ıılııılıı cısco

Microsoft Azure Stack HCI Connectivity to Cisco Nexus 9000 Series Switches in Cisco NX-OS and Cisco® Application Centric Infrastructure (Cisco ACI™) Mode

Contents

Introduction	5
Prerequisites	5
Terminologies	5
Executive Summary	6
Document Purpose	7
Technology Overview	7
About Cisco ACI	7
Cisco ACI Architecture	8
Cisco Nexus 9000 NX-OS based Fabric	9
Solution Design	
Physical Architecture	
Logical Architecture	
Cisco Nexus 9000 Series Switch based Fabric and Benefit	
Cisco ACI Design for Azure Stack HCI Connectivity	
Cisco ACI for Azure Stack HCI Connectivity	17
Azure Stack HCI ACI Tenant Model Overview	19
Cisco NX-OS based Fabric Design for Azure Stack HCI Connectivity	
Cisco NX-OS based Fabric for Azure Stack HCI Connectivity	20
Solution Deployment	
Cisco ACI Configuration for Azure Stack HCI	21
Configuring Leaf Interfaces Connected to Azure Stack HCI Servers	21
Configure QoS	
Configure EPGs	40
Cisco NX-OS based Fabric configuration for Azure Stack HCI	
Configure QoS	47
Configure LLDP	
Configure Networks for Azure Stack HCI	52
Build External Connectivity for Azure Stack HCI servers	58
Appendix	
Design Example with Microsoft Software Defined Networking (SDN) in Azur	e Stack HCI59
Microsoft Azure SDN Components	59
Logical Architecture	60
Cisco ACI Design for PA Network and SLB Connectivity	61
Cisco ACI Design for Azure Stack HCI VNET Connectivity	62
Logical Network	Error! Bookmark not defined.
Azure Stack HCI VNET Connectivity	Error! Bookmark not defined.
Solution Deployment	

Cisco ACI Configuration for PA Network and SLB Connectivity	70
Cisco ACI Configuration for Azure Stack HCI VNET & Gateway VM Connectivity	
Cisco NX-OS based Fabric Design	or! Bookmark not defined.
For more information	
Revision history	

Note: This document contains material and data with multiple dependencies. The information may be updated as and when necessary and is subject to change without notice.

Privileged/Confidential information is contained in this document and may be subject to legal privilege. Access to this material by anyone other than those intended is unauthorized. If you are not the intended recipient (or responsible for delivery of the information to such person), you may not use, copy, distribute, or deliver to anyone this information (or any part of its contents) or take any action in reliance on it. In such case, you should destroy this information and notify Cisco immediately. If you have received this material in error, please notify us immediately and delete the material from any computer. If you or your employer does not consent to this message, please notify us immediately. Our company cannot accept responsibility for any loss or damage arising from the use of this material.

Introduction

This document describes the network design considerations for Microsoft Azure Stack Hyperconverged Infrastructure (HCI) in a Cisco Nexus 9000 Series Switches-based network with Cisco NX-OS and Cisco[®] Application Centric Infrastructure (Cisco ACI[™]).

Prerequisites

This document assumes that you have a basic knowledge of Cisco ACI and Cisco NX-OS VXLAN technologies.

For more information on Cisco ACI, refer to the white papers on Cisco.com: <u>https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-listing.html</u>

For more information on Cisco NX-OS based VXLAN fabrics, refer to the white papers on Cisco.com: https://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/white-paper-listing.html

Terminologies

- Cisco ACI related terminologies

 BD: bridge domain
 EPG: endpoint group
 L3Out: Layer 3 Out or external routed network
 L3Out EPG: subnet-based EPG in L3Out
 VRF: Virtual Routing and Forwarding
 Border leaf: ACI leaf where L3Out is deployed
- Cisco NX-OS related terminologies
 NDFC: Nexus Dashboard Fabric Controller

VXLAN: Virtual Extensible LAN

VNI: Virtual Network Identifier (one to one co-relation between VLAN to VNI)

DAG: Distributed Anycast Gateway

Leaf: Performs VXLAN encapsulation and decapsulation function also referred as Virtual Tunnel End-Point (VTEP). End-hosts are connected to Leafs in the VXLAN fabric

Spine: Provides Underlay Layer-3 connectivity between the leafs in the VXLAN fabric

Border Leaf: Performs similar function to a Leaf. In addition, Border leafs connect the VXLAN fabric to external networks and are placed at the edge of the VXLAN fabric

