Virtual Machine Manager Domains

This chapter contains the following sections:

Cisco ACI VM Networking Support for Virtual Machine Managers

Benefits of ACI VM Networking

Cisco Application Centric Infrastructure(ACI) virtual machine (VM) networking supports hypervisors from multiple vendors. It provides the hypervisors programmable and automated access to high-performance scalable virtualized data center infrastructure.

Programmability and automation are critical features of scalable data center virtualization infrastructure. The Cisco ACI open REST API enables virtual machine integration with and orchestration of the policy model-based Cisco ACI fabric. Cisco ACI VM networking enables consistent enforcement of policies across both virtual and physical workloads that are managed by hypervisors from multiple vendors.

Attachable entity profiles easily enable VM mobility and placement of workloads anywhere in the Cisco ACI fabric. The Cisco Application Policy Infrastructure Controller (APIC) provides centralized troubleshooting, application health score, and virtualization monitoring. Cisco ACI multi-hypervisor VM automation reduces or eliminates manual configuration and manual errors. This enables virtualized data centers to support large numbers of VMs reliably and cost effectively.

Supported Products and Vendors

Cisco ACI supports virtual machine managers (VMMs) from the following products and vendors:

- **Cisco Unified Computing System Manager (UCSM)**

  Integration of Cisco UCSM is supported beginning in Cisco Cisco APIC Release 4.1(1). For information, see the chapter "Cisco ACI with Cisco UCSM Integration in the *Cisco ACI Virtualization Guide, Release 4.1(1)* .

- **Cisco Application Centric Infrastructure (ACI) Virtual Pod (vPod)**

  Cisco ACI vPod is in general availability beginning in Cisco APIC Release 4.0(2). For information, see the Cisco ACI vPod documentation on Cisco.com.

- **Cisco Application Centric Infrastructure Virtual Edge**

  For information, see the Cisco ACI Virtual Edge documentation on Cisco.com.

- **Cisco Application Virtual Switch (AVS)**

  For information, see the chapter "Cisco ACI with Cisco AVS" in the Cisco ACI Virtualization Guide and Cisco AVS documentation on Cisco.com.

- **Cloud Foundry**

  Cloud Foundry integration with Cisco ACI is supported beginning with Cisco APIC Release 3.1(2). For information, see the knowledge base article, Cisco ACI and Cloud Found Integration on Cisco.com.

- **Pivotal Cloud Foundry**

  Pivotal Cloud Foundry integration with Cisco ACI is supported beginning with Cisco APIC Release 3.2(1). For information, see the knowledge base article, Cisco ACI and Cloud Foundry Integration on Cisco.com.

- **Kubernetes**

  For information, see the knowledge base article, Cisco ACI and Kubernetes Integration on Cisco.com.

- **Microsoft System Center Virtual Machine Manager (SCVMM)**

  For information, see the chapters "Cisco ACI with Microsoft SCVMM" and "Cisco ACI with Microsoft Windows Azure Pack" in the *Cisco ACI Virtualization Guide* on Cisco.com

- **OpenShift**

  For information, see the OpenShift documentation on Cisco.com.

- **OpenStack**

  For information, see the OpenStack documentation on Cisco.com.

- **Red Hat Virtualization (RHV)**

  For information, see the knowledge base article, Cisco ACI and Red Hat Integration on Cisco.com.

