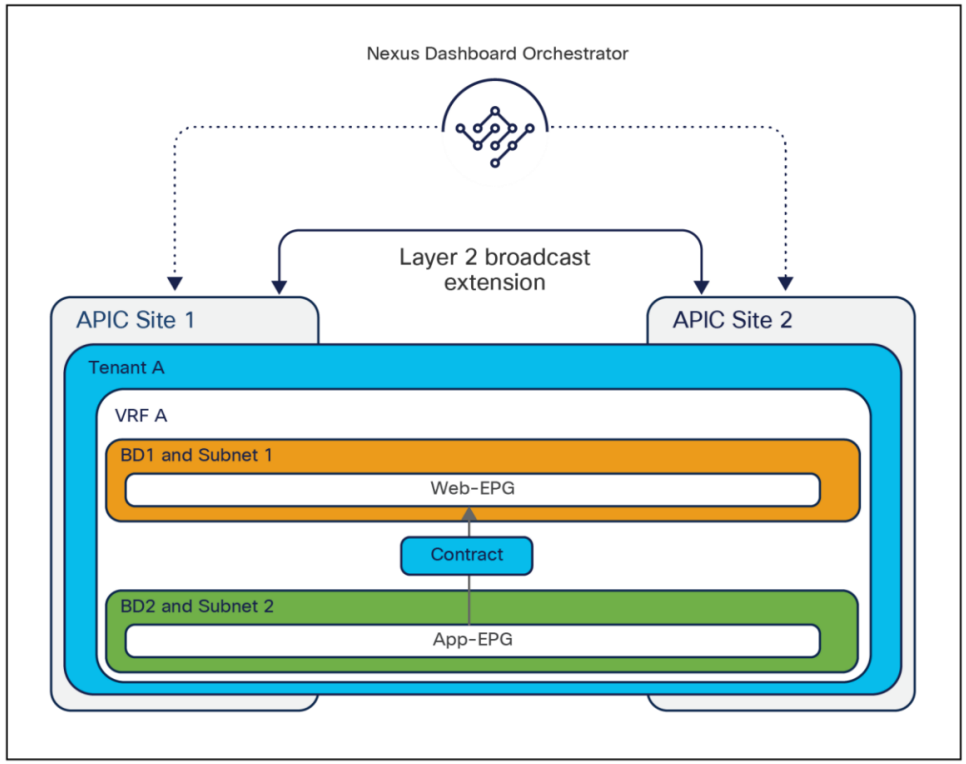


Figure 55.
Layer 2 connectivity across sites with flooding (logical view)

Although Cisco ACI Multi-Pod usually would be positioned as the architectural choice for deployments requiring full Layer 2 flooding between Cisco ACI islands, one advantage offered by the Multi-Site design with this configuration is the capability to tightly control which bridge domains are extended (with or without enabling BUM forwarding) and which are kept local. Cisco Nexus Dashboard Orchestrator, in fact, allows you to differentiate the flooding behavior on a per-bridge-domain level, which is useful in a real-life scenario in which the flooding behavior must be supported only in a subset of the stretched bridge domains.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
019

Note that multicast support is not required within the ISN routing domain to allow the exchange of Layer 2 multi-destination (BUM) traffic across pods in the specific use case in which a bridge domain is stretched with flooding enabled. This is because the Cisco ACI Multi-Site design uses the ingress replication function on the spine nodes of the source site to replicate BUM traffic to all the remote sites on which that bridge domain is stretched. For more information about this capability, refer to the section "Layer 2 BUM traffic handling across sites."

Cisco ACI Multi-Site overlay data plane

After endpoint information is exchanged across sites, the VXLAN data plane is used to allow intersite Layer 2 and Layer 3 communication. Before exploring in detail how this communication can be established, you should understand how Layer 2 multi-destination traffic (usually referred to as BUM) is handled across sites.

Layer 2 BUM traffic handling across sites

The deployment of VXLAN allows the use of a logical abstraction so that endpoints separated by multiple Layer 3 hops can communicate as if they were part of the same logical Layer 2 domain. Thus, those endpoints must be capable of sourcing Layer 2 multi-destination frames so that they can be received by all the other endpoints connected to the same Layer 2 segment, regardless of their actual physical location.

This capability can be achieved in one of two ways: by using the native multicast replication functions offered by the Layer 3 infrastructure interconnecting the endpoints or by enabling ingress replication functions on the source VXLAN TEP (VTEP) devices, which create multiple unicast copies of each BUM frame to be sent to all the remote VTEPs on which those endpoints part of the same Layer 2 domain are connected.

The Cisco ACI Multi-Site design adopts this second approach for intersite BUM forwarding, with the Multi-Site-capable spine switches performing the ingress replication function, because the interconnected fabrics may be deployed around the world, and it would be difficult to ensure proper multicast support across the entire interconnecting network infrastructure (the approach adopted instead in the Cisco ACI Multi-Pod architecture).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
020

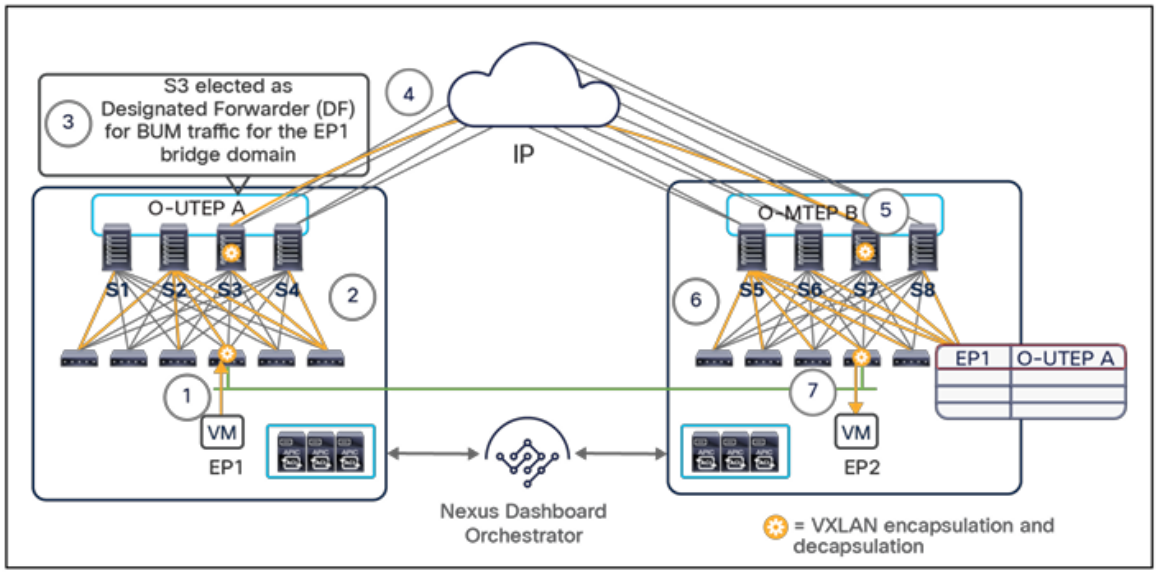
The transmission of Layer 2 BUM frames across sites is required only for the specific bridge domains that are stretched with flooding enabled (that is, the "Intersite BUM Traffic Allow" flag is configured for the bridge domains). There are three different types of Layer 2 BUM traffic, and below is described the intersite forwarding behavior for each of them when BUM is allowed across sites for a given bridge domain. The assumption here is that specific bridge domain configuration knobs are not modified at the APIC level, but the configuration for the bridge domain is only controlled at the Cisco Nexus Dashboard Orchestrator level.

- Layer 2 Broadcast frames (B): Those are always forwarded across sites. A special type of Layer 2 broadcast traffic is ARP; specific considerations can be found as part of the "Intra-subnet unicast communication across sites" section.
- Layer 2 Unknown Unicast frames (U): Those frames, by default, are not flooded across sites but are instead forwarded in unicast mode, assuming that the destination MAC is known in the COOP database of the local spines (else the traffic will be dropped by the receiving spine). However, there is the possibility of changing this behavior on the bridge-domain-specific configuration of Cisco Nexus Dashboard Orchestrator by selecting the "flood" option associated to the "L2 UNKNOWN UNICAST" traffic.
- Layer 2 Multicast frames (M): The same forwarding behavior applies to intra-bridge-domain Layer 3 multicast frames (that is, the source and receivers are in the same or different IP subnets but part of the same bridge domain) or to "true" Layer 2 multicast frames (that is, the destination MAC address is multicast and there is no IP header in the packet). In both cases, the traffic is forwarded across the sites where the bridge domain is stretched once BUM forwarding is enabled for that bridge domain.

  Note: What is mentioned above applies assuming that multicast routing is not enabled. For more considerations about support of multicast routing with Cisco ACI Multi-Site, please refer to the "Multi-Site Layer 3 multicast" section.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
021

Figure 69 shows the sequence of events required to send a Layer 2 BUM frame across sites.



Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
022

Figure 69.
Layer 2 BUM traffic across sites

1. EP1, belonging to a specific bridge domain, generates a Layer 2 BUM frame.
2. Depending on the type of frame and the corresponding bridge domain settings (as previously clarified above), the leaf may need to flood the traffic in that specific Layer 2 domain. As a consequence, the frame is VXLAN-encapsulated and sent to the specific multicast group (called GIPo) associated with the bridge domain within the fabric along one of the specific multi-destination trees associated to that GIPo, so it can reach all the other leaf and spine nodes.
3. One of the spine nodes connected to the external intersite network is elected as the designated forwarder for that specific bridge domain (this election is held between the spine nodes using IS-IS protocol exchanges). The designated forwarder is responsible for replicating each BUM frame for that bridge domain to all the remote sites with the same stretched bridge domain.
4. The designated forwarder makes copies of the BUM frame and sends it to the remote sites. The destination IP address used when VXLAN encapsulating the packet is the special IP address (O-MTEP) identifying each remote site and is used specifically for the transmission of BUM traffic across sites. The O-MTEP is another anycast IP address defined on all the remote spine nodes that are connected to the intersite network (each site uses a unique O-MTEP address). The source IP address for the VXLAN-encapsulated packet is instead the anycast O-UTEP address deployed on all the local spine nodes connected to the intersite network.

   Note: The O-MTEP (referred to as "Overlay Multicast TEP" in the Cisco Nexus Dashboard Orchestrator GUI) is yet another IP address that must be sent to the Layer 3 network connecting the fabrics.
5. One of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain, and sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain (using the multicast group locally assigned to the BD as destination in the VXLAN header).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
023

6. The traffic is forwarded within the site and reaches all the spine and leaf nodes with endpoints actively connected to the specific bridge domain.

7. The receiving leaf nodes use the information contained in the VXLAN header to learn the site location for endpoint EP1 that sourced the BUM frame. They also send the BUM frame to all (or some of) the local interfaces associated with the bridge domain, so that endpoint EP2 (in this example) can receive it.

As previously mentioned, any defined bridge domain is associated with a multicast group address (or a set of multicast addresses), usually referred to as the GIPo address. Depending on the number of configured bridge domains, the same GIPo address may be associated with different bridge domains. Thus, when flooding for one of those bridge domains is enabled across sites, BUM traffic for the other bridge domains using the same GIPo address is also sent across the sites and will then be dropped on the received spine nodes. This behavior can increase the bandwidth utilization in the intersite network.

Because of this behavior, when a bridge domain is configured as stretched with BUM flooding enabled from the Cisco Nexus Dashboard Orchestrator GUI, by default a GIPo address is assigned from a separate range of multicast addresses. This is reflected in the GUI by the "OPTIMIZE WAN BANDWIDTH" flag, which is enabled by default for BDs that are created directly on NDO, as shown in Figure 70.

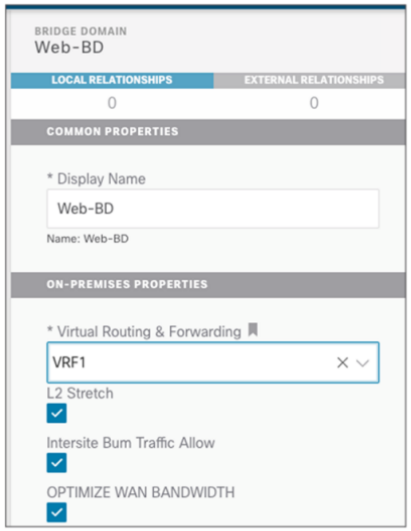


Figure 70.
Optimizing BUM flooding across sites

https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
024

	centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html. *See also*, *e.g.*, evidence and analysis described above with respect to claim 1.
[3] The apparatus according to claim 2, wherein, the domain VSI, when performing the L2 switching, in response to the L2 DA information being a unicast address already learned, transmits the processed packet to a destination virtual port corresponding to the unicast address, and obtains an L3 address of a destination tunnel end point of the packet from the destination virtual port.	The Accused Instrumentalities include the apparatus according to claim 2, wherein, the domain VSI, when performing the L2 switching, in response to the L2 DA information being a unicast address already learned, transmits the processed packet to a destination virtual port corresponding to the unicast address, and obtains an L3 address of a destination tunnel end point of the packet from the destination virtual port. For example, as evidenced below, Cisco ACI Multi-Site performs learned-unicast forwarding whenever the destination MAC address is already known in the system's forwarding state. In the section handling Layer 2 Unknown Unicast frames, the Cisco whitepaper explains that such frames are "not flooded across sites but are instead forwarded in unicast mode, assuming that the destination MAC is known in the COOP database of the local spines." This demonstrates that, when the L2 destination address is a learned unicast, the Bridge Domain's forwarding logic performs a unicast lookup and forwards the packet along a specific output path, which corresponds to transmitting the processed packet to the "destination virtual port" for that unicast address. Additional sections of the same Cisco ACI Multi-Site whitepaper confirm that the forwarding tables and COOP entries used for intersite unicast flows include the L3 tunnel endpoint information. Cisco states that, after endpoint reachability and ARP exchanges, "the leaf and spine nodes would have their Layer 2, Layer 3, and COOP tables properly populated to allow the establishment of intersite unicast traffic flows," and further explains that "the local spine nodes know the remote O-UTEP address identifying the site to which EP2 is connected and can encapsulate the packet and send it across the ISN toward the remote site." Because O-UTEP is the L3 address of the unicast VXLAN tunnel endpoint, these statements show that, following a learned-unicast lookup, the system obtains the L3 address of the destination tunnel endpoint from the forwarding/COOP entry associated with the destination MAC address. Accordingly, when the L2 DA is a unicast address already learned, the accused products perform a unicast lookup within the Bridge Domain (the domain VSI),

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
025

forward the packet to the corresponding virtual port, and obtain the destination tunnel endpoint's L3 address (O-UTEP) for use in VXLAN encapsulation, satisfying Claim 3. *See, e.g.*,

Layer 2 connectivity across sites with flooding

In the next use case, the Cisco ACI Multi-Site design provides traditional Layer 2 stretching of bridge domains across sites, including the capability to flood Layer 2 BUM frames (Figure 54).

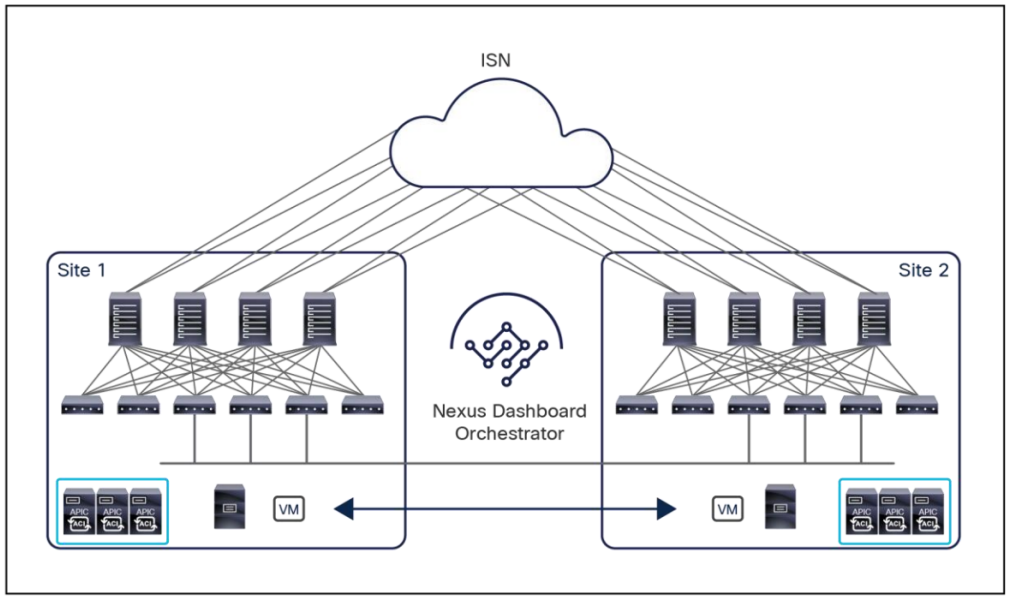


Figure 54.
Layer 2 connectivity across sites with flooding

The need to flood BUM traffic is driven by specific requirements, such as application clustering (which traditionally calls for the use of Layer 2 multicast communication between different application cluster nodes).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
026

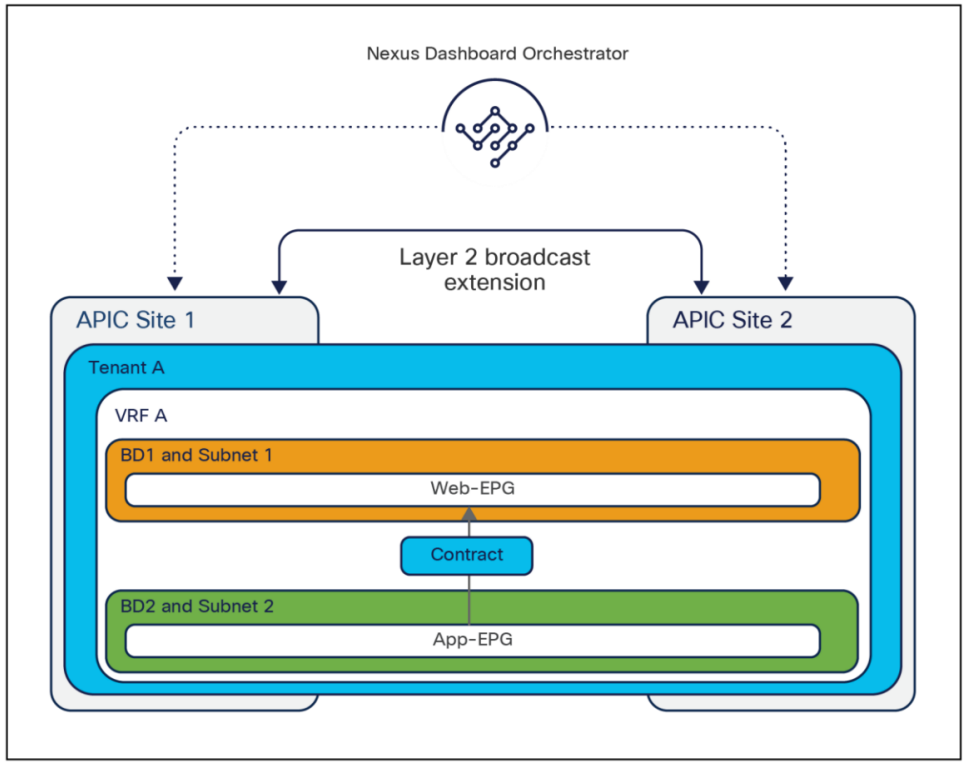


Figure 55.
Layer 2 connectivity across sites with flooding (logical view)

Although Cisco ACI Multi-Pod usually would be positioned as the architectural choice for deployments requiring full Layer 2 flooding between Cisco ACI islands, one advantage offered by the Multi-Site design with this configuration is the capability to tightly control which bridge domains are extended (with or without enabling BUM forwarding) and which are kept local. Cisco Nexus Dashboard Orchestrator, in fact, allows you to differentiate the flooding behavior on a per-bridge-domain level, which is useful in a real-life scenario in which the flooding behavior must be supported only in a subset of the stretched bridge domains.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
027

Note that multicast support is not required within the ISN routing domain to allow the exchange of Layer 2 multi-destination (BUM) traffic across pods in the specific use case in which a bridge domain is stretched with flooding enabled. This is because the Cisco ACI Multi-Site design uses the ingress replication function on the spine nodes of the source site to replicate BUM traffic to all the remote sites on which that bridge domain is stretched. For more information about this capability, refer to the section "Layer 2 BUM traffic handling across sites."