External Connectivity: Provide L3 connectivity outside of the VXLAN fabric

- Microsoft Azure Stack HCI related terminologies
 - RDMA: Remote Direct Memory Access
 - RoCE: RDMA over Converged Ethernet
 - SET: Switch Embedded Teaming
 - SMB: Server Message Block

Storage Spaces Direct: A feature of the Microsoft Azure Stack HCI and Windows Server that enables you to cluster servers with an internal storage into a software-defined storage solution. Storage Spaces Direct uses SMB3, including SMB Direct and SMB Multichannel over Ethernet to communicate between servers

SMB Direct: The Windows Server includes a feature called SMB Direct, which supports the use of network adapters that have RDMA capability. Network adapters with RDMA capability can function at full speed with lower latency without compromising CPU utilization. SMB Direct requires network adapters with RDMA capability on the servers and RDMA over Converged Ethernet (RoCEv2) on the network

Executive Summary

Beginning with Cisco ACI Release 6.0(3e) and NX-OS 10.3(2)F, Nexus 9000 Series Switches support the Microsoft <u>Azure Stack HCI requirements</u>. This document details the Microsoft Azure Stack HCI network design with Cisco Nexus 9000 Series Switches in either Cisco ACI or Cisco NX-OS mode.



Figure 1.

Topology example with Nexus 9000 Series Switches in Cisco ACI mode

Note: Cisco Application Policy Infrastructure Controller (APIC) can be connected to leaf switches directly or connected through the Layer 3 network via spine switches.



Figure 2.

Topology example with Nexus 9000 Series Switches in Cisco NX-OS mode

Document Purpose

You must ensure that there are direct connections from the Microsoft Azure Stack HCI servers to the Cisco Nexus 9000 Top-of-Rack (ToR) switches and accessibility to the data center among other required tasks, when installing the Microsoft Azure Stack HCI.

This document provides information, education, and guidance for connecting the Microsoft Azure Stack HCI servers to an existing Cisco Nexus 9000 Series Switch-based network in the data centers. The document provides fundamental information and recommended configurations based on internal testing of the solution. This document does not cover the installation and configuration of Cisco ACI or NX-OS based infrastructure nor does it detail the setup of Microsoft Azure Stack HCI.

This document uses Cisco UCS C240 M6/M7 servers as the Microsoft Azure Stack HCl servers. For Cisco UCS configuration and design considerations, refer to the Cisco Validated Design (CVD) on cisco.com: https://www.cisco.com/c/en/us/td/docs/unified computing/ucs/UCS CVDs/ucs mas hci m7.html.

The Microsoft Azure Stack HCI internal networks are not managed using a Cisco controller such as Cisco APIC and NDFC in this solution. The Azure Stack HCI system is connected to the Nexus 9000 Series Switch-based network, which acts as the gateway to allow the Azure Stack HCI Virtual Machines (VM) to connect with other VMs, the external network, and other internal network services in the datacenter.

Technology Overview

This section introduces the technologies that are used in the solution, which are described in this document.

About Cisco ACI

Cisco ACI is an evolutionary leap from SDN's initial vision of operational efficiency through network agility and programmability. Cisco ACI has industry leading innovations in management automation, programmatic policies, and dynamic workload provisioning. The ACI fabric accomplishes this with a combination of hardware, policy-based control systems, and closely coupled software to provide advantages that are not possible in other architectures. Cisco ACI takes a policy-based systems approach to operationalizing the data center network. The policy is centered around the needs (reachability, access to services, and security policies) of the applications. Cisco ACI delivers a resilient fabric to satisfy today's dynamic applications.

Cisco ACI Architecture

The Cisco ACI fabric is a leaf-and-spine architecture where each leaf connects to every spine using highspeed 40/100/400-Gbps Ethernet links, with no direct connection between the spine switches or leaf switches. The ACI fabric is a routed fabric with a VXLAN overlay network, where every leaf is VXLAN Tunnel Endpoint (VTEP). Cisco ACI provides both Layer 2 (L2) and Layer 3 (L3) forwarding across this routed fabric infrastructure.

The following are the ACI fabric components:

Cisco APIC: Cisco Application Policy Infrastructure Controller (APIC) is the unifying point of automation and management for the Cisco ACI fabric. Cisco APIC is a centralized, clustered controller that provides centralized access to all fabric information, optimizes the application lifecycle for scale and performance, and supports flexible application provisioning across physical and virtual resources. Cisco APIC exposes northbound APIs through XML and JSON and provides both a command-line interface (CLI) and a GUI, which utilize the APIs to manage the fabric.