- **VMware Virtual Distributed Switch (VDS)**

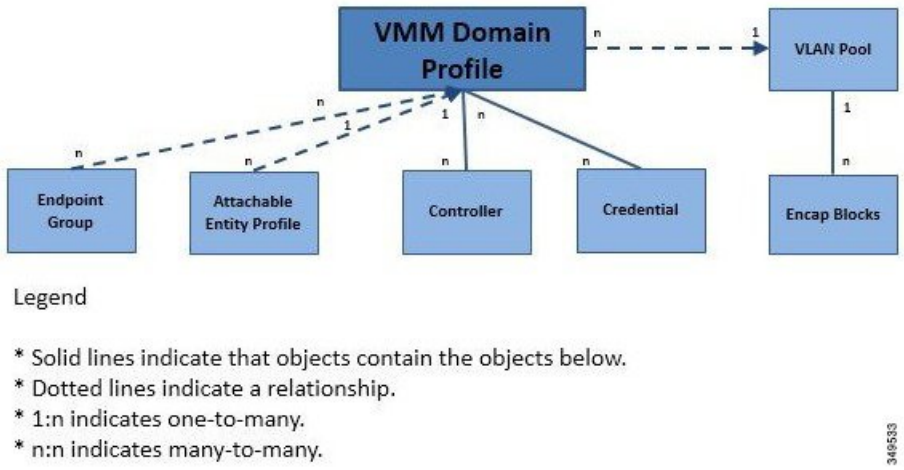
  For information, see the chapter "Cisco "ACI with VMware VDS Integration" in the Cisco ACI Virtualization Guide.

See the Cisco ACI Virtualization Compatibility Matrix for the most current list of verified interoperable products.

VMM Domain Policy Model

VMM domain profiles (`vmmDomP`) specify connectivity policies that enable virtual machine controllers to connect to the ACI fabric. The figure below provides an overview of the `vmmDomP` policy.

Figure 1: VMM Domain Policy Model Overview



Virtual Machine Manager Domain Main Components

ACI fabric virtual machine manager (VMM) domains enable an administrator to configure connectivity policies for virtual machine controllers. The essential components of an ACI VMM domain policy include the following:

- **Virtual Machine Manager Domain Profile**—Groups VM controllers with similar networking policy requirements. For example, VM controllers can share VLAN pools and application endpoint groups (EPGs). The APIC communicates with the controller to publish network configurations such as port groups that are then applied to the virtual workloads. The VMM domain profile includes the following essential components:
  - **Credential**—Associates a valid VM controller user credential with an APIC VMM domain.
  - **Controller**—Specifes how to connect to a VM controller that is part of a policy enforcement domain. For example, the controller specifies the connection to a VMware vCenter that is part a VMM domain.

Note A single VMM domain can contain multiple instances of VM controllers, but they must be from the same vendor (for example, from VMware or from Microsoft.

- **EPG Association**—Endpoint groups regulate connectivity and visibility among the endpoints within the scope of the VMM domain policy. VMM domain EPGs behave as follows:
  - The APIC pushes these EPGs as port groups into the VM controller.

    - An EPG can span multiple VMM domains, and a VMM domain can contain multiple EPGs.

- **Attachable Entity Profile Association**—Associates a VMM domain with the physical network infrastructure. An attachable entity profile (AEP) is a network interface template that enables deploying VM controller policies on a large set of leaf switch ports. An AEP specifies which switches and ports are available, and how they are configured.
- **VLAN Pool Association**—A VLAN pool specifies the VLAN IDs or ranges used for VLAN encapsulation that the VMM domain consumes.

Virtual Machine Manager Domains

An APIC VMM domain profile is a policy that defines a VMM domain. The VMM domain policy is created in APIC and pushed into the leaf switches.

VMM domains provide the following:

- A common layer in the ACI fabric that enables scalable fault-tolerant support for multiple VM controller platforms.
- VMM support for multiple tenants within the ACI fabric.

VMM domains contain VM controllers such as VMware vCenter or Microsoft SCVMM Manager and the credential(s) required for the ACI API to interact with the VM controller. A VMM domain enables VM mobility within the domain but not across domains. A single VMM domain can contain multiple instances of VM controllers but they must be the same kind. For example, a VMM domain can contain many VMware vCenters managing multiple controllers each running multiple VMs but it may not also contain SCVMM Managers. A VMM domain inventories controller elements (such as pNICs, vNICs, VM names, and so forth) and pushes policies into the controller(s), creating port groups, and other necessary elements. The ACI VMM domain listens for controller events such as VM mobility and responds accordingly.