Cisco ACI Multi-Site overlay data plane

After endpoint information is exchanged across sites, the VXLAN data plane is used to allow intersite Layer 2 and Layer 3 communication. Before exploring in detail how this communication can be established, you should understand how Layer 2 multi-destination traffic (usually referred to as BUM) is handled across sites.

Layer 2 BUM traffic handling across sites

The deployment of VXLAN allows the use of a logical abstraction so that endpoints separated by multiple Layer 3 hops can communicate as if they were part of the same logical Layer 2 domain. Thus, those endpoints must be capable of sourcing Layer 2 multi-destination frames so that they can be received by all the other endpoints connected to the same Layer 2 segment, regardless of their actual physical location.

This capability can be achieved in one of two ways: by using the native multicast replication functions offered by the Layer 3 infrastructure interconnecting the endpoints or by enabling ingress replication functions on the source VXLAN TEP (VTEP) devices, which create multiple unicast copies of each BUM frame to be sent to all the remote VTEPs on which those endpoints part of the same Layer 2 domain are connected.

The Cisco ACI Multi-Site design adopts this second approach for intersite BUM forwarding, with the Multi-Site-capable spine switches performing the ingress replication function, because the interconnected fabrics may be deployed around the world, and it would be difficult to ensure proper multicast support across the entire interconnecting network infrastructure (the approach adopted instead in the Cisco ACI Multi-Pod architecture).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
028

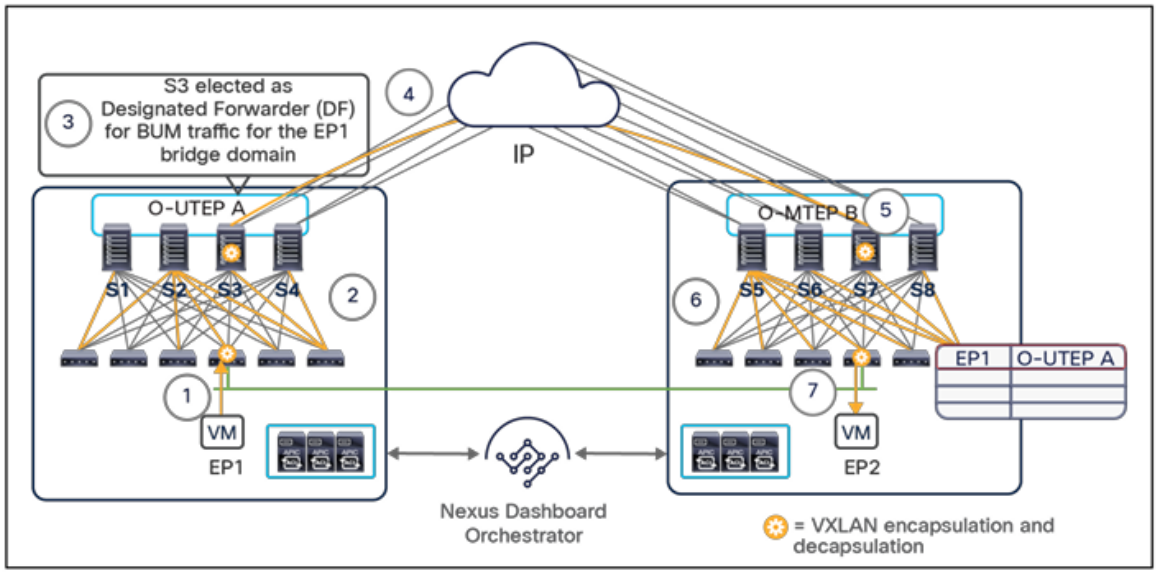
The transmission of Layer 2 BUM frames across sites is required only for the specific bridge domains that are stretched with flooding enabled (that is, the "Intersite BUM Traffic Allow" flag is configured for the bridge domains). There are three different types of Layer 2 BUM traffic, and below is described the intersite forwarding behavior for each of them when BUM is allowed across sites for a given bridge domain. The assumption here is that specific bridge domain configuration knobs are not modified at the APIC level, but the configuration for the bridge domain is only controlled at the Cisco Nexus Dashboard Orchestrator level.

- Layer 2 Broadcast frames (B): Those are always forwarded across sites. A special type of Layer 2 broadcast traffic is ARP; specific considerations can be found as part of the "Intra-subnet unicast communication across sites" section.
- Layer 2 Unknown Unicast frames (U): Those frames, by default, are not flooded across sites but are instead forwarded in unicast mode, assuming that the destination MAC is known in the COOP database of the local spines (else the traffic will be dropped by the receiving spine). However, there is the possibility of changing this behavior on the bridge-domain-specific configuration of Cisco Nexus Dashboard Orchestrator by selecting the "flood" option associated to the "L2 UNKNOWN UNICAST" traffic.
- Layer 2 Multicast frames (M): The same forwarding behavior applies to intra-bridge-domain Layer 3 multicast frames (that is, the source and receivers are in the same or different IP subnets but part of the same bridge domain) or to "true" Layer 2 multicast frames (that is, the destination MAC address is multicast and there is no IP header in the packet). In both cases, the traffic is forwarded across the sites where the bridge domain is stretched once BUM forwarding is enabled for that bridge domain.

  Note: What is mentioned above applies assuming that multicast routing is not enabled. For more considerations about support of multicast routing with Cisco ACI Multi-Site, please refer to the "Multi-Site Layer 3 multicast" section.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
029

Figure 69 shows the sequence of events required to send a Layer 2 BUM frame across sites.



Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
030

Figure 69.
Layer 2 BUM traffic across sites

1. EP1, belonging to a specific bridge domain, generates a Layer 2 BUM frame.
2. Depending on the type of frame and the corresponding bridge domain settings (as previously clarified above), the leaf may need to flood the traffic in that specific Layer 2 domain. As a consequence, the frame is VXLAN-encapsulated and sent to the specific multicast group (called GIPo) associated with the bridge domain within the fabric along one of the specific multi-destination trees associated to that GIPo, so it can reach all the other leaf and spine nodes.
3. One of the spine nodes connected to the external intersite network is elected as the designated forwarder for that specific bridge domain (this election is held between the spine nodes using IS-IS protocol exchanges). The designated forwarder is responsible for replicating each BUM frame for that bridge domain to all the remote sites with the same stretched bridge domain.
4. The designated forwarder makes copies of the BUM frame and sends it to the remote sites. The destination IP address used when VXLAN encapsulating the packet is the special IP address (O-MTEP) identifying each remote site and is used specifically for the transmission of BUM traffic across sites. The O-MTEP is another anycast IP address defined on all the remote spine nodes that are connected to the intersite network (each site uses a unique O-MTEP address). The source IP address for the VXLAN-encapsulated packet is instead the anycast O-UTEP address deployed on all the local spine nodes connected to the intersite network.

   Note: The O-MTEP (referred to as "Overlay Multicast TEP" in the Cisco Nexus Dashboard Orchestrator GUI) is yet another IP address that must be sent to the Layer 3 network connecting the fabrics.
5. One of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain, and sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain (using the multicast group locally assigned to the BD as destination in the VXLAN header).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
031

6. The traffic is forwarded within the site and reaches all the spine and leaf nodes with endpoints actively connected to the specific bridge domain.
7. The receiving leaf nodes use the information contained in the VXLAN header to learn the site location for endpoint EP1 that sourced the BUM frame. They also send the BUM frame to all (or some of) the local interfaces associated with the bridge domain, so that endpoint EP2 (in this example) can receive it.

As previously mentioned, any defined bridge domain is associated with a multicast group address (or a set of multicast addresses), usually referred to as the GIPo address. Depending on the number of configured bridge domains, the same GIPo address may be associated with different bridge domains. Thus, when flooding for one of those bridge domains is enabled across sites, BUM traffic for the other bridge domains using the same GIPo address is also sent across the sites and will then be dropped on the received spine nodes. This behavior can increase the bandwidth utilization in the intersite network.

Because of this behavior, when a bridge domain is configured as stretched with BUM flooding enabled from the Cisco Nexus Dashboard Orchestrator GUI, by default a GIPo address is assigned from a separate range of multicast addresses. This is reflected in the GUI by the "OPTIMIZE WAN BANDWIDTH" flag, which is enabled by default for BDs that are created directly on NDO, as shown in Figure 70.

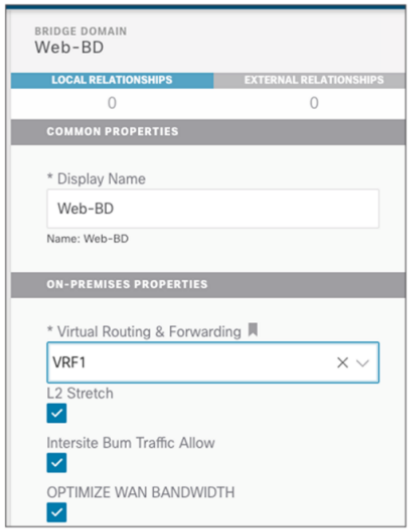


Figure 70.
Optimizing BUM flooding across sites

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
032

- Overlay Unicast TEP (O-UTEP): This common anycast address is shared by all the spine nodes in the same pod and is used to source and receive unicast VXLAN data-plane traffic. Each pod is characterized by an O-UTEP address that, essentially, uniquely identifies the site (for single pod fabric deployments).

  Note: When deploying a Multi-Pod fabric, each pod gets assigned a unique O-UTEP address. For more information please refer to the "Integration of Cisco ACI Multi-Pod and Multi-Site" section.

Layer 2 and Layer 3 unicast communication

Assuming endpoint reachability information has been exchanged across sites, and assuming ARP exchange has also successfully completed, the leaf and spine nodes would have their Layer 2, Layer 3, and COOP tables properly populated to allow the establishment of intersite unicast traffic flows (Figure 94).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
033

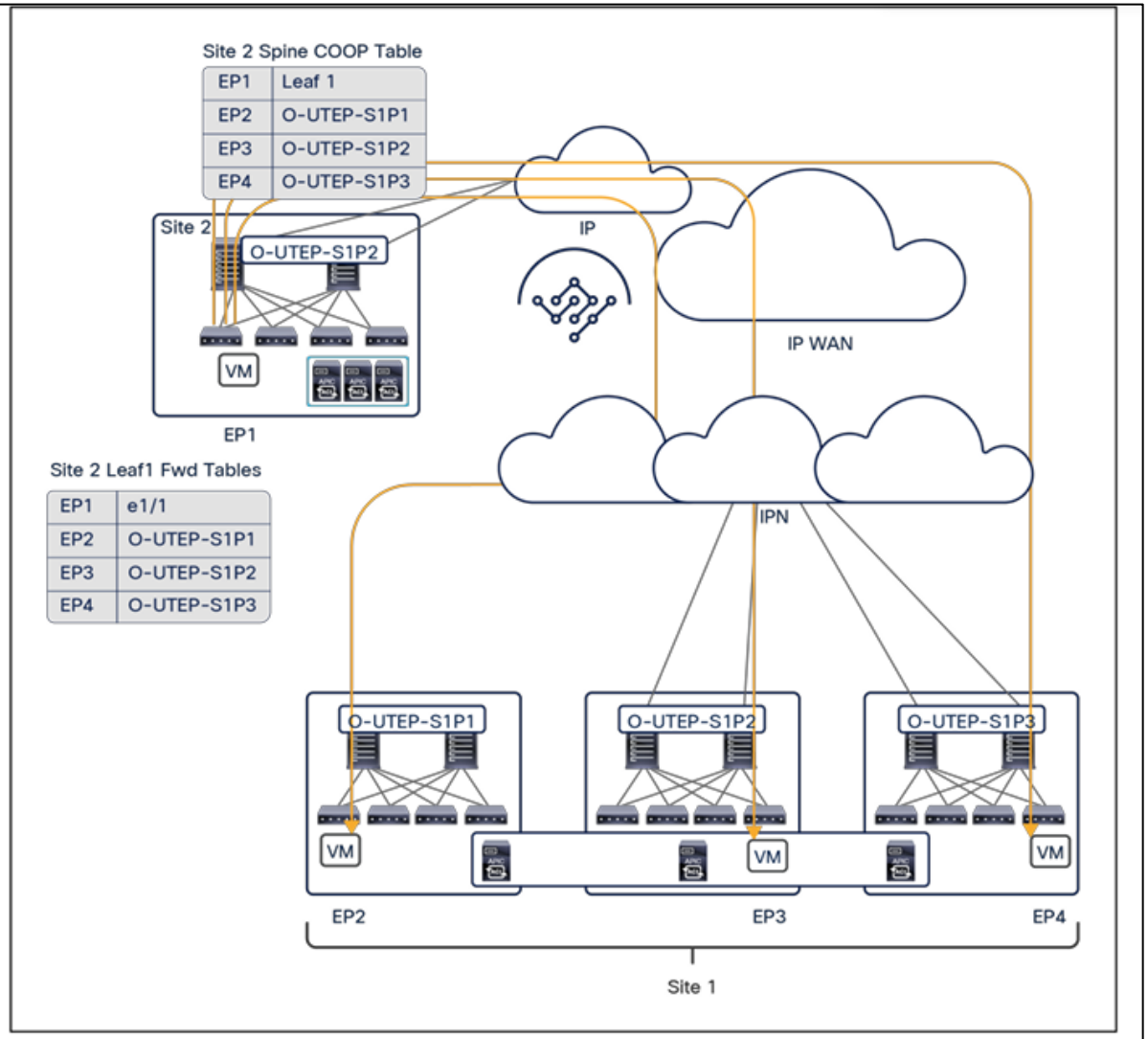


Figure 94.
Unicast traffic ingressing a Multi-Pod fabric that is part of Cisco ACI Multi-Site