Leaf Switches: The ACI leaf provides physical connectivity for servers, storage devices, and other access layer components, and enforces the ACI policies. Leaf switches also provide connectivity to an existing enterprise or a service provider infrastructure. The leaf switches provide options starting at 1G up through 400G Ethernet ports for connectivity.

Spine Switches: In ACI, spine switches provide the mapping database function and connectivity between leaf switches. A spine switch can be the modular Cisco Nexus 9500 series equipped with ACI ready line cards or a fixed form-factor switch, such as the Cisco Nexus 9332D-GX2B. Spine switches provide high-density 40/100/400 Gigabit Ethernet connectivity to the leaf switches.



Figure 3. Cisco ACI Fabric Components

Cisco Nexus 9000 NX-OS based Fabric

Cisco NX-OS based fabric is another option for building a data center by using the Nexus 9000 series switches. These switches act as independent devices and have their own control-plane and data-plane. Nexus 9000 series switches running NX-OS offer various data Center fabric options, such as VXLAN, L3 Routed or traditional (2-tier or 3-tier) LAN.

This document only focuses on connecting the Azure Stack HCI to the VXLAN fabric. However, NX-OS based L3 Routed or traditional LAN fabrics can also be used.

The following are the Cisco NX-OS based VXLAN fabric components:

NDFC: Cisco Nexus Dashboard Fabric Controller (NDFC) is an Orchestration and Automation tool to build and manage data center fabrics. Cisco NDFC can be used either in LAN or SAN mode. In LAN mode, NDFC supports various fabric templates to create VXLAN, VXLAN Multisite, L3 Routed Fabric, and traditional LAN and IP Fabrics for media. Cisco NDFC offers the following day 0 to day 2 operations:

- Day 0: Bootstrap (POAP) support for the devices and pre-provisioning of the fabrics.
- Day 1: Automation of new Greenfield fabrics as well as support for Brownfield fabrics, deployment for Networks & VRFs, and L4-L7 service insertion.
- Day 2: Image Management, RMA workflow, Change Control & Rollback, monitoring of devices health and interfaces.

Cisco NDFC is optional. A VXLAN fabric can also be managed through the traditional CLI. But using Cisco NDFC has its own advantages. As stated above Cisco NDFC provides full automation support for all types of data center fabrics by eliminating the chance for human errors.

Nexus 9000 Series Switches: Nexus 9000 switches are data center switches for a hybrid cloud networking foundation. The Cisco Nexus 9000 Series offers modular and fixed form-factors and can deliver 1Gig to 800 Gig of line-rate forwarding.

VXLAN EVPN Fabric: VXLAN EVPN is the de-facto standard of building large scale data center fabrics, which provides seamless mobility of the hosts, tenant isolation, large name space for L2 segments, and traffic entropy across all the ECMP paths.

Spine Switches: In the VXLAN fabric, spine switches provide connectivity between all the leaf switches by using high speed links. Spines are not used to connect end-hosts.

Leaf Switches: Leaf switches function as VTEP and are responsible for the encapsulation and decapsulation of the VXLAN header. End-hosts are terminated on the leaf switches.





Solution Design

Prior to implementing the solution, it is important to understand the logical architecture of the Microsoft Azure Stack HCI and how it maps to the underlying physical architecture. This section describes the logical and physical connectivity of the Microsoft Azure Stack HCI, and the Nexus 9000 Series Switch based network with either the Cisco ACI or Cisco NX-OS mode.

Physical Architecture

Each Cisco UCS C240 M6/M7 server is connected to a pair of Cisco Nexus 9000 Top-of-Rack (ToR) switches using dual 100-Gb connections. In this example, the Cisco Nexus 9000 Series Switch based data center network carries all the Azure Stack HCI network traffic (management for host operating system, cluster communication, compute, and storage traffic). You can also use different networks.

Physical server management, such as Cisco Integrated Management Controller (CIMC) on Cisco UCS C series, is facilitated through an Out-of-band (OOB) management network that connects the server's dedicated management port to an OOB management switch with 1GbE links.

The following diagram illustrates a high-level physical architecture design:



Figure 5.

Physical Architecture (Cisco ACI or NX-OS mode)

In the case of Cisco NX-OS mode, the use of spine-leaf topology is not mandatory though it's a common design option whereas the Cisco ACI mode requires spine-leaf topology. Although downstream vPC is not used to connect the Microsoft Azure Stack HCI server to a pair of ToR switches, the use of vPC peer-link is recommended.