VMM Domain VLAN Pool Association

VLAN pools represent blocks of traffic VLAN identifiers. A VLAN pool is a shared resource and can be consumed by multiple domains such as VMM domains and Layer 4 to Layer 7 services.

Each pool has an allocation type (static or dynamic), defined at the time of its creation. The allocation type determines whether the identifiers contained in it will be used for automatic assignment by the Cisco APIC (dynamic) or set explicitly by the administrator (static). By default, all blocks contained within a VLAN pool have the same allocation type as the pool but users can change the allocation type for encapsulation blocks contained in dynamic pools to static. Doing so excludes them from dynamic allocation.

A VMM domain can associate with only one dynamic VLAN pool. By default, the assignment of VLAN identifiers to EPGs that are associated with VMM domains is done dynamically by the Cisco APIC. While dynamic allocation is the default and preferred configuration, an administrator can statically assign a VLAN identifier to an endpoint group (EPG) instead. In that case, the identifiers used must be selected from encapsulation blocks in the VLAN pool associated with the VMM domain, and their allocation type must be changed to static.

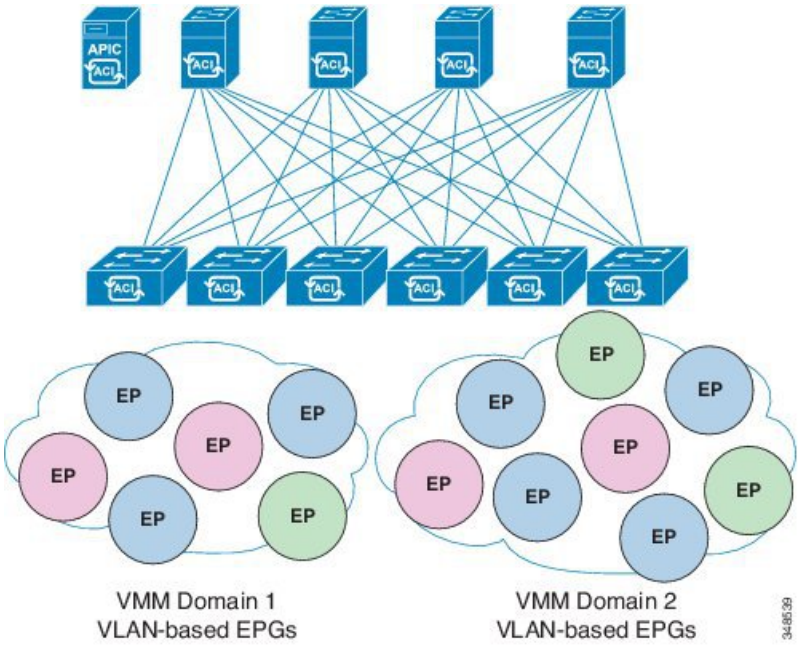
The Cisco APIC provisions VMM domain VLAN on leaf ports based on EPG events, either statically binding on leaf ports or based on VM events from controllers such as VMware vCenter or Microsoft SCVMM.

Note In dynamic VLAN pools, if a VLAN is disassociated from an EPG, it is automatically reassociated with the EPG in five minutes.

VMM Domain EPG Association

The Cisco Application Centric Infrastructure (ACI) fabric associates tenant application profile endpoint groups (EPGs) to virtual machine manager (VMM) domains, The Cisco ACI does so either automatically by an orchestration component such as Microsoft Azure, or by a Cisco Application Policy Infrastructure Controller (APIC) administrator creating such configurations. An EPG can span multiple VMM domains, and a VMM domain can contain multiple EPGs.