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
034

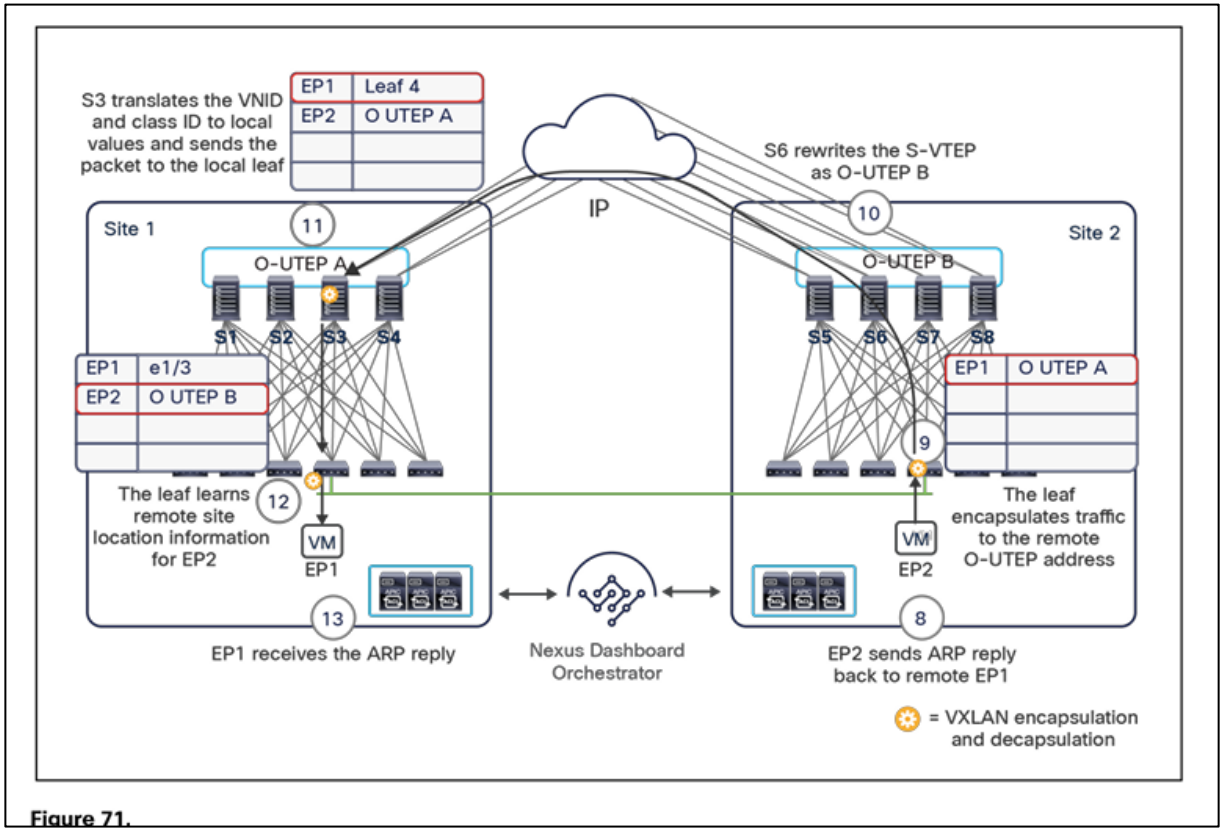


Figure 71.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
035

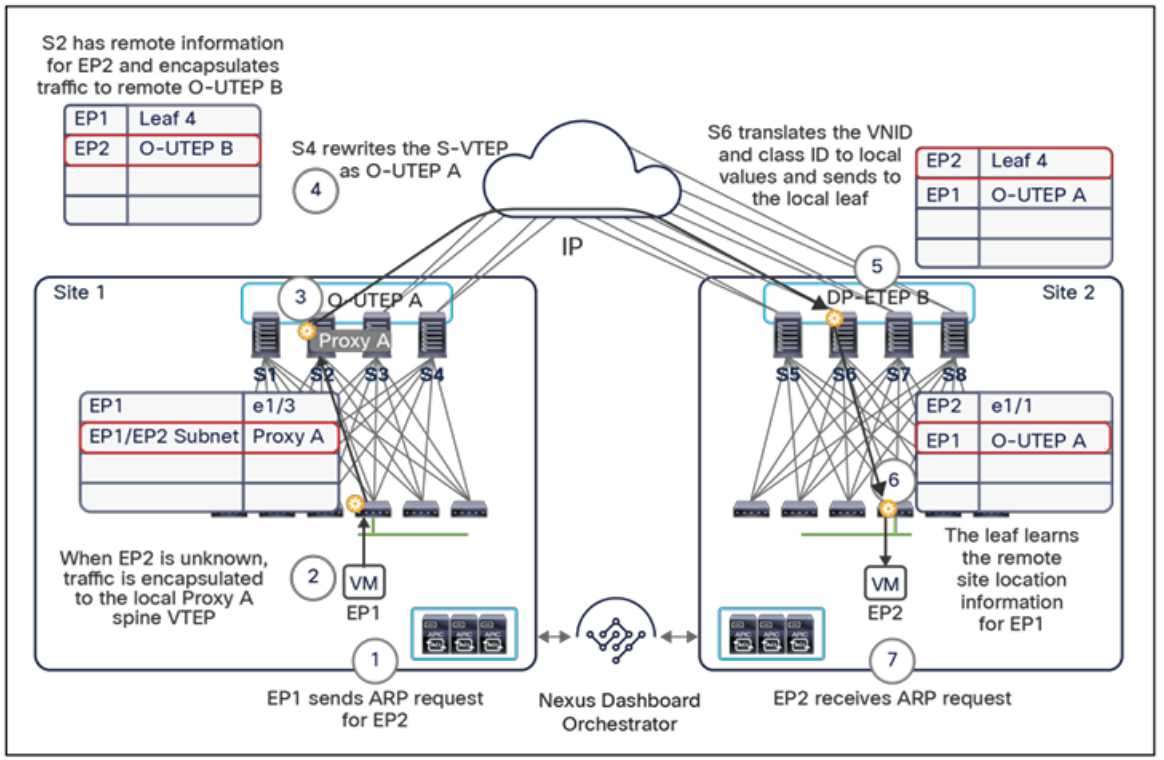


Figure 72.
ARP request across sites without flooding

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
036

	4. The capability of forwarding ARP requests across sites in "unicast mode" is mainly dependent on the knowledge in the COOP database of the IP address of the remote endpoint (information received via the MP-BGP EVPN control plane with the remote spines). Considering initially the case where the IP address of the remote endpoint is known (that is, EP2 is not a "silent host"), the local spine nodes know the remote O-UTEP address identifying the site to which EP2 is connected and can encapsulate the packet and send it across the ISN toward the remote site. It is worth noticing that the spines also rewrite the source IP address of the VXLAN-encapsulated packet, replacing the VTEP address of the leaf node with the local O-UTEP A address identifying the local site. This operation is very important, because, as previously mentioned, when describing the EVPN control-plane exchange across sites, only the EVPN-RID and O-UTEP/O-MTEP addresses of the spine nodes should be visible in the external IP network. https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html. *See also*, *e.g.*, evidence and analysis described above with respect to claims 1 and 2.
[4] The apparatus according to claim 1, wherein the tunnel packet end point obtains a VNID (Virtual Network ID) of the tunnel packet and L3 address information of a sending tunnel end point from the tunnel header of the tunnel packet, processes the tunnel packet such that the L2 switching may be performed, and transmits the processed packet to the domain VSI corresponding to the VNID through a virtual port associated with the sending tunnel end point.	The Accused Instrumentalities include the apparatus according to claim 1, wherein the tunnel packet end point obtains a VNID (Virtual Network ID) of the tunnel packet and L3 address information of a sending tunnel end point from the tunnel header of the tunnel packet, processes the tunnel packet such that the L2 switching may be performed, and transmits the processed packet to the domain VSI corresponding to the VNID through a virtual port associated with the sending tunnel end point. For example, as evidenced below, the Cisco ACI Multi-Site whitepaper describes that inbound VXLAN packets carry both the VNID and the source tunnel endpoint address in their tunnel headers. Cisco explains that "the VXLAN Network Identifier (VNID) identifies the bridge domain (BD) (for Layer 2 communication)," and the BD is the Layer 2 switching instance. Figure 69 shows that when a VXLAN packet is received at the spine, the device decapsulates the packet, extracts the VNID, translates it to the locally significant VNID value for the BD, and "sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain." Additional portions of the whitepaper explain that the tunnel

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
037

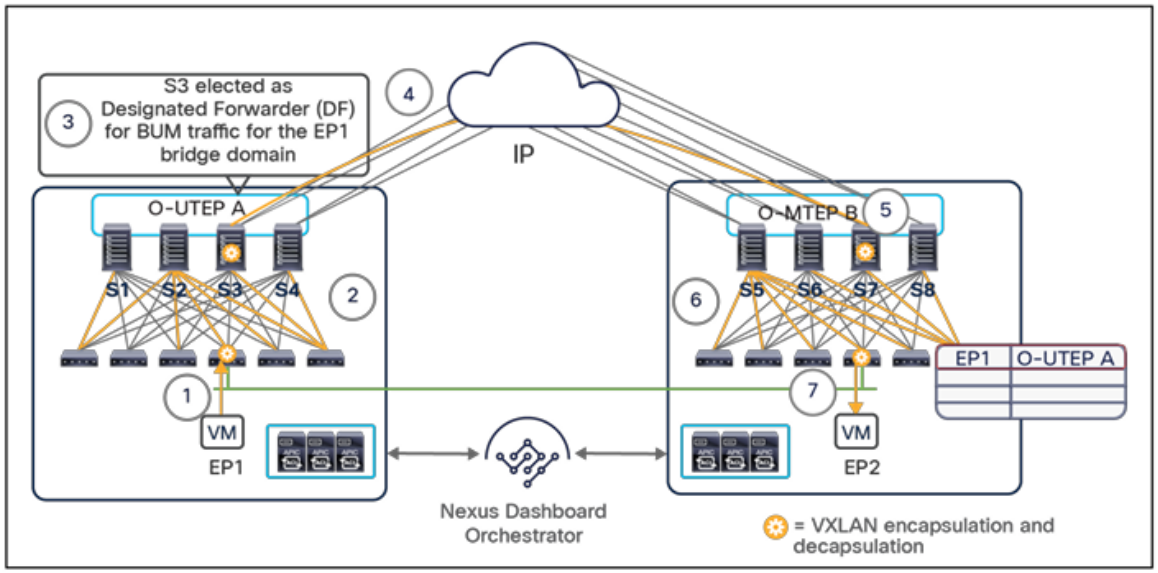
header also contains the L3 address of the sending tunnel endpoint, such as the O-UTEP address associated with the originating site. These operations demonstrate that the accused products: (1) obtain the VNID and L3 address of the sending tunnel endpoint from the VXLAN tunnel header upon receipt of an intersite packet; (2) process the packet by decapsulating it and translating the VNID so that Layer 2 switching may be performed in the BD; and (3) deliver the processed packet into the BD (the domain VSI) through the ingress context associated with the sending tunnel endpoint (i.e., the virtual port mapped to the sender's O-UTEP). This satisfies all elements of Claim 4. *See, e.g.*,

- Intersite control plane: Endpoint reachability information is exchanged across sites using a Multiprotocol-BGP (MP-BGP) Ethernet VPN (EVPN) control plane. This approach allows the exchange of MAC and IP address information for the endpoints that communicate across sites. MP-BGP EVPN sessions are established between the spine nodes deployed in separate fabrics that are managed by the same instance of Cisco Nexus Dashboard Orchestrator, as discussed in more detail in the section "Cisco ACI Multi-Site overlay control plane."
- Intersite data plane: All communication (Layer 2 or Layer 3) between endpoints connected to different sites is achieved by establishing site-to-site Virtual Extensible LAN (VXLAN) tunnels across a generic IP network that interconnects the various sites. As discussed in the section "Intersite connectivity deployment considerations", this IP network has no specific functional requirements other than the capability to support routing and increased Maximum Transmission Unit (MTU) size (given the overhead from the VXLAN encapsulation).

The definition of the O-UTEP and O-MTEP ensures that all the resources connected to a specific site are always seen as reachable from the remote sites through one of those two IP addresses (depending on if it is unicast traffic or BUM/multicast communication). This implies that when observing east-west VXLAN traffic exchanged between two sites in the ISN, all the packets will always be sourced from the O-UTEP address identifying the source site and destined to the O-UTEP (or O-MTEP) of the destination site. Notice that this does not prevent taking advantage of multiple ECMP links that may interconnect the two sites: when building the VXLAN encapsulated packet, a hashing of the L2/L3/L4 headers for the original frame generated by the endpoint is used as a UDP source port in the external header, and this allows you to build "entropy" in the packet. Different application flows between the same sites (and even between the same pair of endpoints) would cause the creation of different UDP source port values: as long as the network devices inside the ISN consider the L4 port information to choose the link to forward traffic, it will hence be possible to load-balance VXLAN packets across multiple paths.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
038

Figure 69 shows the sequence of events required to send a Layer 2 BUM frame across sites.



Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
039

Figure 69.
Layer 2 BUM traffic across sites

1. EP1, belonging to a specific bridge domain, generates a Layer 2 BUM frame.
2. Depending on the type of frame and the corresponding bridge domain settings (as previously clarified above), the leaf may need to flood the traffic in that specific Layer 2 domain. As a consequence, the frame is VXLAN-encapsulated and sent to the specific multicast group (called GIPo) associated with the bridge domain within the fabric along one of the specific multi-destination trees associated to that GIPo, so it can reach all the other leaf and spine nodes.
3. One of the spine nodes connected to the external intersite network is elected as the designated forwarder for that specific bridge domain (this election is held between the spine nodes using IS-IS protocol exchanges). The designated forwarder is responsible for replicating each BUM frame for that bridge domain to all the remote sites with the same stretched bridge domain.
4. The designated forwarder makes copies of the BUM frame and sends it to the remote sites. The destination IP address used when VXLAN encapsulating the packet is the special IP address (O-MTEP) identifying each remote site and is used specifically for the transmission of BUM traffic across sites. The O-MTEP is another anycast IP address defined on all the remote spine nodes that are connected to the intersite network (each site uses a unique O-MTEP address). The source IP address for the VXLAN-encapsulated packet is instead the anycast O-UTEP address deployed on all the local spine nodes connected to the intersite network.

   Note: The O-MTEP (referred to as "Overlay Multicast TEP" in the Cisco Nexus Dashboard Orchestrator GUI) is yet another IP address that must be sent to the Layer 3 network connecting the fabrics.
5. One of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain, and sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain (using the multicast group locally assigned to the BD as destination in the VXLAN header).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
040

The VXLAN Network Identifier (VNID) identifies the bridge domain (BD) (for Layer 2 communication) or the Virtual Routing and Forwarding (VRF) instance (for Layer 3 traffic) of the endpoint sourcing the traffic (for intra-VRF communication). The class ID is the unique identifier of the source Endpoint Group (EPG). However, these values are locally significant within a fabric. Because a completely separate and independent APIC domain and fabric are deployed at the destination site, a translation function (also referred to as "name-space normalization") must be applied before the traffic is forwarded inside the receiving site, to help ensure that locally significant values identifying that same source EPG, bridge domain, and VRF instance are used.

To better understand the need for this name-space normalization function, it is important to clarify what happens when a specific contract is defined between two EPGs deployed in separate sites in order to allow intersite communications between endpoints that are part of those EPGs. As shown in Figure 7, when the desired configuration (intent) is defined in Cisco Nexus Dashboard Orchestrator and then pushed to the different APIC domains, specific copies of EPGs, named shadow EPGs, are automatically created in each APIC domain. This ensures that the whole configuration centrally defined in NDO can be locally instantiated in each site and the security policy properly enforced, even when each EPG is only locally defined and not stretched across sites (specific VNIDs and class IDs are assigned to the shadow objects in each APIC domain).

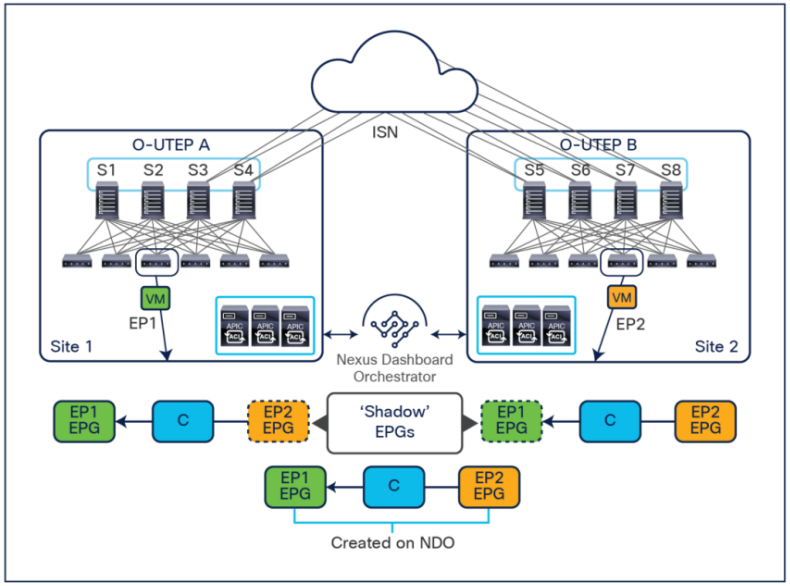


Figure 7.
Creation of shadow EPGs

In the example above, the yellow EP2 EPG (and its associated BD) is created as a "shadow" EPG in APIC domain 1, whereas the green EP1 EPG (and its associated BD) is a "shadow" EPG in APIC domain 2. Up to Cisco ACI Release 5.0(1), the shadow EPGs and BDs are not easily distinguishable in the APIC GUI, so it is quite important to be aware of their existence and role.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
041

Since both APIC domains are completely independent from each other, it is logical to expect that different VNID and class-ID values would be assigned to a given EPG (the "real" and the "shadow" copy) across sites. This implies that a translation of those values is required on the spines receiving data-plane traffic from a remote site before injecting the traffic into the local site, as highlighted in Figure 8, below.

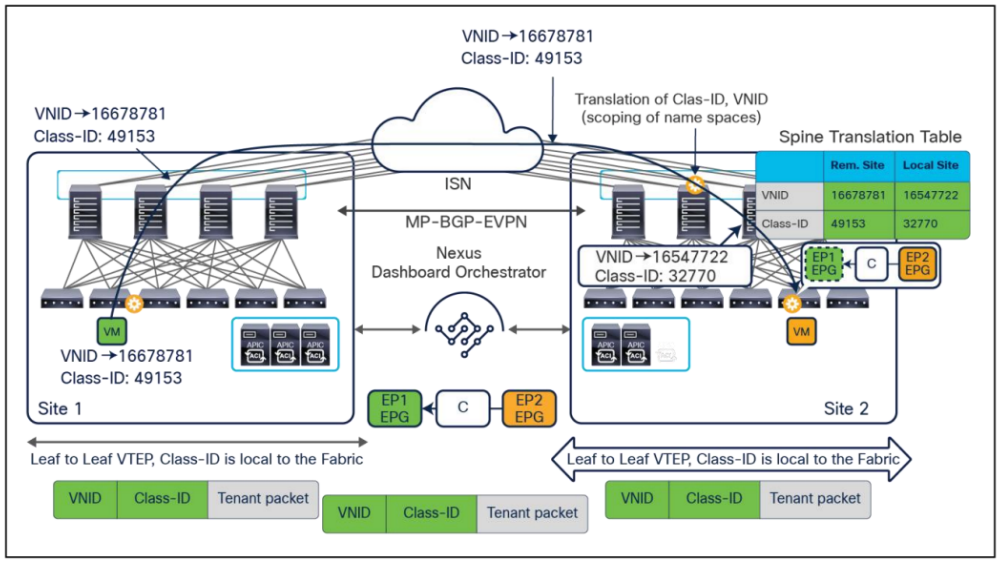


Figure 8.
Name-space translation function on the receiving spine

When the policy is created on Cisco Nexus Dashboard Orchestrator stating that "EP1 EPG" must communicate with "EP2 EPG," the Nexus Dashboard Orchestrator receives from each APIC controller the specific identifiers (pcTag, L2VNI, L3VNI) assigned to the local and shadow objects, and instructs the APIC controllers to program proper translation rules in the local spines. The end result is that the configured policy can then correctly be applied on the leaf node before sending the traffic to the destination endpoint. Additionally, the creation of a contract between EPGs locally deployed in separate sites usually results also in the configuration of the BD subnets associated to the EPGs on the leaf nodes of remote sites (in order to enable the proxy-path via the local spine nodes). More details about this can be found in the "Cisco ACI Multi-Site overlay data plane" section.

Note: The example in Figure 8 shows the policy being applied on the leaf node in the receiving site. This is normally the case until the leaf node in the source site learns via the data-plane the location information of the remote endpoint. From that moment on, the policy can be applied directly on the ingress leaf. The use of service-graph with Policy-Based Redirection (PBR) is a specific case where this may not be always the case, as discussed in more detail in the "Network services integration" section.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
042

Intra-subnet unicast communication across sites

The first requirement before intra-subnet IP communication across sites can be achieved is to complete the ARP exchange between source and destination endpoints. As previously mentioned, ARP represents a special type of Layer 2 broadcast traffic, and its forwarding across sites can be controlled at the bridge domain level.

Note: The assumption for the discussion below is that the bridge domain has associated a single IP subnet, even if there may be scenarios where that is not the case and multiple subnets are defined for the same bridge domain.