Note: As the only difference between ACI based fabric and NX-OS based fabric is a vPC peer-link, this document uses the topology illustration with a vPC peer-link. This vPC peer-link doesn't exist in the ACI fabric.

Physical connectivity considerations include the following:

- Microsoft recommends a 10+ Gigabit Ethernet network with remote-direct memory access (RDMA).
 For UCS C240 M6/M7 based Azure Stack HCl, the NVIDIA ConnectX-6X dual Port 100 Gigabit Ethernet NIC card is required. (Cisco VIC is currently not an option).
 - Microsoft requires that all server nodes be configured the same.
 - Up to 16 Azure Stack HCl servers per cluster.
- The Microsoft Azure Stack HCI server interfaces are connected to a pair of ToR switches with individual links, not Virtual Port Channel (vPC).
- The pair of ToR switches don't have to be dedicated to Azure Stack HCI connectivity.
- The ToR switches are configured for a MTU size of 9216. The MTU size for the packets sent on the network are controlled by the endpoints.

Logical Architecture

The network infrastructure for Azure Stack HCI consists of several logical networks:

- Tenant (Compute) Network: The tenant network is a VLAN trunk that carries one or more VLANs that
 provide access to the tenant virtual machines. Each VLAN is provisioned in the ToR switch and the
 SET switch that is running on the physical server. Each tenant VLAN is expected have an IP subnet
 assigned to it.
- Management Network (native VLAN is preferred but tagged VLAN is also supported): The
 management network is a VLAN that carries network traffic to the parent partition. This management
 network is used to access the host operating system. The connectivity to the management network
 is provided by the management (Mgmt) vNIC in the parent partition. Fault tolerance for the
 management vNIC is provided by the SET switch. A bandwidth limit can be assigned to the
 management, as necessary.
- Storage Network: The storage network carries RoCEv2 network traffic that is used for Storage Spaces Direct, storage replication, and Live Migration network traffic. The storage network has a Storage A and a Storage B segment, each with its own IP subnet. This design keeps the east-west RDMA isolated to the ToR switches.

In this document, the storage network is also used as a preferred path for cluster communication. (If both Storage A and Storage B segments are not available, the management network is used for cluster communication).

The following diagrams illustrate the tenant and management network (Figure 6) and storage network (Figure 7). For tenant and management network, ToRs provide the gateway functionality.

The default gateway of servers running on Azures Stack HCl are the anycast gateways provided by the ToRs.



Note: vPC peer-link doesn't exist for ACI fabric.

Figure 6.

Azure Stack HCI Logical Architecture (tenant and management network)

Unlike tenant and management networks, storage networks require separate VLANs to connect a pair of ToRs. For example, VLAN 10 is used to connect Leaf1 (Storage A segment) and VLAN 20 is used to connect Leaf2 (Storage B segment).



Note: vPC peer-link doesn't exist for ACI fabric.

Figure 7.

Azure Stack HCI Logical Architecture (storage network)

Storage network design considerations include the following:

- The storage network is used for Layer 2 communication only, where gateways on the ToR switches are not required.
- The storage network carries RoCEv2 traffic that is used for Storage Spaces Direct, storage replication, and Live Migration network traffic. Also used as a preferred path for cluster communication in this document.
- RoCE requires Data Center Bridging (DCB) to make the network lossless (DCB is optional for iWARP). If DCB is used, PFC and ETS configuration needs to be implemented in the network.
- As the storage network is also used as a preferred path for cluster communication in this document a different QoS configuration is required for storage traffic and cluster communication traffic. For example, Cos 4 is for storage traffic and Cos 7 is for cluster communication traffic.
 The following table shows the <u>QoS recommendations provided by Microsoft:</u>

Table 1.	Azure Stack HC	network QoS	recommendation
----------	----------------	-------------	----------------

	Cluster Communication Traffic	Storage traffic	Default (Tenant and Management Networks)
Purpose	Bandwidth reservation for cluster heatbeats	Bandwidth reservation for lossless RDMA communication used for Storage Spaces Direct	For all other traffic such as tenant networks.
Flow Control (PFC enabled)	No	Yes	No
Traffic Class	7	3 or 4	0 (default)
Bandwidth reservation	1% for 25GbE or higher RDMA networks 2% for 10GbE or lower RDMA networks	50%	Default (no host configuration required)

Note: Although the storage network is also used as a preferred path for cluster communication in this document, cluster communication could take any available network called as a preferred path. This path is chosen based on the metric role that is defined in the cluster network configured through Microsoft Network ATC. (Microsoft Network ATC provides an intent-based approach (management, compute, or storage) to host network deployment on the Azure Stack HCI servers. See <u>Microsoft Network ATC</u> <u>document</u> for details.) In this example, three cluster networks exist: Storage A, Storage B, and Management.