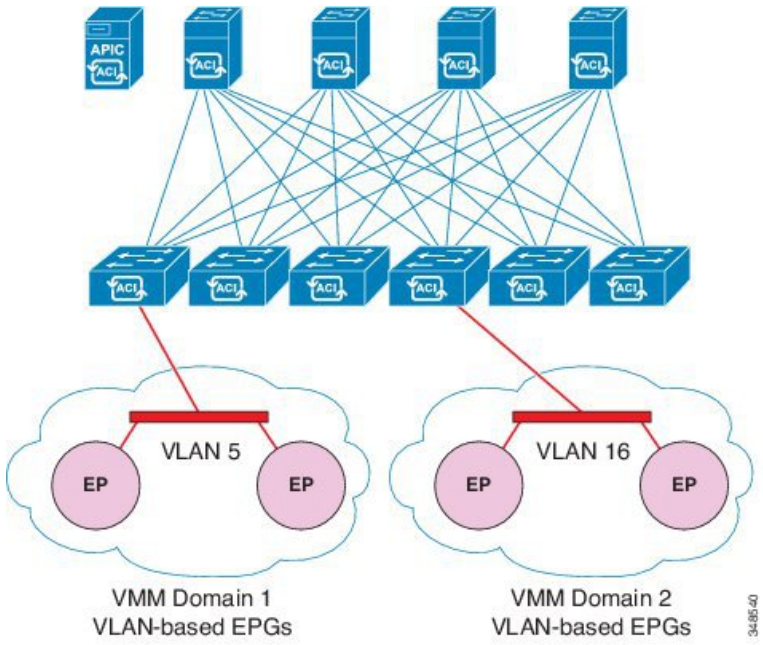
Figure 2: VMM Domain EPG Association



In the preceding illustration, end points (EPs) of the same color are part of the same EPG. For example, all the green EPs are in the same EPG although they are in two different VMM domains.

See the latest *Verified Scalability Guide for Cisco ACI* for virtual network and VMM domain EPG capacity information.

Figure 3: VMM Domain EPG VLAN Consumption



Note When multiple VMM domains with an overlapping VLAN ID range are connected to the same leaf switch, those domains should use the same VLAN pool. With the same VLAN pool, Cisco APIC can make sure to pick a different VLAN ID for each domain-to-EPG association. Otherwise, Cisco APIC might pick a VLAN ID that is already used on the switch for another domain-to-EPG association, which causes the VLAN deployment fail.

When multiple VMM domains with an overlapping VLAN ID range are connected to the same leaf switch and those domains use the same VLAN pool, you can have multiple VMM domains associated with the same EPG. However, each domain-to-EPG association deploys a different VLAN ID, respectively, even though the VLANs are for the same EPG and potentially are on the same port. If using VLAN IDs in this manner is suboptimal to your requirements, you can use the same VMM domain with multiple VMM controllers instead of having multiple VMM domains.

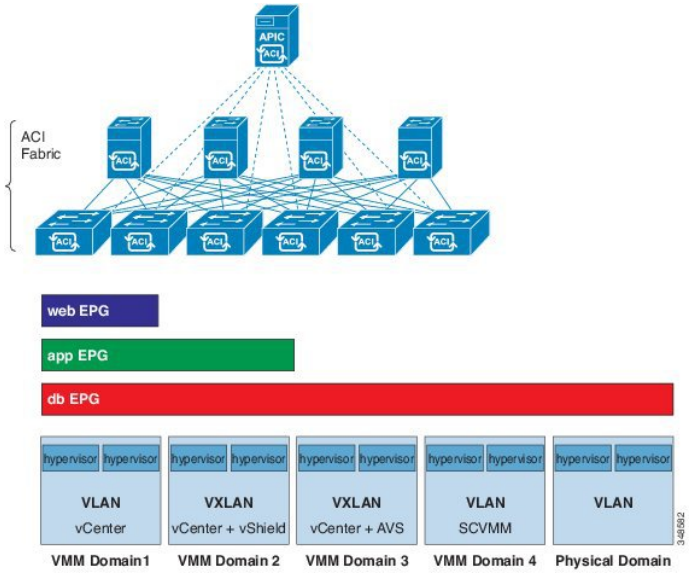
EPGs can use multiple VMM domains in the following ways:

- An EPG within a VMM domain is identified by using an encapsulation identifier. Cisco APIC can manage the identifier automatically, or the administrator can statically select it. An example is a VLAN, a Virtual Network ID (VNID).
- An EPG can be mapped to multiple physical (for baremetal servers) or virtual domains. It can use different VLAN or VNID encapsulations in each domain.