There are two different scenarios to consider:

- ARP flooding is enabled in the bridge domain: This is the default configuration for stretched bridge domains created on Cisco Nexus Dashboard Orchestrator with BUM forwarding **enabled**. In this case, the behavior is identical to that discussed in the previous section and shown in Figure 70. The ARP request will reach the destination endpoints in remote sites, which will allow the remote leaf nodes to learn the site location of the source endpoint. As a consequence, the ARP unicast reply will be directly VXLAN-encapsulated to the O-UTEP address identifying the EP1 site, and one of the receiving spine nodes will perform the VNID and class-ID translation and send the frame toward the local leaf node to which EP1 is connected, as shown in Figure 71.

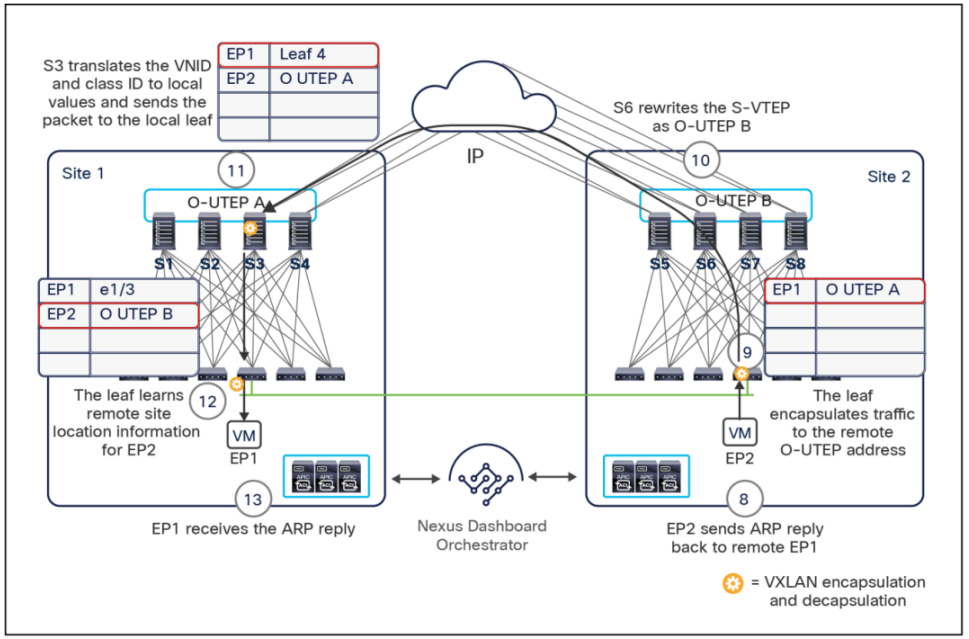


Figure 71.
ARP reply delivered to EP1 in site 1

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
043

4. The capability of forwarding ARP requests across sites in “unicast mode” is mainly dependent on the knowledge in the COOP database of the IP address of the remote endpoint (information received via the MP-BGP EVPN control plane with the remote spines). Considering initially the case where the IP address of the remote endpoint is known (that is, EP2 is not a “silent host”), the local spine nodes know the remote O-UTEP address identifying the site to which EP2 is connected and can encapsulate the packet and send it across the ISN toward the remote site. It is worth noticing that the spines also rewrite the source IP address of the VXLAN-encapsulated packet, replacing the VTEP address of the leaf node with the local O-UTEP A address identifying the local site. This operation is very important, because, as previously mentioned, when describing the EVPN control-plane exchange across sites, only the EVPN-RID and O-UTEP/O-MTEP addresses of the spine nodes should be visible in the external IP network.
5. The VXLAN frame is received by one of the remote spine nodes, which translates the original VNID and class-ID values to locally significant ones and encapsulates the ARP request and sends it toward the local leaf nodes to which EP2 is connected.
6. The leaf node receives the frame, decapsulates it, and learns the class-ID and site location information for remote endpoint EP1.
7. The frame is then sent out the interface on which the EP2 is learned and reaches the endpoint.

- When traffic originates from a source in a site and is forwarded to remote sites through the ISN, the proper translation entries must be created on the spines (for the VRF VNIDs and the EPG class-IDs), as already discussed for the unicast-communication use cases. In addition, if the source bridge domain is not stretched across sites, it is also required to configure the source IP subnet on the leaf nodes of the remote sites, to ensure a successful Reverse Path Forwarding (RPF) check when multicast traffic is received. In order to avoid having to create this configuration for all the EPGs and/or bridge domains that are Layer 3 multicast enabled, but only for the ones with sources connected, the user must explicitly indicate on Cisco Nexus Dashboard Orchestrator what are the EPGs containing multicast sources.

It is important to notice how differently from regular intersite (east-west) communication between endpoints, the spines on the receiving site 1 are not involved in any VNID/Class-ID translations for intersite communication to external prefixes (Figure 111).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
044

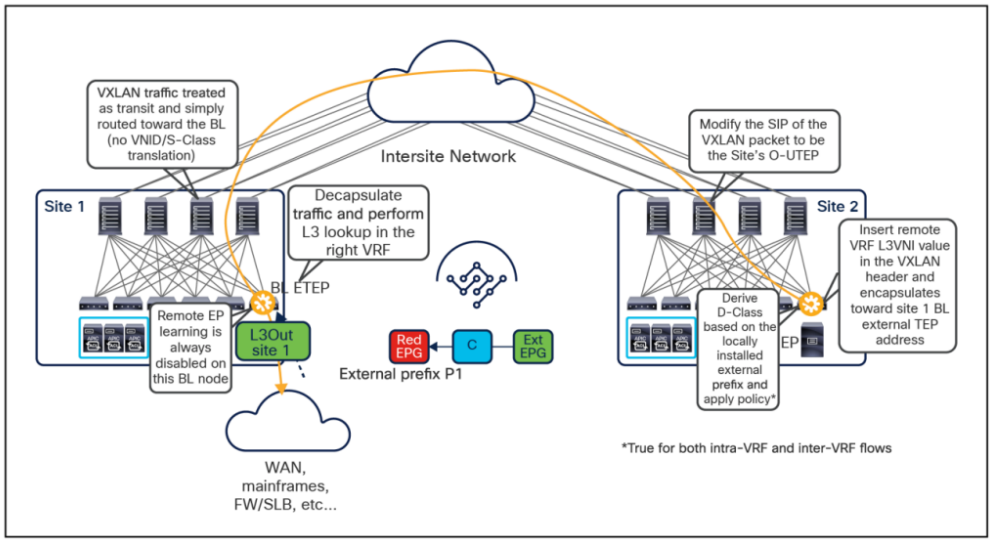


Figure 111.
No need for VNID/S-Class translation on the receiving spines

This is because the VXLAN tunnel is established directly to the BL node and the spines in site 1 (as well as the ones in site 2) just become a routed hop in the underlay infrastructure allowing for the VXLAN communication to happen. This has two consequences:

- The information about the L3VNI value for the L3Out in site 1 must be propagated from site 1 to site 2 via the MP-BGP VPNv4 control plane. This allows the compute leaf in site 2 to add this information in the VXLAN header of the packet and ensures that when the BL node in site 1 receives the traffic, it can perform the Layer 3 lookup into the right VRF before sending the traffic out of the L3Out connection. Depending on the use case, the VRF could be the same of the endpoint in site 2 or a different VRF.
- The security policy is always applied on the compute leaf in site 2, as the class-id for the external EPG deployed in site 1 is locally programmed there.
- The Red EPG endpoint information can never be learned on the BL node in site 1 based on data-plane activity (as normally happens); this is because, as previously mentioned, no class-ID/VNI translation happens on the spines in site 1 (since the traffic is destined directly to the VTEP of the BL node), therefore the source class-id identifying the Red EPG in site 2 may have no meaning (or a complete different meaning) in the APIC domain in site 1.

The return flow from site 1 to site 2 is depicted in Figure 112, below.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
045

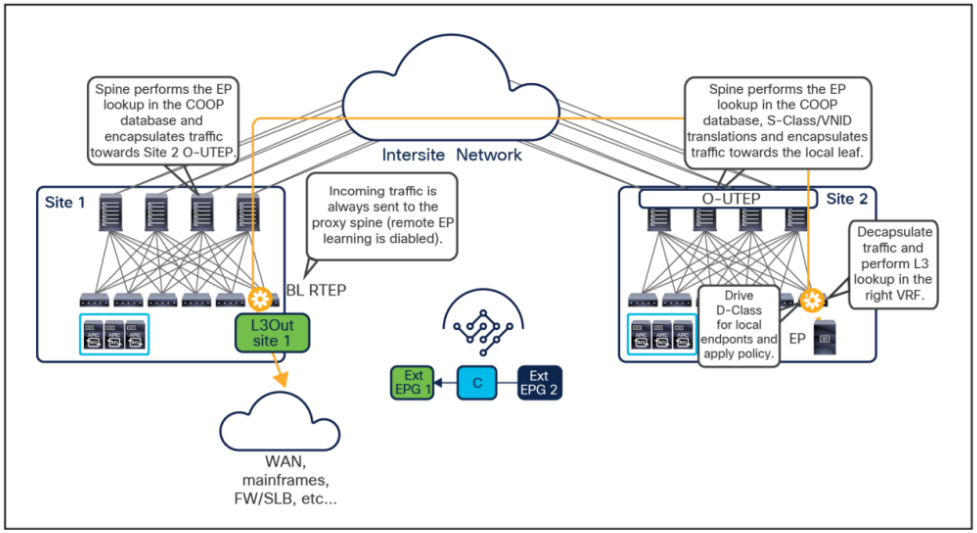


Figure 112.
Return traffic from the L3Out to the remote endpoint

The following are some specific points to highlight:

- Traffic received from the L3Out connection and destined to the remote endpoint part of the Red EPG is always encapsulated by the BL node toward the local proxy spine service, since as previously mentioned there will never be specific endpoint information learned on the BL node. The BL node will insert in the VXLAN header the L3VNI identifying the VRF of the endpoint destination assigned in the local site.
- The local spine would then encapsulate to the spine nodes in the remote site where the destination endpoint has been discovered. This is the normal behavior for intersite communication already described in the "Cisco ACI Multi-Site overlay data plane" section and allows the receiving spines to perform the normal S-Class/VNID translation services that are normally done for east-west communication between endpoints.
- The receiving compute leaf node will then be able to perform the Layer 3 lookup in the right VRF, apply the policy and, if allowed, forward the traffic toward the destination endpoint.

https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
046

	See also, *e.g.*, evidence and analysis described above with respect to claim 1.
[5] The apparatus according to claim 1, wherein the tunnel packet generator is configured, for the 5 packet for which the L2 switching has been performed, to generate a VN header (virtual network header) including a VNID (Virtual Network ID) corresponding to the domain VSI and a L3 tunnel header based on a destination tunnel end point L3 address obtained from a destination virtual port of the packet for which the L2 switching has been performed, and to combine the VN header and the L3 tunnel header with the processed packet for which the L2 switching has been performed, and transmit the same.	The Accused Instrumentalities include the apparatus according to claim 1, wherein the tunnel packet generator is configured, for the 5 packet for which the L2 switching has been performed, to generate a VN header (virtual network header) including a VNID (Virtual Network ID) corresponding to the domain VSI and a L3 tunnel header based on a destination tunnel end point L3 address obtained from a destination virtual port of the packet for which the L2 switching has been performed, and to combine the VN header and the L3 tunnel header with the processed packet for which the L2 switching has been performed, and transmit the same. For example, as evidenced below, Cisco ACI Multi-Site encapsulates intersite traffic using VXLAN headers that contain both the VXLAN Network Identifier (VNID) and the L3 tunnel endpoint addresses (O-UTEP or O-MTEP). The Cisco whitepaper states that "the VXLAN Network Identifier (VNID) identifies the bridge domain (BD) (for Layer 2 communication)," confirming that the VNID in the VXLAN header identifies the domain VSI used for switching the processed packet. Cisco further explains that the spine and leaf nodes "rewrite the source IP address of the VXLAN-encapsulated packet" and that, for intersite unicast flows, "the leaf encapsulates traffic to the remote O-UTEP address," demonstrating that the tunnel packet generator creates an outer L3 header based on the destination tunnel endpoint's L3 address. As further evidenced below, figures 69, 71, 111, and 112 all depict the tunnel packet generation process, showing the addition of the VXLAN header—including VNID and O-UTEP/O-MTEP tunnel endpoint addresses—to the packet after Layer 2 switching has occurred in the Bridge Domain. The resulting VXLAN-encapsulated packet is then transmitted across the intersite network. These operations satisfy Claim 5's requirement that the tunnel packet generator combine a VNID-based virtual network header and an L3 tunnel header with the processed packet and transmit the new tunnel packet. *See, e.g.*,

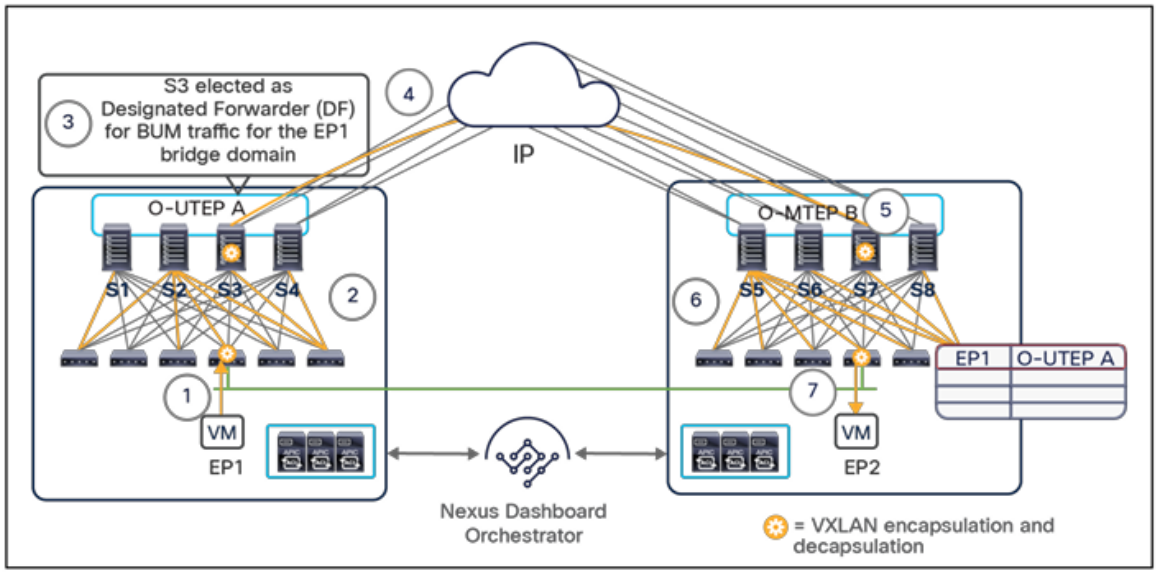
Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
047

- Intersite control plane: Endpoint reachability information is exchanged across sites using a Multiprotocol-BGP (MP-BGP) Ethernet VPN (EVPN) control plane. This approach allows the exchange of MAC and IP address information for the endpoints that communicate across sites. MP-BGP EVPN sessions are established between the spine nodes deployed in separate fabrics that are managed by the same instance of Cisco Nexus Dashboard Orchestrator, as discussed in more detail in the section "Cisco ACI Multi-Site overlay control plane."
- Intersite data plane: All communication (Layer 2 or Layer 3) between endpoints connected to different sites is achieved by establishing site-to-site Virtual Extensible LAN (VXLAN) tunnels across a generic IP network that interconnects the various sites. As discussed in the section "Intersite connectivity deployment considerations", this IP network has no specific functional requirements other than the capability to support routing and increased Maximum Transmission Unit (MTU) size (given the overhead from the VXLAN encapsulation).

The definition of the O-UTEP and O-MTEP ensures that all the resources connected to a specific site are always seen as reachable from the remote sites through one of those two IP addresses (depending on if it is unicast traffic or BUM/multicast communication). This implies that when observing east-west VXLAN traffic exchanged between two sites in the ISN, all the packets will always be sourced from the O-UTEP address identifying the source site and destined to the O-UTEP (or O-MTEP) of the destination site. Notice that this does not prevent taking advantage of multiple ECMP links that may interconnect the two sites: when building the VXLAN encapsulated packet, a hashing of the L2/L3/L4 headers for the original frame generated by the endpoint is used as a UDP source port in the external header, and this allows you to build "entropy" in the packet. Different application flows between the same sites (and even between the same pair of endpoints) would cause the creation of different UDP source port values: as long as the network devices inside the ISN consider the L4 port information to choose the link to forward traffic, it will hence be possible to load-balance VXLAN packets across multiple paths.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
048

Figure 69 shows the sequence of events required to send a Layer 2 BUM frame across sites.



Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
049

Figure 69.
Layer 2 BUM traffic across sites

1. EP1, belonging to a specific bridge domain, generates a Layer 2 BUM frame.
2. Depending on the type of frame and the corresponding bridge domain settings (as previously clarified above), the leaf may need to flood the traffic in that specific Layer 2 domain. As a consequence, the frame is VXLAN-encapsulated and sent to the specific multicast group (called GIPo) associated with the bridge domain within the fabric along one of the specific multi-destination trees associated to that GIPo, so it can reach all the other leaf and spine nodes.
3. One of the spine nodes connected to the external intersite network is elected as the designated forwarder for that specific bridge domain (this election is held between the spine nodes using IS-IS protocol exchanges). The designated forwarder is responsible for replicating each BUM frame for that bridge domain to all the remote sites with the same stretched bridge domain.
4. The designated forwarder makes copies of the BUM frame and sends it to the remote sites. The destination IP address used when VXLAN encapsulating the packet is the special IP address (O-MTEP) identifying each remote site and is used specifically for the transmission of BUM traffic across sites. The O-MTEP is another anycast IP address defined on all the remote spine nodes that are connected to the intersite network (each site uses a unique O-MTEP address). The source IP address for the VXLAN-encapsulated packet is instead the anycast O-UTEP address deployed on all the local spine nodes connected to the intersite network.

   Note: The O-MTEP (referred to as "Overlay Multicast TEP" in the Cisco Nexus Dashboard Orchestrator GUI) is yet another IP address that must be sent to the Layer 3 network connecting the fabrics.
5. One of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain, and sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain (using the multicast group locally assigned to the BD as destination in the VXLAN header).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
050

The VXLAN Network Identifier (VNID) identifies the bridge domain (BD) (for Layer 2 communication) or the Virtual Routing and Forwarding (VRF) instance (for Layer 3 traffic) of the endpoint sourcing the traffic (for intra-VRF communication). The class ID is the unique identifier of the source Endpoint Group (EPG). However, these values are locally significant within a fabric. Because a completely separate and independent APIC domain and fabric are deployed at the destination site, a translation function (also referred to as "name-space normalization") must be applied before the traffic is forwarded inside the receiving site, to help ensure that locally significant values identifying that same source EPG, bridge domain, and VRF instance are used.