Note By default, the Cisco APIC dynamically manages the allocatiion of a VLAN for an EPG. VMware DVS administrators have the option to configure a specific VLAN for an EPG. In that case, the VLAN is chosen from a static allocation block within the pool that is associated with the VMM domain.

Applications can be deployed across VMM domains.

Figure 4: Multiple VMM Domains and Scaling of EPGs in the Fabric



While live migration of VMs within a VMM domain is supported, live migration of VMs across VMM domains is not supported.

Note When you change the VRF on a bridge domain that is linked to an EPG with an associated VMM domain, the port-group is deleted and then added back on vCenter. This results in the EPG being undeployed from the VMM domain. This is expected behavior.

Trunk Port Group

A trunk port group is used to aggregate the traffic of EPGs. Currently, it is supported under a VMware domain only. The trunk port group's naming scheme does not follow an EPG's T|A|E format. The name can be any ASCII string, as a trunk port group is not tenant-aware.

The aggregation of EPGs under the same domain is based on a VLAN range, which is specified as encapsulation blocks contained in the trunk port group. Whenever a EPG's encapsulation is changed or a trunk port group's encapsulation block is changed, the aggregation will be re-evaluated to determine if the EGP should be aggregated. A trunk port group controls the deployment in leafs of network resources, such as VLANs, allocated to EPGs being aggregated, including both the base EPG and uSeg EPG. In the case of a uSeg EPG, the trunk port group's VLAN ranges need to include both the primary and secondary VLANs.

Note Cisco ACI does not support IP fragmentation. Therefore, when you configure Layer 3 Outside (L3Out) connections to external routers, or multipod connections through an Inter-Pod Network (IPN), it is critical that the MTU is set appropriately on both sides. On some platforms, such as ACI, Cisco NX-OS, and Cisco IOS, the configurable MTU value takes into account the IP headers (resulting in a max packet size to be set as 9216 bytes for ACI and 9000 for NX-OS and IOS). However, other platforms such as IOS-XR configure the MTU value exclusive of packet headers (resulting in a max packet size of 8986 bytes).

For the appropriate MTU values for each platform, see the relevant configuration guides.

Cisco highly recommends you test the MTU using CLI-based commands. For example, on the Cisco NX-OS CLI, use a command such as `ping 1.1.1.1 df-bit packet-size 9000 source-interface ethernet 1/1`.

Caution If you install 1 Gigabit Ethernet (GE) or 10GE links between the leaf and spine switches in the fabric, there is risk of packets being dropped instead of forwarded, because of inadequate bandwidth. To avoid the risk, use 40GE or 100GE links between the leaf and spine switches.

Note Multiple Spanning Tree (MST) is not supported on interfaces configured with the Per Port VLAN feature (configuring multiple EPGs on a leaf switch using the same VLAN ID with localPort scope).

Note If you are using Cisco ACI Multi-Site with this Cisco APIC cluster/fabric, look for a cloud icon on the object names in the navigation bar. This indicates that the information is derived from Multi-Site. It is recommended to only make changes from the Multi-Site GUI. Please review the Multi-Site documentation before making changes here.

Note For a Cisco APIC REST API query of event records, the APIC system limits the response to a maximum of 500,000 event records. If the response is more than 500,000 events, it returns an error. Use filters to refine your queries. For more information, see Composing Query Filter Expressions.

See the *Cisco APIC Layer 4 to Layer 7 Services Deployment Guide* or the *Cisco ACI Virtualization Guide* for more information.