To better understand the need for this name-space normalization function, it is important to clarify what happens when a specific contract is defined between two EPGs deployed in separate sites in order to allow intersite communications between endpoints that are part of those EPGs. As shown in Figure 7, when the desired configuration (intent) is defined in Cisco Nexus Dashboard Orchestrator and then pushed to the different APIC domains, specific copies of EPGs, named shadow EPGs, are automatically created in each APIC domain. This ensures that the whole configuration centrally defined in NDO can be locally instantiated in each site and the security policy properly enforced, even when each EPG is only locally defined and not stretched across sites (specific VNIDs and class IDs are assigned to the shadow objects in each APIC domain).

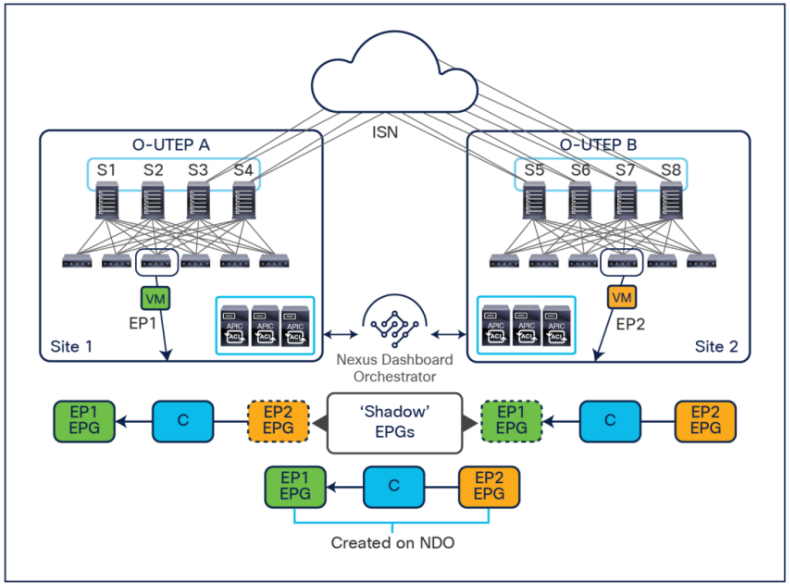


Figure 7.
Creation of shadow EPGs

In the example above, the yellow EP2 EPG (and its associated BD) is created as a "shadow" EPG in APIC domain 1, whereas the green EP1 EPG (and its associated BD) is a "shadow" EPG in APIC domain 2. Up to Cisco ACI Release 5.0(1), the shadow EPGs and BDs are not easily distinguishable in the APIC GUI, so it is quite important to be aware of their existence and role.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
051

Since both APIC domains are completely independent from each other, it is logical to expect that different VNID and class-ID values would be assigned to a given EPG (the "real" and the "shadow" copy) across sites. This implies that a translation of those values is required on the spines receiving data-plane traffic from a remote site before injecting the traffic into the local site, as highlighted in Figure 8, below.

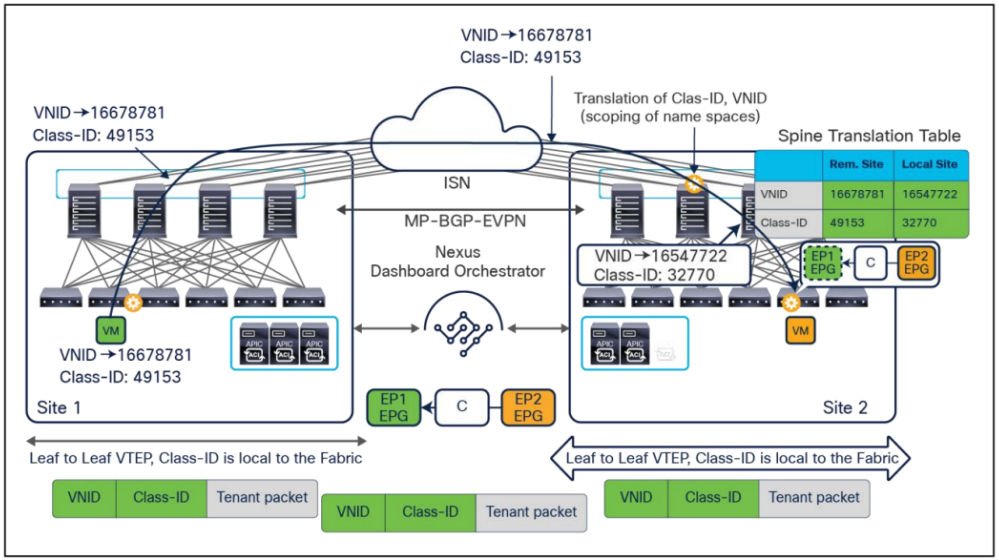


Figure 8.
Name-space translation function on the receiving spine

When the policy is created on Cisco Nexus Dashboard Orchestrator stating that "EP1 EPG" must communicate with "EP2 EPG," the Nexus Dashboard Orchestrator receives from each APIC controller the specific identifiers (pcTag, L2VNI, L3VNI) assigned to the local and shadow objects, and instructs the APIC controllers to program proper translation rules in the local spines. The end result is that the configured policy can then correctly be applied on the leaf node before sending the traffic to the destination endpoint. Additionally, the creation of a contract between EPGs locally deployed in separate sites usually results also in the configuration of the BD subnets associated to the EPGs on the leaf nodes of remote sites (in order to enable the proxy-path via the local spine nodes). More details about this can be found in the "Cisco ACI Multi-Site overlay data plane" section.

Note: The example in Figure 8 shows the policy being applied on the leaf node in the receiving site. This is normally the case until the leaf node in the source site learns via the data-plane the location information of the remote endpoint. From that moment on, the policy can be applied directly on the ingress leaf. The use of service-graph with Policy-Based Redirection (PBR) is a specific case where this may not be always the case, as discussed in more detail in the "Network services integration" section.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
052

Intra-subnet unicast communication across sites

The first requirement before intra-subnet IP communication across sites can be achieved is to complete the ARP exchange between source and destination endpoints. As previously mentioned, ARP represents a special type of Layer 2 broadcast traffic, and its forwarding across sites can be controlled at the bridge domain level.

Note: The assumption for the discussion below is that the bridge domain has associated a single IP subnet, even if there may be scenarios where that is not the case and multiple subnets are defined for the same bridge domain.

There are two different scenarios to consider:

- ARP flooding is enabled in the bridge domain: This is the default configuration for stretched bridge domains created on Cisco Nexus Dashboard Orchestrator with BUM forwarding **enabled**. In this case, the behavior is identical to that discussed in the previous section and shown in Figure 70. The ARP request will reach the destination endpoints in remote sites, which will allow the remote leaf nodes to learn the site location of the source endpoint. As a consequence, the ARP unicast reply will be directly VXLAN-encapsulated to the O-UTEP address identifying the EP1 site, and one of the receiving spine nodes will perform the VNID and class-ID translation and send the frame toward the local leaf node to which EP1 is connected, as shown in Figure 71.

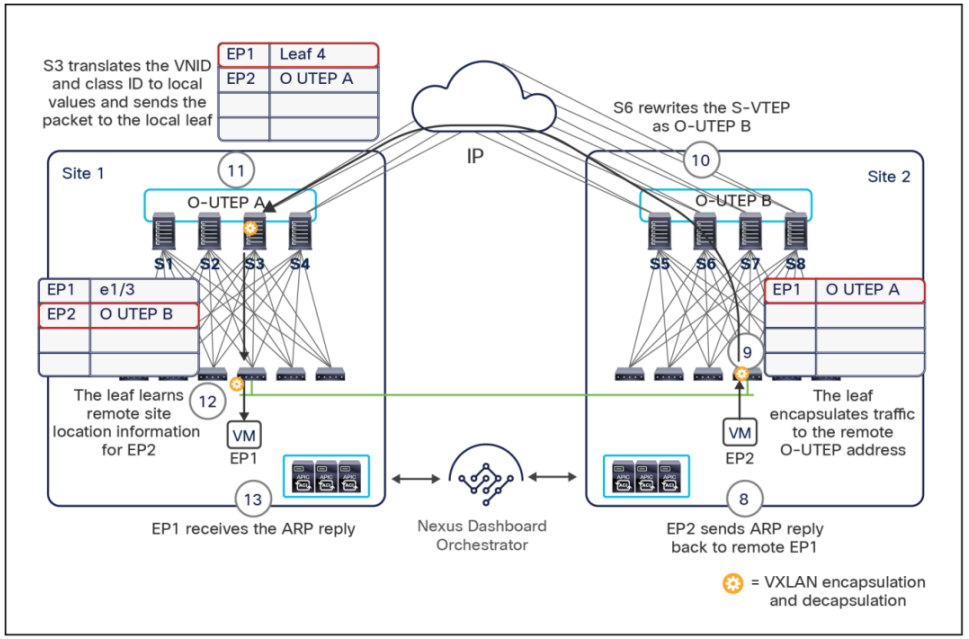


Figure 71.
ARP reply delivered to EP1 in site 1

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
053

4. The capability of forwarding ARP requests across sites in “unicast mode” is mainly dependent on the knowledge in the COOP database of the IP address of the remote endpoint (information received via the MP-BGP EVPN control plane with the remote spines). Considering initially the case where the IP address of the remote endpoint is known (that is, EP2 is not a “silent host”), the local spine nodes know the remote O-UTEP address identifying the site to which EP2 is connected and can encapsulate the packet and send it across the ISN toward the remote site. It is worth noticing that the spines also rewrite the source IP address of the VXLAN-encapsulated packet, replacing the VTEP address of the leaf node with the local O-UTEP A address identifying the local site. This operation is very important, because, as previously mentioned, when describing the EVPN control-plane exchange across sites, only the EVPN-RID and O-UTEP/O-MTEP addresses of the spine nodes should be visible in the external IP network.
5. The VXLAN frame is received by one of the remote spine nodes, which translates the original VNID and class-ID values to locally significant ones and encapsulates the ARP request and sends it toward the local leaf nodes to which EP2 is connected.
6. The leaf node receives the frame, decapsulates it, and learns the class-ID and site location information for remote endpoint EP1.
7. The frame is then sent out the interface on which the EP2 is learned and reaches the endpoint.

- When traffic originates from a source in a site and is forwarded to remote sites through the ISN, the proper translation entries must be created on the spines (for the VRF VNIDs and the EPG class-IDs), as already discussed for the unicast-communication use cases. In addition, if the source bridge domain is not stretched across sites, it is also required to configure the source IP subnet on the leaf nodes of the remote sites, to ensure a successful Reverse Path Forwarding (RPF) check when multicast traffic is received. In order to avoid having to create this configuration for all the EPGs and/or bridge domains that are Layer 3 multicast enabled, but only for the ones with sources connected, the user must explicitly indicate on Cisco Nexus Dashboard Orchestrator what are the EPGs containing multicast sources.

It is important to notice how differently from regular intersite (east-west) communication between endpoints, the spines on the receiving site 1 are not involved in any VNID/Class-ID translations for intersite communication to external prefixes (Figure 111).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
054

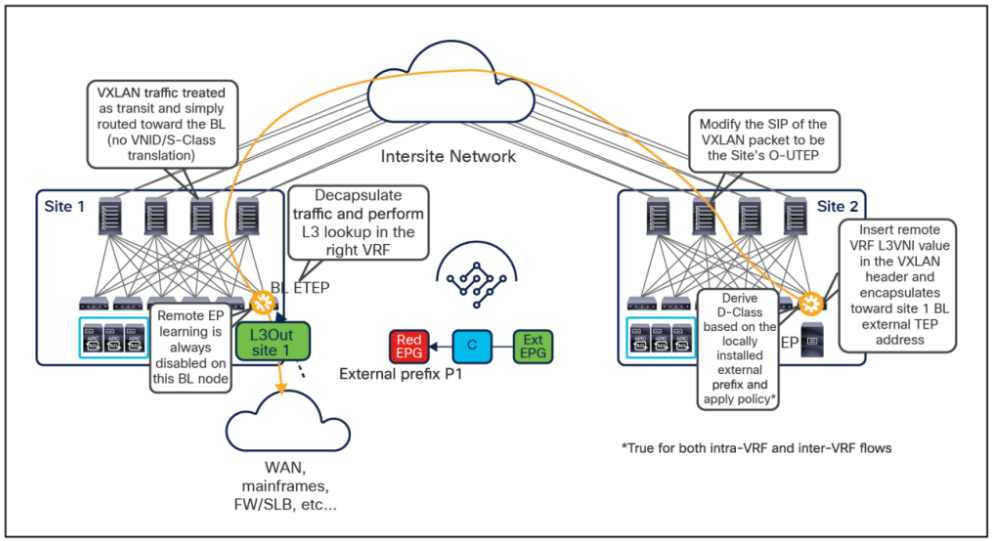


Figure 111.
No need for VNID/S-Class translation on the receiving spines

This is because the VXLAN tunnel is established directly to the BL node and the spines in site 1 (as well as the ones in site 2) just become a routed hop in the underlay infrastructure allowing for the VXLAN communication to happen. This has two consequences:

- The information about the L3VNI value for the L3Out in site 1 must be propagated from site 1 to site 2 via the MP-BGP VPNv4 control plane. This allows the compute leaf in site 2 to add this information in the VXLAN header of the packet and ensures that when the BL node in site 1 receives the traffic, it can perform the Layer 3 lookup into the right VRF before sending the traffic out of the L3Out connection. Depending on the use case, the VRF could be the same of the endpoint in site 2 or a different VRF.
- The security policy is always applied on the compute leaf in site 2, as the class-id for the external EPG deployed in site 1 is locally programmed there.
- The Red EPG endpoint information can never be learned on the BL node in site 1 based on data-plane activity (as normally happens); this is because, as previously mentioned, no class-ID/VNI translation happens on the spines in site 1 (since the traffic is destined directly to the VTEP of the BL node), therefore the source class-id identifying the Red EPG in site 2 may have no meaning (or a complete different meaning) in the APIC domain in site 1.

The return flow from site 1 to site 2 is depicted in Figure 112, below.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
055

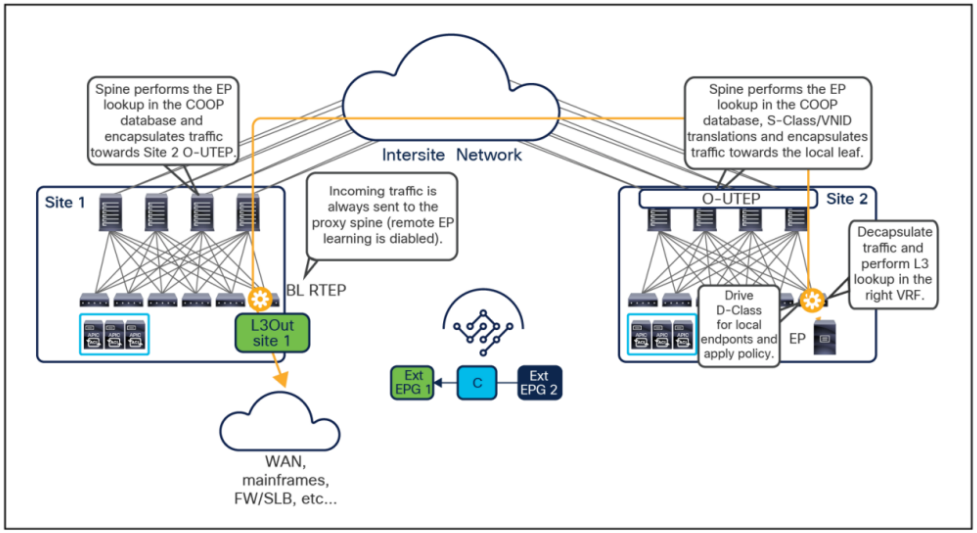


Figure 112.
Return traffic from the L3Out to the remote endpoint

The following are some specific points to highlight:

- Traffic received from the L3Out connection and destined to the remote endpoint part of the Red EPG is always encapsulated by the BL node toward the local proxy spine service, since as previously mentioned there will never be specific endpoint information learned on the BL node. The BL node will insert in the VXLAN header the L3VNI identifying the VRF of the endpoint destination assigned in the local site.
- The local spine would then encapsulate to the spine nodes in the remote site where the destination endpoint has been discovered. This is the normal behavior for intersite communication already described in the “Cisco ACI Multi-Site overlay data plane” section and allows the receiving spines to perform the normal S-Class/VNID translation services that are normally done for east-west communication between endpoints.
- The receiving compute leaf node will then be able to perform the Layer 3 lookup in the right VRF, apply the policy and, if allowed, forward the traffic toward the destination endpoint.

https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
056

	See also, *e.g.*, evidence and analysis described above with respect to claim 1.
[6] The apparatus according to claim 1, further comprising a domain VSI manager configured to receive a VNID (virtual network ID) of a virtual network from a network virtualization manager that manages network virtualization of an entire network and to generate the domain VSI (domain virtual switching instance) corresponding thereto.	The Accused Instrumentalities include the apparatus according to claim 1, further comprising a domain VSI manager configured to receive a VNID (virtual network ID) of a virtual network from a network virtualization manager that manages network virtualization of an entire network and to generate the domain VSI (domain virtual switching instance) corresponding thereto. As evidenced below, the Cisco ACI Multi-Site whitepaper explains that all participating APIC domains and switches receive their VNID assignments from a centralized orchestration system, the Cisco Nexus Dashboard Orchestrator (NDO). Cisco states that the APIC domains are "managed by the same instance of Cisco Nexus Dashboard Orchestrator," ensuring that VNIDs are consistently allocated across sites. The whitepaper also explains that "the VXLAN Network Identifier (VNID) identifies the bridge domain (BD) (for Layer 2 communication)," confirming that each VNID corresponds to a specific BD instantiated on the switches. Because the BD is the Layer 2 switching construct in ACI and serves as the claimed "domain VSI," this evidence shows that the system receives the VNID assigned by the network virtualization manager (NDO/APIC) and generates the corresponding BD instance on each switch. Accordingly, the accused products include a domain VSI manager configured to receive a VNID from the network virtualization manager and to generate the domain VSI corresponding to that VNID, satisfying Claim 6. *See, e.g.*,

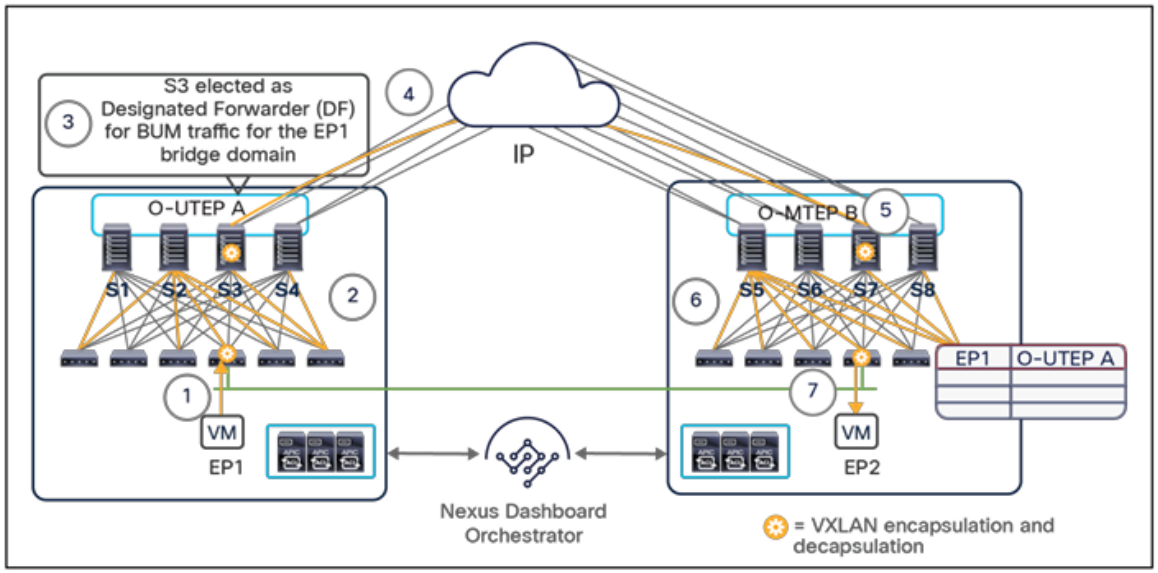
Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
057

- Intersite control plane: Endpoint reachability information is exchanged across sites using a Multiprotocol-BGP (MP-BGP) Ethernet VPN (EVPN) control plane. This approach allows the exchange of MAC and IP address information for the endpoints that communicate across sites. MP-BGP EVPN sessions are established between the spine nodes deployed in separate fabrics that are managed by the same instance of Cisco Nexus Dashboard Orchestrator, as discussed in more detail in the section "Cisco ACI Multi-Site overlay control plane."
- Intersite data plane: All communication (Layer 2 or Layer 3) between endpoints connected to different sites is achieved by establishing site-to-site Virtual Extensible LAN (VXLAN) tunnels across a generic IP network that interconnects the various sites. As discussed in the section "Intersite connectivity deployment considerations", this IP network has no specific functional requirements other than the capability to support routing and increased Maximum Transmission Unit (MTU) size (given the overhead from the VXLAN encapsulation).

The definition of the O-UTEP and O-MTEP ensures that all the resources connected to a specific site are always seen as reachable from the remote sites through one of those two IP addresses (depending on if it is unicast traffic or BUM/multicast communication). This implies that when observing east-west VXLAN traffic exchanged between two sites in the ISN, all the packets will always be sourced from the O-UTEP address identifying the source site and destined to the O-UTEP (or O-MTEP) of the destination site. Notice that this does not prevent taking advantage of multiple ECMP links that may interconnect the two sites: when building the VXLAN encapsulated packet, a hashing of the L2/L3/L4 headers for the original frame generated by the endpoint is used as a UDP source port in the external header, and this allows you to build "entropy" in the packet. Different application flows between the same sites (and even between the same pair of endpoints) would cause the creation of different UDP source port values: as long as the network devices inside the ISN consider the L4 port information to choose the link to forward traffic, it will hence be possible to load-balance VXLAN packets across multiple paths.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
058

Figure 69 shows the sequence of events required to send a Layer 2 BUM frame across sites.



Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
059

Figure 69.
Layer 2 BUM traffic across sites

1. EP1, belonging to a specific bridge domain, generates a Layer 2 BUM frame.
2. Depending on the type of frame and the corresponding bridge domain settings (as previously clarified above), the leaf may need to flood the traffic in that specific Layer 2 domain. As a consequence, the frame is VXLAN-encapsulated and sent to the specific multicast group (called GIPo) associated with the bridge domain within the fabric along one of the specific multi-destination trees associated to that GIPo, so it can reach all the other leaf and spine nodes.
3. One of the spine nodes connected to the external intersite network is elected as the designated forwarder for that specific bridge domain (this election is held between the spine nodes using IS-IS protocol exchanges). The designated forwarder is responsible for replicating each BUM frame for that bridge domain to all the remote sites with the same stretched bridge domain.
4. The designated forwarder makes copies of the BUM frame and sends it to the remote sites. The destination IP address used when VXLAN encapsulating the packet is the special IP address (O-MTEP) identifying each remote site and is used specifically for the transmission of BUM traffic across sites. The O-MTEP is another anycast IP address defined on all the remote spine nodes that are connected to the intersite network (each site uses a unique O-MTEP address). The source IP address for the VXLAN-encapsulated packet is instead the anycast O-UTEP address deployed on all the local spine nodes connected to the intersite network.

   Note: The O-MTEP (referred to as "Overlay Multicast TEP" in the Cisco Nexus Dashboard Orchestrator GUI) is yet another IP address that must be sent to the Layer 3 network connecting the fabrics.
5. One of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain, and sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain (using the multicast group locally assigned to the BD as destination in the VXLAN header).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
060

The VXLAN Network Identifier (VNID) identifies the bridge domain (BD) (for Layer 2 communication) or the Virtual Routing and Forwarding (VRF) instance (for Layer 3 traffic) of the endpoint sourcing the traffic (for intra-VRF communication). The class ID is the unique identifier of the source Endpoint Group (EPG). However, these values are locally significant within a fabric. Because a completely separate and independent APIC domain and fabric are deployed at the destination site, a translation function (also referred to as "name-space normalization") must be applied before the traffic is forwarded inside the receiving site, to help ensure that locally significant values identifying that same source EPG, bridge domain, and VRF instance are used.

To better understand the need for this name-space normalization function, it is important to clarify what happens when a specific contract is defined between two EPGs deployed in separate sites in order to allow intersite communications between endpoints that are part of those EPGs. As shown in Figure 7, when the desired configuration (intent) is defined in Cisco Nexus Dashboard Orchestrator and then pushed to the different APIC domains, specific copies of EPGs, named shadow EPGs, are automatically created in each APIC domain. This ensures that the whole configuration centrally defined in NDO can be locally instantiated in each site and the security policy properly enforced, even when each EPG is only locally defined and not stretched across sites (specific VNIDs and class IDs are assigned to the shadow objects in each APIC domain).

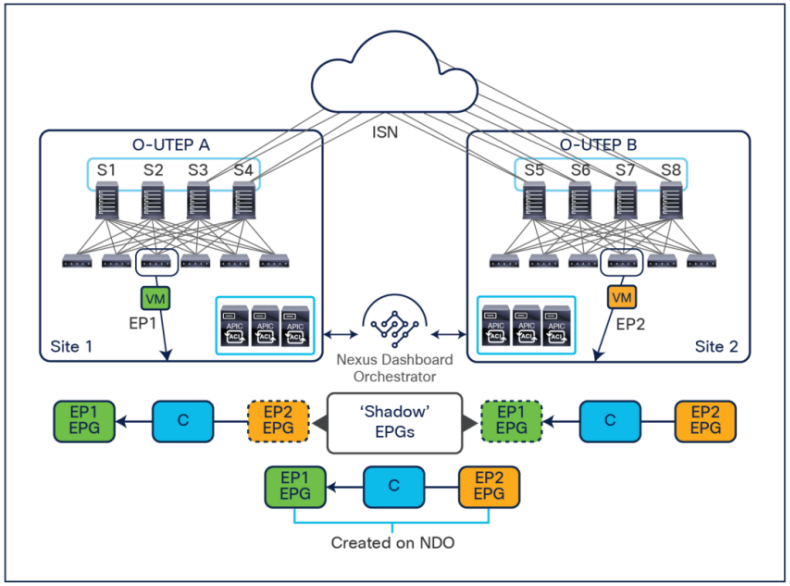


Figure 7.
Creation of shadow EPGs

In the example above, the yellow EP2 EPG (and its associated BD) is created as a "shadow" EPG in APIC domain 1, whereas the green EP1 EPG (and its associated BD) is a "shadow" EPG in APIC domain 2. Up to Cisco ACI Release 5.0(1), the shadow EPGs and BDs are not easily distinguishable in the APIC GUI, so it is quite important to be aware of their existence and role.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
061

Since both APIC domains are completely independent from each other, it is logical to expect that different VNID and class-ID values would be assigned to a given EPG (the "real" and the "shadow" copy) across sites. This implies that a translation of those values is required on the spines receiving data-plane traffic from a remote site before injecting the traffic into the local site, as highlighted in Figure 8, below.

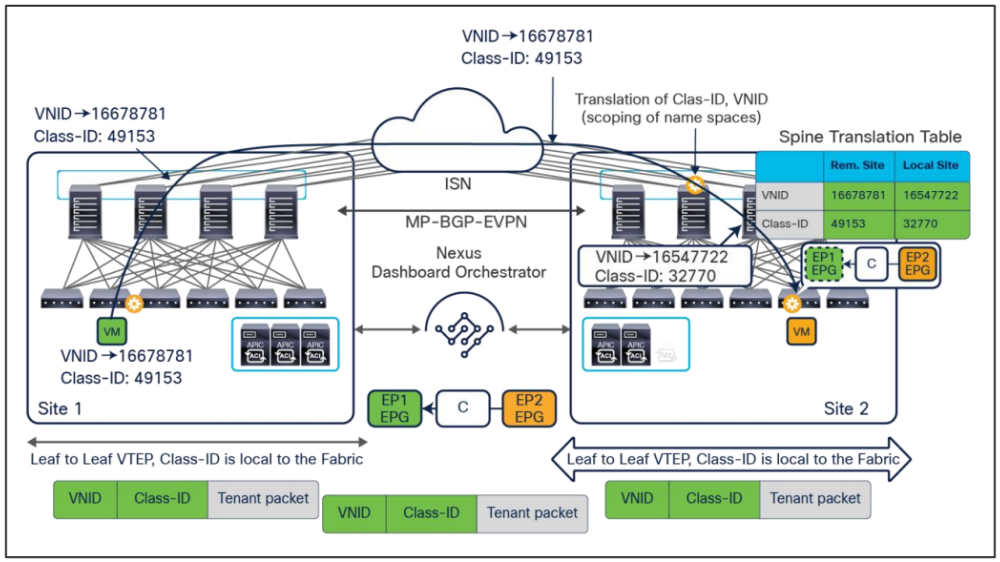


Figure 8.
Name-space translation function on the receiving spine

When the policy is created on Cisco Nexus Dashboard Orchestrator stating that "EP1 EPG" must communicate with "EP2 EPG," the Nexus Dashboard Orchestrator receives from each APIC controller the specific identifiers (pcTag, L2VNI, L3VNI) assigned to the local and shadow objects, and instructs the APIC controllers to program proper translation rules in the local spines. The end result is that the configured policy can then correctly be applied on the leaf node before sending the traffic to the destination endpoint. Additionally, the creation of a contract between EPGs locally deployed in separate sites usually results also in the configuration of the BD subnets associated to the EPGs on the leaf nodes of remote sites (in order to enable the proxy-path via the local spine nodes). More details about this can be found in the "Cisco ACI Multi-Site overlay data plane" section.

Note: The example in Figure 8 shows the policy being applied on the leaf node in the receiving site. This is normally the case until the leaf node in the source site learns via the data-plane the location information of the remote endpoint. From that moment on, the policy can be applied directly on the ingress leaf. The use of service-graph with Policy-Based Redirection (PBR) is a specific case where this may not be always the case, as discussed in more detail in the "Network services integration" section.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
062

Intra-subnet unicast communication across sites

The first requirement before intra-subnet IP communication across sites can be achieved is to complete the ARP exchange between source and destination endpoints. As previously mentioned, ARP represents a special type of Layer 2 broadcast traffic, and its forwarding across sites can be controlled at the bridge domain level.

Note: The assumption for the discussion below is that the bridge domain has associated a single IP subnet, even if there may be scenarios where that is not the case and multiple subnets are defined for the same bridge domain.

There are two different scenarios to consider:

- ARP flooding is enabled in the bridge domain: This is the default configuration for stretched bridge domains created on Cisco Nexus Dashboard Orchestrator with BUM forwarding **enabled**. In this case, the behavior is identical to that discussed in the previous section and shown in Figure 70. The ARP request will reach the destination endpoints in remote sites, which will allow the remote leaf nodes to learn the site location of the source endpoint. As a consequence, the ARP unicast reply will be directly VXLAN-encapsulated to the O-UTEP address identifying the EP1 site, and one of the receiving spine nodes will perform the VNID and class-ID translation and send the frame toward the local leaf node to which EP1 is connected, as shown in Figure 71.

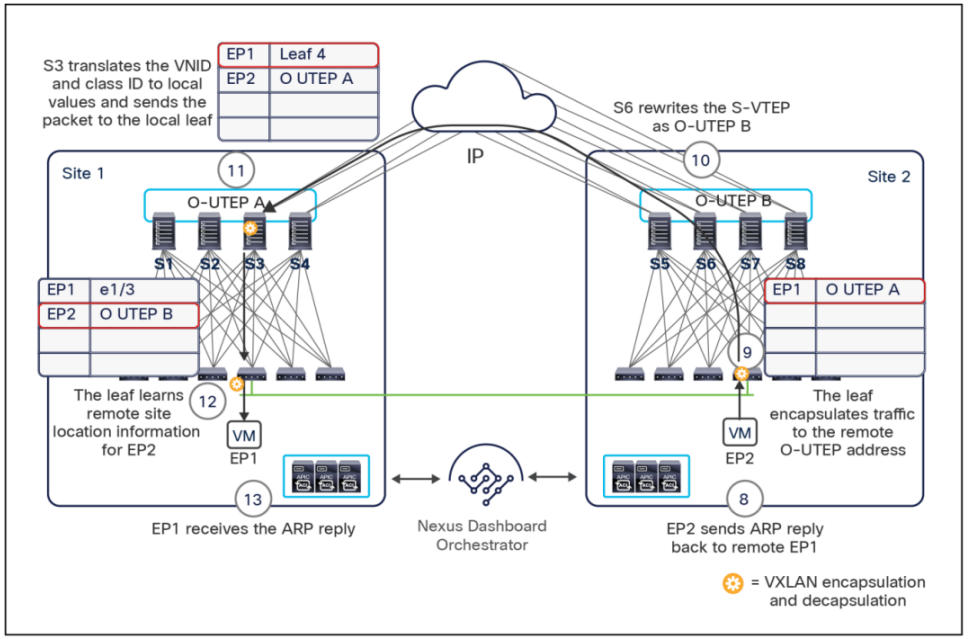


Figure 71.
ARP reply delivered to EP1 in site 1

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
063

4. The capability of forwarding ARP requests across sites in “unicast mode” is mainly dependent on the knowledge in the COOP database of the IP address of the remote endpoint (information received via the MP-BGP EVPN control plane with the remote spines). Considering initially the case where the IP address of the remote endpoint is known (that is, EP2 is not a “silent host”), the local spine nodes know the remote O-UTEP address identifying the site to which EP2 is connected and can encapsulate the packet and send it across the ISN toward the remote site. It is worth noticing that the spines also rewrite the source IP address of the VXLAN-encapsulated packet, replacing the VTEP address of the leaf node with the local O-UTEP A address identifying the local site. This operation is very important, because, as previously mentioned, when describing the EVPN control-plane exchange across sites, only the EVPN-RID and O-UTEP/O-MTEP addresses of the spine nodes should be visible in the external IP network.
5. The VXLAN frame is received by one of the remote spine nodes, which translates the original VNID and class-ID values to locally significant ones and encapsulates the ARP request and sends it toward the local leaf nodes to which EP2 is connected.
6. The leaf node receives the frame, decapsulates it, and learns the class-ID and site location information for remote endpoint EP1.
7. The frame is then sent out the interface on which the EP2 is learned and reaches the endpoint.

- When traffic originates from a source in a site and is forwarded to remote sites through the ISN, the proper translation entries must be created on the spines (for the VRF VNIDs and the EPG class-IDs), as already discussed for the unicast-communication use cases. In addition, if the source bridge domain is not stretched across sites, it is also required to configure the source IP subnet on the leaf nodes of the remote sites, to ensure a successful Reverse Path Forwarding (RPF) check when multicast traffic is received. In order to avoid having to create this configuration for all the EPGs and/or bridge domains that are Layer 3 multicast enabled, but only for the ones with sources connected, the user must explicitly indicate on Cisco Nexus Dashboard Orchestrator what are the EPGs containing multicast sources.

It is important to notice how differently from regular intersite (east-west) communication between endpoints, the spines on the receiving site 1 are not involved in any VNID/Class-ID translations for intersite communication to external prefixes (Figure 111).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
064

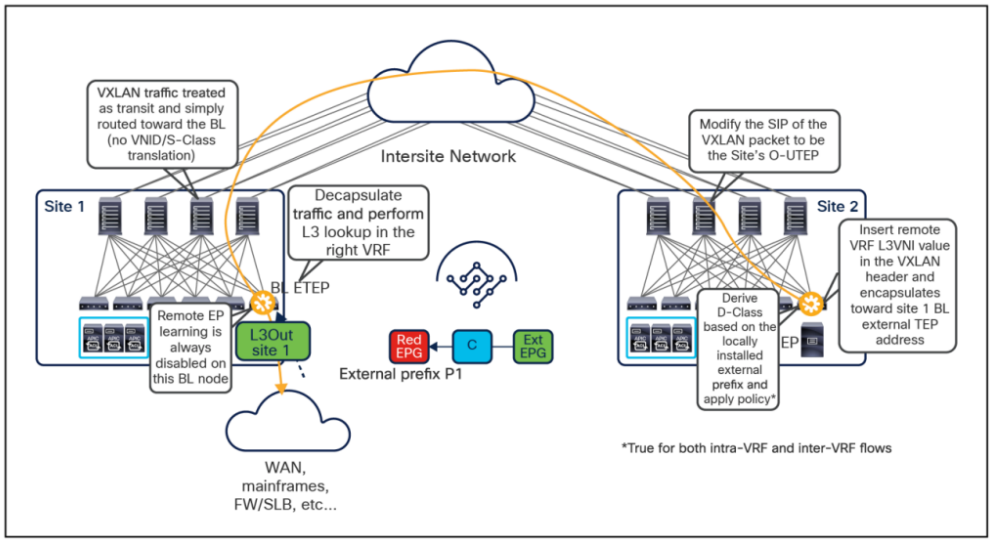


Figure 111.
No need for VNID/S-Class translation on the receiving spines

This is because the VXLAN tunnel is established directly to the BL node and the spines in site 1 (as well as the ones in site 2) just become a routed hop in the underlay infrastructure allowing for the VXLAN communication to happen. This has two consequences:

- The information about the L3VNI value for the L3Out in site 1 must be propagated from site 1 to site 2 via the MP-BGP VPNv4 control plane. This allows the compute leaf in site 2 to add this information in the VXLAN header of the packet and ensures that when the BL node in site 1 receives the traffic, it can perform the Layer 3 lookup into the right VRF before sending the traffic out of the L3Out connection. Depending on the use case, the VRF could be the same of the endpoint in site 2 or a different VRF.
- The security policy is always applied on the compute leaf in site 2, as the class-id for the external EPG deployed in site 1 is locally programmed there.
- The Red EPG endpoint information can never be learned on the BL node in site 1 based on data-plane activity (as normally happens); this is because, as previously mentioned, no class-ID/VNI translation happens on the spines in site 1 (since the traffic is destined directly to the VTEP of the BL node), therefore the source class-id identifying the Red EPG in site 2 may have no meaning (or a complete different meaning) in the APIC domain in site 1.