EPG Policy Resolution and Deployment Immediacy

Whenever an EPG associates to a VMM domain, the administrator can choose the resolution and deployment preferences to specify when a policy should be pushed into leaf switches.

Resolution Immediacy

- Pre-provision—Specifies that a policy (for example, VLAN, VXLAN binding, contracts, or filters) is downloaded to a leaf switch even before a VM controller is attached to the virtual switch (for example, VMware VDS). This pre-provisions the configuration on the switch.

  This helps the situation where management traffic for hypervisors/VM controllers are also using the virtual switch associated to APIC VMM domain (VMM switch).

  Deploying a VMM policy such as VLAN on ACI leaf switch requires APIC to collect CDP/LLDP information from both hypervisors via VM controller and ACI leaf switch. However if VM Controller is supposed to use the same VMM policy (VMM switch) to communicate with its hypervisors or even APIC, the CDP/LLDP information for hypervisors can never be collected because the policy required for VM controller/hypervisor management traffic is not deployed yet.

  When using pre-provision immediacy, policy is downloaded to ACI leaf switch regardless of CDP/LLDP neighborship. Even without a hypervisor host connected to the VMM switch.

- Immediate—Specifies that EPG policies (including contracts and filters) are downloaded to the associated leaf switch software upon ESXi host attachment to a DVS. LLDP or OpFlex permissions are used to resolve the VM controller to leaf node attachments.

  The policy will be downloaded to leaf when you add host to the VMM switch. CDP/LLDP neighborship from host to leaf is required.

- On Demand—Specifies that a policy (for example, VLAN, VXLAN bindings, contracts, or filters) is pushed to the leaf node only when an ESXi host is attached to a DVS and a VM is placed in the port group (EPG).

  The policy will be downloaded to leaf when host is added to VMM switch and virtual machine needs to be placed into port group (EPG). CDP/LLDP neighborship from host to leaf is required.

  With both immediate and on demand, if host and leaf lose LLDP/CDP neighborship the policies are removed.

Deployment Immediacy

Once the policies are downloaded to the leaf software, deployment immediacy can specify when the policy is pushed into the hardware policy content-addressable memory (CAM).

- Immediate—Specifies that the policy is programmed in the hardware policy CAM as soon as the policy is downloaded in the leaf software.
- On demand—Specifies that the policy is programmed in the hardware policy CAM only when the first packet is received through the data path. This process helps to optimize the hardware space.

Note When you use on demand deployment immediacy with MAC-pinned VPCs, the EPG contracts are not pushed to the leaf ternary content-addressble memory (TCAM) until the first endpoint is learned in the EPG on each leaf. This can cause uneven TCAM utilization across VPC peers. (Normally, the contract would be pushed to both peers.)

Guidelines for Deleting VMM Domains

Follow the sequence below to assure that the APIC request to delete a VMM domain automatically triggers the associated VM controller (for example VMware vCenter or Microsoft SCVMM) to complete the process normally, and that no orphan EPGs are stranded in the ACI fabric.

1. The VM administrator must detach all the VMs from the port groups (in the case of VMware vCenter) or VM networks (in the case of SCVMM), created by the APIC.

   In the case of Cisco AVS, the VM admin also needs to delete vmk interfaces associated with the Cisco AVS.

2. The ACI administrator deletes the VMM domain in the APIC. The APIC triggers deletion of VMware VDS or Cisco AVS or SCVMM logical switch and associated objects.

Note The VM administrator should not delete the virtual switch or associated objects (such as port groups or VM networks); allow the APIC to trigger the virtual switch deletion upon completion of step 2 above. EPGs could be orphaned in the APIC if the VM administrator deletes the virtual switch from the VM controller before the VMM domain is deleted in the APIC.

If this sequence is not followed, the VM controller does delete the virtual switch associated with the APIC VMM domain. In this scenario, the VM administrator must manually remove the VM and vtep associations from the VM controller, then delete the virtual switch(es) previously associated with the APIC VMM domain.