The return flow from site 1 to site 2 is depicted in Figure 112, below.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
065

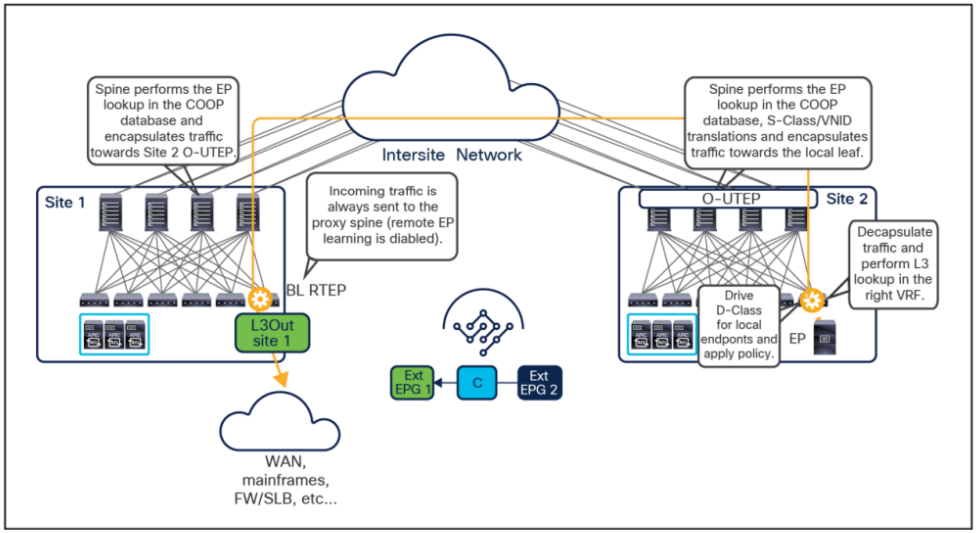


Figure 112.
Return traffic from the L3Out to the remote endpoint

The following are some specific points to highlight:

- Traffic received from the L3Out connection and destined to the remote endpoint part of the Red EPG is always encapsulated by the BL node toward the local proxy spine service, since as previously mentioned there will never be specific endpoint information learned on the BL node. The BL node will insert in the VXLAN header the L3VNI identifying the VRF of the endpoint destination assigned in the local site.
- The local spine would then encapsulate to the spine nodes in the remote site where the destination endpoint has been discovered. This is the normal behavior for intersite communication already described in the “Cisco ACI Multi-Site overlay data plane” section and allows the receiving spines to perform the normal S-Class/VNID translation services that are normally done for east-west communication between endpoints.
- The receiving compute leaf node will then be able to perform the Layer 3 lookup in the right VRF, apply the policy and, if allowed, forward the traffic toward the destination endpoint.

https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
066

	See also, *e.g.*, evidence and analysis described above with respect to claim 1.
[7] An operating method of a network virtualization apparatus, the method comprising:	To the extent that the preamble is limiting, the Accused Instrumentalities perform an operating method of a network virtualization apparatus. *See, e.g.*, *See* evidence and analysis above, e.g., limitation 1[pre].
[7a] collecting tunnel end point information within a network virtualization domain, and connecting each tunnel end point with a tunnel based on the tunnel end point information collected;	The Accused Instrumentalities perform an operating method of a network virtualization apparatus, the method comprising "collecting tunnel end point information within a network virtualization domain, and connecting each tunnel end point with a tunnel based on the tunnel end point information collected." *See, e.g.,* *See* evidence and analysis above, e.g., limitation 1[a].
[7b] receiving a tunnel packet and processing the packet such that a L2 switching may be performed;	The Accused Instrumentalities perform an operating method of a network virtualization apparatus, the method comprising "receiving a tunnel packet and processing the packet such that a L2 switching may be performed." *See, e.g.,* *See* evidence and analysis above, e.g., limitation 1[b].
[7c] performing the L2 switching on the processed packet;	The Accused Instrumentalities perform an operating method of a network virtualization apparatus, the method comprising "performing the L2 switching on the processed packet." *See, e.g.,* *See* evidence and analysis above, e.g., limitation 1[b], 1[c].
[7d] adding a tunnel header to the packet for which the L2 switching has been performed to generate a new tunnel packet, and transmitting the new tunnel packet; and	The Accused Instrumentalities perform an operating method of a network virtualization apparatus, the method comprising "adding a tunnel header to the packet for which the L2 switching has been performed to generate a new tunnel packet, and transmitting the new tunnel packet." *See, e.g.,* *See* evidence and analysis above, e.g., limitation 1[d].
[7e] connecting tunnel end points within the network virtualization domain to other, tunnel	The Accused Instrumentalities perform an operating method of a network virtualization apparatus, the method comprising connecting tunnel end points within the network

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
067

end points within another network virtualization network domain with a unicast tunnel using the new tunnel packet.	virtualization domain to other, tunnel end points within another network virtualization network domain with a unicast tunnel using the new tunnel packet. For example, as evidenced below, the Cisco ACI Multi-Site whitepaper explains that intersite communication between different ACI fabrics is carried through VXLAN tunnels that connect the tunnel endpoints of one site to the tunnel endpoints of another. In the "intersite control plane" section, Cisco states that endpoint reachability information is exchanged across separate sites using MP-BGP EVPN sessions established between the spine nodes of each fabric. The accompanying "intersite data plane" bullet further explains that "all communication (Layer 2 or Layer 3) between endpoints connected to different sites is achieved by establishing site-to-site Virtual Extensible LAN (VXLAN) tunnels across a generic IP network that interconnects the various sites." These VXLAN tunnels are formed between the O-UTEP tunnel endpoint address identifying the source site and the O-UTEP (or O-MTEP) address identifying the remote site. Figures 69, 94, and 95, together with their associated text, depict this intersite connectivity, showing that the designated forwarder spine VXLAN-encapsulates outgoing frames and "sends [them] to the remote sites" using the O-UTEP or O-MTEP address of the remote tunnel endpoint. Additional sections of the whitepaper describe the unicast scenario in which the system encapsulates traffic "to the remote O-UTEP address" for delivery to the remote site, confirming that unicast VXLAN tunneling is used between distinct ACI fabrics. These passages demonstrate that the accused products use the newly generated VXLAN tunnel packet to connect tunnel endpoints within one network virtualization domain to tunnel endpoints in another network virtualization domain through a unicast VXLAN tunnel, satisfying Claim 7(e). *See, e.g.*,

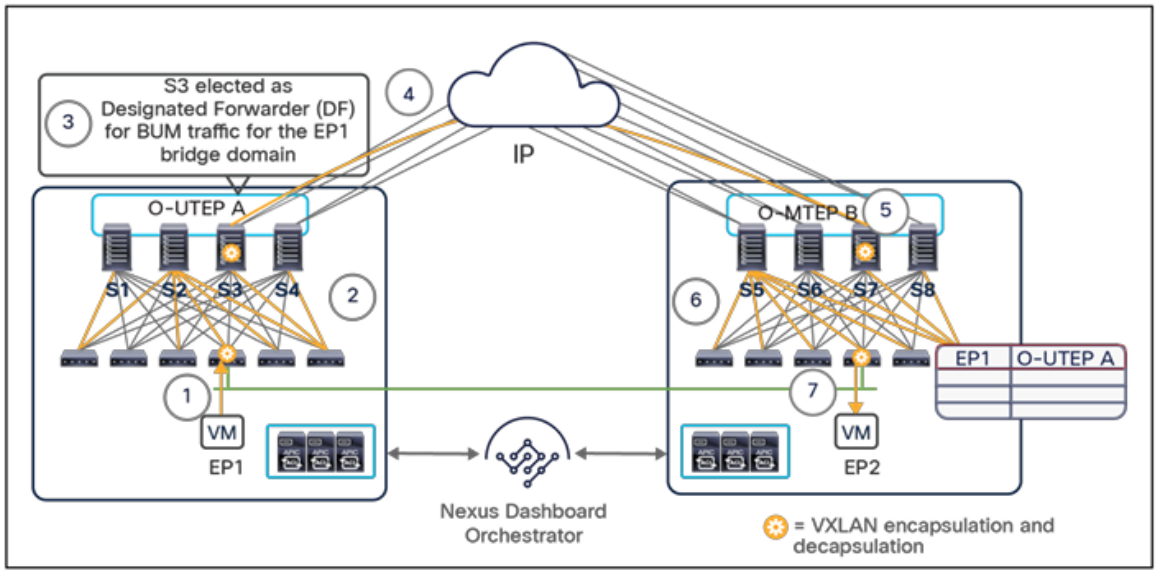
Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
068

- Intersite control plane: Endpoint reachability information is exchanged across sites using a Multiprotocol-BGP (MP-BGP) Ethernet VPN (EVPN) control plane. This approach allows the exchange of MAC and IP address information for the endpoints that communicate across sites. MP-BGP EVPN sessions are established between the spine nodes deployed in separate fabrics that are managed by the same instance of Cisco Nexus Dashboard Orchestrator, as discussed in more detail in the section "Cisco ACI Multi-Site overlay control plane."
- Intersite data plane: All communication (Layer 2 or Layer 3) between endpoints connected to different sites is achieved by establishing site-to-site Virtual Extensible LAN (VXLAN) tunnels across a generic IP network that interconnects the various sites. As discussed in the section "Intersite connectivity deployment considerations", this IP network has no specific functional requirements other than the capability to support routing and increased Maximum Transmission Unit (MTU) size (given the overhead from the VXLAN encapsulation).

The definition of the O-UTEP and O-MTEP ensures that all the resources connected to a specific site are always seen as reachable from the remote sites through one of those two IP addresses (depending on if it is unicast traffic or BUM/multicast communication). This implies that when observing east-west VXLAN traffic exchanged between two sites in the ISN, all the packets will always be sourced from the O-UTEP address identifying the source site and destined to the O-UTEP (or O-MTEP) of the destination site. Notice that this does not prevent taking advantage of multiple ECMP links that may interconnect the two sites: when building the VXLAN encapsulated packet, a hashing of the L2/L3/L4 headers for the original frame generated by the endpoint is used as a UDP source port in the external header, and this allows you to build "entropy" in the packet. Different application flows between the same sites (and even between the same pair of endpoints) would cause the creation of different UDP source port values: as long as the network devices inside the ISN consider the L4 port information to choose the link to forward traffic, it will hence be possible to load-balance VXLAN packets across multiple paths.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
069

Figure 69 shows the sequence of events required to send a Layer 2 BUM frame across sites.



Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
070

Figure 69.
Layer 2 BUM traffic across sites

1. EP1, belonging to a specific bridge domain, generates a Layer 2 BUM frame.
2. Depending on the type of frame and the corresponding bridge domain settings (as previously clarified above), the leaf may need to flood the traffic in that specific Layer 2 domain. As a consequence, the frame is VXLAN-encapsulated and sent to the specific multicast group (called GIPo) associated with the bridge domain within the fabric along one of the specific multi-destination trees associated to that GIPo, so it can reach all the other leaf and spine nodes.
3. One of the spine nodes connected to the external intersite network is elected as the designated forwarder for that specific bridge domain (this election is held between the spine nodes using IS-IS protocol exchanges). The designated forwarder is responsible for replicating each BUM frame for that bridge domain to all the remote sites with the same stretched bridge domain.
4. The designated forwarder makes copies of the BUM frame and sends it to the remote sites. The destination IP address used when VXLAN encapsulating the packet is the special IP address (O-MTEP) identifying each remote site and is used specifically for the transmission of BUM traffic across sites. The O-MTEP is another anycast IP address defined on all the remote spine nodes that are connected to the intersite network (each site uses a unique O-MTEP address). The source IP address for the VXLAN-encapsulated packet is instead the anycast O-UTEP address deployed on all the local spine nodes connected to the intersite network.

   Note: The O-MTEP (referred to as "Overlay Multicast TEP" in the Cisco Nexus Dashboard Orchestrator GUI) is yet another IP address that must be sent to the Layer 3 network connecting the fabrics.
5. One of the remote spine nodes receives the packet, translates the VNID value contained in the header to the locally significant VNID value associated with the same bridge domain, and sends the traffic to the site along one of the local multi-destination trees defined for the bridge domain (using the multicast group locally assigned to the BD as destination in the VXLAN header).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
071

	To support the IP mobility use case, the Cisco ACI Multi-Site solution provides several important functions: • The relocated endpoint must be allowed to communicate with endpoints that are part of the same (or a different) IP subnet that may still be connected in the original site. To address this requirement, when an ARP request is sourced from the original site for the relocated endpoint, the Cisco ACI Multi-Site architecture must deliver that ARP request to the endpoint in the remote site. If the IP address of the migrated endpoint has been discovered by the fabric (which is normally the case for cold migration scenarios), the ARP request can be delivered across sites in unicast mode (by performing VXLAN encapsulation directly to the spines' anycast TEP address identifying the new site in which the destination endpoint is now connected). If the IP address of the migrated endpoint is initially not discovered (which could be the case for hot migration if the migrated endpoint remains “silent” after the move), an “ARP Glean” functionality is performed by the spines in the original site to force the silent host to originate an ARP reply allowing its discovery. At that point, a newly originated ARP request would be delivered in unicast mode as described in the previous paragraph.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
072

Cisco ACI Multi-Site overlay data plane

After endpoint information is exchanged across sites, the VXLAN data plane is used to allow intersite Layer 2 and Layer 3 communication. Before exploring in detail how this communication can be established, you should understand how Layer 2 multi-destination traffic (usually referred to as BUM) is handled across sites.

Layer 2 BUM traffic handling across sites

The deployment of VXLAN allows the use of a logical abstraction so that endpoints separated by multiple Layer 3 hops can communicate as if they were part of the same logical Layer 2 domain. Thus, those endpoints must be capable of sourcing Layer 2 multi-destination frames so that they can be received by all the other endpoints connected to the same Layer 2 segment, regardless of their actual physical location.

This capability can be achieved in one of two ways: by using the native multicast replication functions offered by the Layer 3 infrastructure interconnecting the endpoints or by enabling ingress replication functions on the source VXLAN TEP (VTEP) devices, which create multiple unicast copies of each BUM frame to be sent to all the remote VTEPs on which those endpoints part of the same Layer 2 domain are connected.

The Cisco ACI Multi-Site design adopts this second approach for intersite BUM forwarding, with the Multi-Site-capable spine switches performing the ingress replication function, because the interconnected fabrics may be deployed around the world, and it would be difficult to ensure proper multicast support across the entire interconnecting network infrastructure (the approach adopted instead in the Cisco ACI Multi-Pod architecture).

The transmission of Layer 2 BUM frames across sites is required only for the specific bridge domains that are stretched with flooding enabled (that is, the “Intersite BUM Traffic Allow” flag is configured for the bridge domains). There are three different types of Layer 2 BUM traffic, and below is described the intersite forwarding behavior for each of them when BUM is allowed across sites for a given bridge domain. The assumption here is that specific bridge domain configuration knobs are not modified at the APIC level, but the configuration for the bridge domain is only controlled at the Cisco Nexus Dashboard Orchestrator level.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
073

- Layer 2 Broadcast frames (B): Those are always forwarded across sites. A special type of Layer 2 broadcast traffic is ARP; specific considerations can be found as part of the "Intra-subnet unicast communication across sites" section.
- Layer 2 Unknown Unicast frames (U): Those frames, by default, are not flooded across sites but are instead forwarded in unicast mode, assuming that the destination MAC is known in the COOP database of the local spines (else the traffic will be dropped by the receiving spine). However, there is the possibility of changing this behavior on the bridge-domain-specific configuration of Cisco Nexus Dashboard Orchestrator by selecting the "flood" option associated to the "L2 UNKNOWN UNICAST" traffic.
- Layer 2 Multicast frames (M): The same forwarding behavior applies to intra-bridge-domain Layer 3 multicast frames (that is, the source and receivers are in the same or different IP subnets but part of the same bridge domain) or to "true" Layer 2 multicast frames (that is, the destination MAC address is multicast and there is no IP header in the packet). In both cases, the traffic is forwarded across the sites where the bridge domain is stretched once BUM forwarding is enabled for that bridge domain.

  Note: What is mentioned above applies assuming that multicast routing is not enabled. For more considerations about support of multicast routing with Cisco ACI Multi-Site, please refer to the "Multi-Site Layer 3 multicast" section.

Data-plane considerations

As previously discussed in the "Cisco ACI Multi-Site overlay data plane" section, different types of east-west communication can be established between sites, all leveraging the use of VXLAN data-plane encapsulation to simplify the configuration and functionality required by the intersite network. This remains valid when integrating Cisco ACI Multi-Pod and Multi-Site, with just a few additional design considerations discussed in the following sections.

Layer 2 and Layer 3 unicast communication

Assuming endpoint reachability information has been exchanged across sites, and assuming ARP exchange has also successfully completed, the leaf and spine nodes would have their Layer 2, Layer 3, and COOP tables properly populated to allow the establishment of intersite unicast traffic flows (Figure 94).

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
074

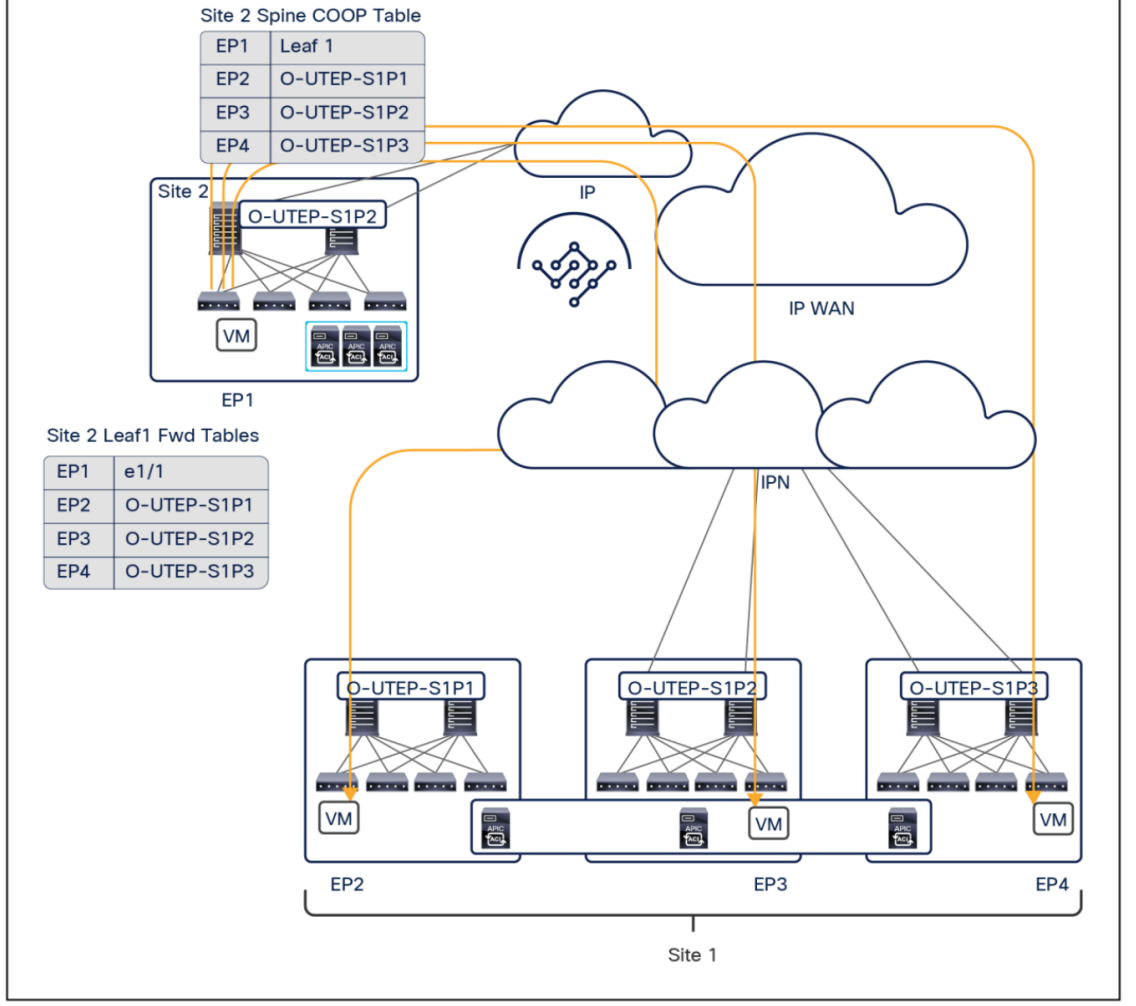


Figure 94.
Unicast traffic ingressing a Multi-Pod fabric that is part of Cisco ACI Multi-Site

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
075

When integrating a Cisco ACI Multi-Pod fabric in a Multi-Site architecture, each pod will be assigned a unique Overlay Unicast TEP address (O-UTEP) to be associated to all the Multi-Site-capable spines deployed in that pod. This ensures that Layer 2 or Layer 3 traffic received from a remote site can be steered directly to the specific pod where the destination endpoint is connected.

Note: The Data-Plane TEP used for inter-Pod communication must be different from the O-UTEP address used for inter-fabric communication.

Figure 95 shows, instead, how Layer 2 BUM traffic can be sent from a remote site toward a Multi-Pod fabric deployed as a site.

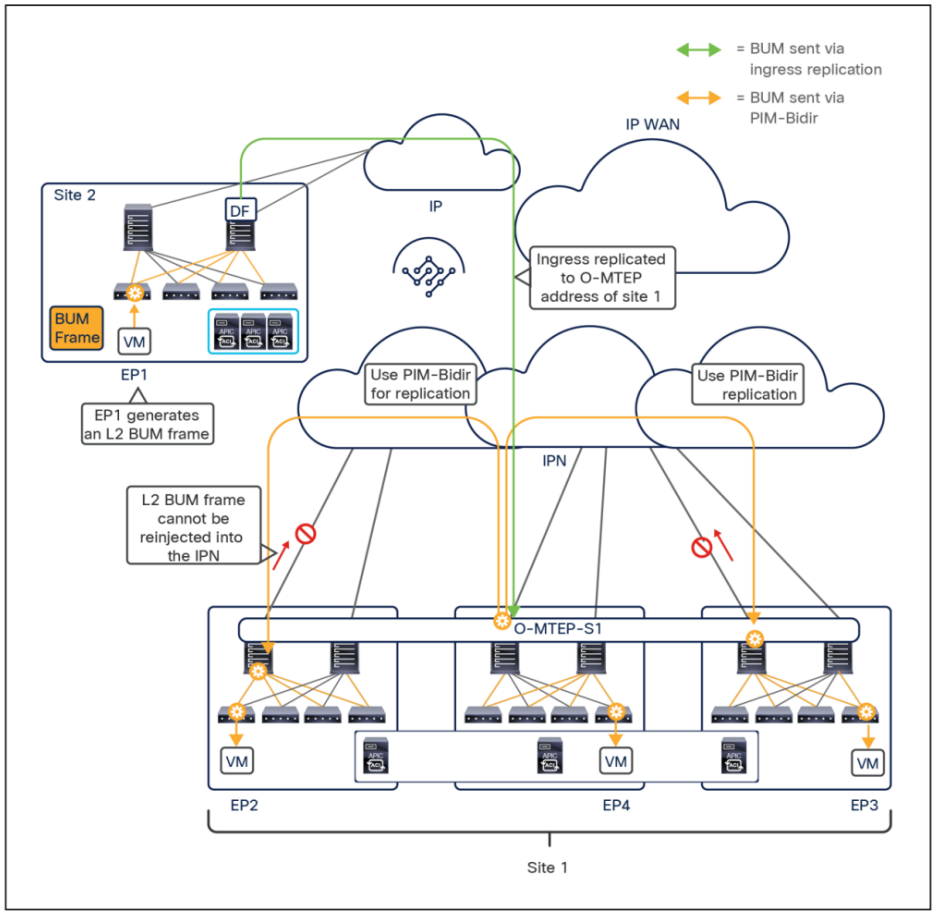


Figure 95.
Layer 2 BUM traffic ingressing a Multi-Pod fabric that is part of Cisco ACI Multi-Site

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
076

	Differently from the unicast scenario, a single Overlay Multicast TEP address is associated to all the Multi-Site-capable spines that are part of the same Multi-Pod fabric (that is, across all the deployed pods). This implies that ingress L2 BUM traffic can be delivered to the spines of any pod, solely based on routing information available in the external backbone network. The receiving spine would then have to perform two operations; it would have to: • Forward the BUM traffic inside the pod along an FTAG tree associated to the bridge domain the traffic belongs to • Forward the BUM traffic to the other pods through the IPN When the spines in other pods receive the BUM traffic, they would then be able to forward it locally inside their pods while not being allowed to reinject the traffic into the IPN, to avoid traffic duplication. https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.pdf; *see also* https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html.
[8] The method according to claim 7, wherein the performing the L2 switching comprises: obtaining L2 SA (source address) information for the processed packet such that the L2 switching may be performed; determining whether or not L2 DA (destination address) information of the processed packet is a unicast address already learned; and performing a flooding based on a type of a input virtual port, when the L2 DA information is not the unicast address already learned, or indicates multicast or broadcast address, wherein the flooding is performed on only domain virtual ports within a domain VSI when the type of the input virtual port being an edge virtual port, and the flooding is performed on only edge virtual ports within	The Accused Instrumentalities perform the method according to claim 7, wherein the performing the L2 switching comprises: obtaining L2 SA (source address) information for the processed packet such that the L2 switching may be performed; determining whether or not L2 DA (destination address) information of the processed packet is a unicast address already learned; and performing a flooding based on a type of a input virtual port, when the L2 DA information is not the unicast address already learned, or indicates multicast or broadcast address, wherein the flooding is performed on only domain virtual ports within a domain VSI when the type of the input virtual port being an edge virtual port, and the flooding is performed on only edge virtual ports within the domain VSI when the type of the input virtual port being a domain virtual port. *See, e.g.,* *See* evidence and analysis above, e.g., Claims 2 and 7.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
077

the domain VSI when the type of the input virtual port being a domain virtual port.	
[9] The method according to claim 8, wherein the determining whether or not the L2 DA (destination address) information of the processed packet is a unicast address already learned further comprises: in response to the L2 DA information being the unicast address already learned, transmitting the processed packet to a destination virtual port corresponding to the unicast address, and obtaining the L3 address of destination tunnel end point of the packet from the destination virtual port.	The Accused Instrumentalities perform the method according to claim 8, wherein the determining whether or not the L2 DA (destination address) information of the processed packet is a unicast address already learned further comprises: in response to the L2 DA information being the unicast address already learned, transmitting the processed packet to a destination virtual port corresponding to the unicast address, and obtaining the L3 address of destination tunnel end point of the packet from the destination virtual port. *See, e.g.,* *See* evidence and analysis above, e.g., Claims 3, 7, and 8.
[10] The method according to claim 7, wherein the receiving the tunnel packet and processing the packet such that the L2 switching may be performed comprises: obtaining a VNID (Virtual Network ID) of the tunnel packet and L3 address information of a sending tunnel end point from the tunnel header of the tunnel packet; processing the tunnel packet such that the L2 switching may be performed; and transmitting the processed packet to the domain VSI corresponding to the VNID through a virtual port associated with the sending tunnel end point.	The Accused Instrumentalities perform the method according to claim 7, wherein the receiving the tunnel packet and processing the packet such that the L2 switching may be performed comprises: obtaining a VNID (Virtual Network ID) of the tunnel packet and L3 address information of a sending tunnel end point from the tunnel header of the tunnel packet; processing the tunnel packet such that the L2 switching may be performed; and transmitting the processed packet to the domain VSI corresponding to the VNID through a virtual port associated with the sending tunnel end point. *See, e.g.,* *See* evidence and analysis above, e.g., Claims 4 and 7.
[11] The method according to claim 7, wherein the adding a tunnel header to the packet for which the L2 switching has been performed to generate a new tunnel packet, and transmitting the new tunnel packet comprises: generating, for the packet for which the L2 switching has been	The Accused Instrumentalities perform the method according to claim 7, wherein the adding a tunnel header to the packet for which the L2 switching has been performed to generate a new tunnel packet, and transmitting the new tunnel packet comprises: generating, for the packet for which the L2 switching has been performed, a VN header (Virtual Network Header) including a VNID (Virtual Network ID) corresponding to a domain VSI, and an L3 tunnel header based on a destination tunnel end point L3 address obtained from a destination virtual port of the

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
078

performed, a VN header (Virtual Network Header) including a VNID (Virtual Network ID) corresponding to a domain VSI, and an L3 tunnel header based on a destination tunnel end point L3 address obtained from a destination virtual port of the packet for which the L2 switching has been performed; and combining the VN header and the L3 tunnel header with the packet for which the L2 switching has been performed, and transmitting the same.	packet for which the L2 switching has been performed; and combining the VN header and the L3 tunnel header with the packet for which the L2 switching has been performed, and transmitting the same. *See, e.g.,* *See* evidence and analysis above, e.g., Claims 5 and 7.
[12] The method according to claim 7, prior to collecting tunnel end points information within a network, and connecting each tunnel end point with a tunnel based on the tunnel end point information collected, further comprising receiving a VNID (Virtual Network ID) of the virtual network from a network virtualization manager that manages network virtualization of an entire network and generating a domain VSI (Domain Virtual Switching Instance) corresponding thereto.	The Accused Instrumentalities perform the method according to claim 7, wherein prior to collecting tunnel end points information within a network, and connecting each tunnel end point with a tunnel based on the tunnel end point information collected, further comprising receiving a VNID (Virtual Network ID) of the virtual network from a network virtualization manager that manages network virtualization of an entire network and generating a domain VSI (Domain Virtual Switching Instance) corresponding thereto.. *See, e.g.,* *See* evidence and analysis above, e.g., Claims 6 and 7.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
079

Since both APIC domains are completely independent from each other, it is logical to expect that different VNID and class-ID values would be assigned to a given EPG (the “real” and the “shadow” copy) across sites. This implies that a translation of those values is required on the spines receiving data-plane traffic from a remote site before injecting the traffic into the local site, as highlighted in Figure 8, below.

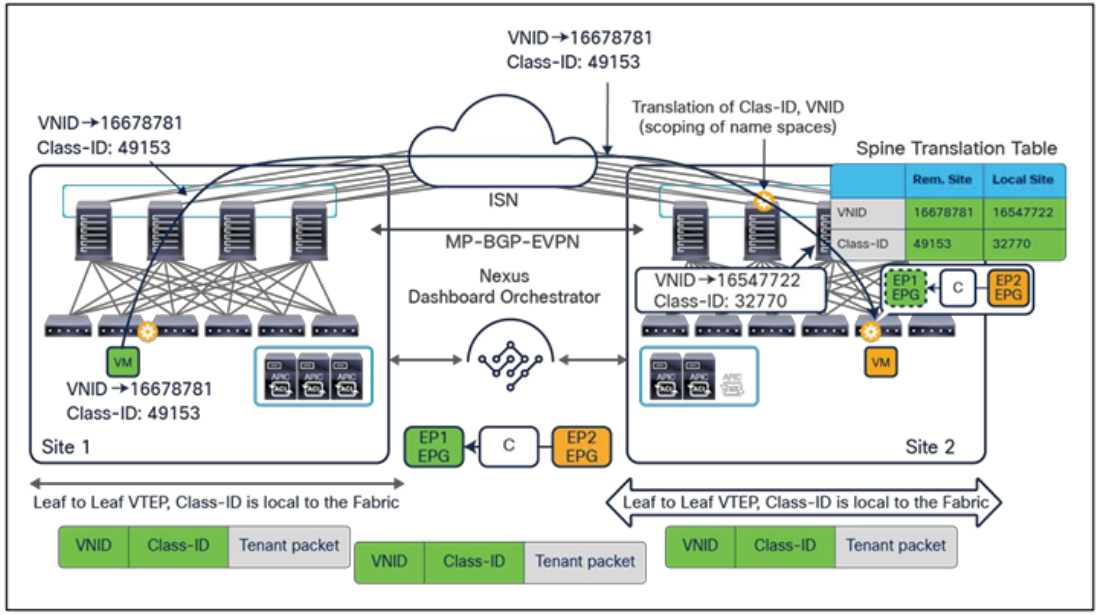


Figure 8.
Name-space translation function on the receiving spine

Layer 2 and Layer 3 unicast communication

Assuming endpoint reachability information has been exchanged across sites, and assuming ARP exchange has also successfully completed, the leaf and spine nodes would have their Layer 2, Layer 3, and COOP tables properly populated to allow the establishment of intersite unicast traffic flows (Figure 94).

As described in the Cisco ACI Multi-Site whitepaper, VNID assignment occurs as part of the initial configuration and policy distribution performed by the Cisco Nexus Dashboard Orchestrator (NDO) and per-site Cisco APIC controllers, before EVPN-based endpoint reachability information is exchanged between sites. The whitepaper explains that the APIC

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
080

	domains are "managed by the same instance of Cisco Nexus Dashboard Orchestrator" and that this orchestration ensures that "the VNIDs are consistent across sites," which must occur in advance of intersite discovery so the fabrics agree on a common virtual-network identifier for each Bridge Domain (BD). The whitepaper further states that "the VXLAN Network Identifier (VNID) identifies the bridge domain (BD) (for Layer 2 communication)," establishing that VNID receipt and BD instantiation precede any data-plane learning or EVPN-based TEP reachability exchange. Together, these statements show that the system receives the VNID from the network virtualization manager before it collects tunnel endpoint information via EVPN/COOP and generates the corresponding BD, satisfying the temporal requirement of Claim 12.
[13 pre] A network virtualization apparatus comprising:	To the extent that the preamble is limiting, the Accused Instrumentalities include a network virtualization apparatus. *See, e.g.,* *See* evidence and analysis above, e.g., Claim 1[pre].
[13a] a tunnel manager configured to collect tunnel end point information within a network, and to connect each tunnel end point with a tunnel based on the tunnel end point information collected;	The Accused Instrumentalities include a network virtualization apparatus which includes a tunnel manager configured to collect tunnel end point information within a network, and to connect each tunnel end point with a tunnel based on the tunnel end point information collected. *See, e.g.,* *See* evidence and analysis above, e.g., Claim 1[a].
[13b] a tunnel packet end point configured to receive a tunnel packet and process the packet such that an L2 switching may be performed, and to transmit the packet to a domain VSI (virtual switching instance);	The Accused Instrumentalities include a network virtualization apparatus which includes a tunnel packet end point configured to receive a tunnel packet and process the packet such that an L2 switching may be performed, and to transmit the packet to a domain VSI (virtual switching instance). *See, e.g.,* *See* evidence and analysis above, e.g., Claim 1[b].

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
081

[13c] the domain VSI configured to perform the L2 switching on the processed packet, and to transmit the packet to a tunnel packet generator; and	The Accused Instrumentalities include a network virtualization apparatus which includes the domain VSI configured to perform the L2 switching on the processed packet, and to transmit the packet to a tunnel packet generator. *See, e.g.,* *See* evidence and analysis above, e.g., Claim 1[c].
[13d] the tunnel packet generator configured to add a tunnel header to the packet for which the L2 switching has been performed to generate a new unicast tunnel packet, and to transmit the new unicast tunnel packet.	The Accused Instrumentalities include a network virtualization apparatus which includes the tunnel packet generator configured to add a tunnel header to the packet for which the L2 switching has been performed to generate a new unicast tunnel packet, and to transmit the new unicast tunnel packet. *See, e.g.,* *See* evidence and analysis above, e.g., Claim limitations 1[d] and 7(e). Specifically, the evidence cited for Claim 1[d], together with the unicast-specific VXLAN encapsulation evidence already cited under Claim 7(e)—including the O-UTEP definition, the unicast scenario text, Figure 71, Figure 94, and Figure 95—demonstrates that the accused products generate and transmit VXLAN packets using unicast tunnel endpoints (O-UTEP) for intersite unicast communication. Accordingly, the evidence for Claim 1 and Claim 7(e) together fully satisfies all limitations of Claim 13.

Cisco Systems, Inc. v. OptimNet LLC
Exhibit 1016
082