PS C:\Users\Administrator.MIHIGUCH> Get	t-Clus	terNetwo	ork
Name	State	Metric	Role
<pre>mgmt_compute_storage(Management)</pre>	Up	68800	ClusterAndClient
mgmt compute storage(Storage VLAN1601)	Up	19200	Cluster
mgmt_compute_storage(Storage_VLAN1602)	Up	19201	Cluster

Figure 8.

Azure Stack HCI Cluster Networks. The inside of an Azure Stack HCI server has the following network components:

- SET Switch: This is a virtual switch with embedded teaming capabilities. The SET switch provides teaming capabilities for network traffic that does not use the SMB-Multichannel. SMB Direct (RDMA) traffic uses SMB-Multichannel* to leverage available network connections for bandwidth and redundancy instead of the teaming feature in the SET switch.
- Guest Partition: The tenant virtual machines run in the guest partition on the Hyper-V host. Each virtual machine runs in isolation from others and does not have direct access to the physical hardware in the host. Network connectivity is provided to the tenant virtual machine by connecting synthetic NIC in the virtual machine to the SET switch on the host.
- Parent Partition: The parent partition is the host operating system that runs the virtualization
 management stack and has access to the physical server hardware. The parent partition has one
 management vNIC and two storage vNICs as shown in the example below. An optional dedicated
 vNIC for backup operations can be added, if needed.

* SMB Multichannel is part of the Server Message Block (SMB) 3.0 protocol, which increases the network performance and the availability of file servers. SMB Multichannel enables file servers to use multiple network connections simultaneously.

The following diagrams illustrate a logical network diagram within an Azure Stack HCl server. In this example, Storage A and Storage B are for the parent partition only, whereas management network is available for both parent partition and VMs in the guest partition. By default, the "Allow management operating system to share this network adapter" option is enabled on vNIC on the SET switch. In this example, it's enabled on the management vNIC (Yellow) whereas it's disabled on the storage vNICs (Green and Purple).



Note: vPC peer-link doesn't exist for ACI fabric.

Figure 9.

Azure Stack HCI Logical Architecture (SET Switch, Guest, and Parent Partitions)

MAC addresses for the VM virtual NICs are dynamically assigned, and the SET switch selects one of the available uplinks (physical NICs on the server) based on the source MAC address. This behavior provides load balancing and fault tolerance. The following diagram illustrates an example of how traffic from virtual machine A with virtual NIC MAC-A uses physical NIC1 as the uplink whereas traffic from virtual machine B with virtual NIC MAC-B uses physical NIC2 as the uplink. If the path using physical NIC1 is not available, all traffic goes through the other path.



Figure 10.

Load balancing behavior based on MAC address.

A consequence of this behavior is that some of the east-west network traffic that is not storage traffic transverses the spine (in the case of ACI) or vPC peer-link (in the case of NX-OS).



Figure 11. Traffic flow example

The network needs to allow the required traffic. Firewall requirements for Azure Stack HCl can be found at <u>https://learn.microsoft.com/en-us/azure-stack/hci/concepts/firewall-requirements.</u>

Cisco Nexus 9000 Series Switch based Fabric and Benefit

The table below lists the main features and benefits of the Nexus 9000 Series Switches based data center fabric.

Features	Benefit	ACI/NX-OS
Single point of Management	The use of the controller (APIC or NDFC) provides single point of configuration management and policy definition, which simplifies the operational aspects of the fabric.	ACI: APIC NX-OS: NDFC
Anycast Gateway	The fabric operates as an anycast gateway for the VMs on Azure Stack HCI servers and other physical/virtual servers. Layer 3 gateway functionality is provided by ToR switches instead of core or aggregation switches.	Both
VXLAN	The use of the VXLAN provides seamless Layer 2 and Layer 3 connectivity between servers, independently from the physical Leaf location. It also provides multi- tenancy.	Both
Multi-Pod/Multi-Site	Multi-Pod/Multi-Site fabric provides seamless Layer 2 and Layer 3 connectivity between endpoints, independently from the physical locations across data centers.	ACI: Multi-Pod, Multi-Site and Remote Leaf NX-OS: Multi-Site
Service Chaining	The use of Service Chaining capability provides flexible traffic redirection to L4- L7 service devices such as firewalls and load balancers.	ACI: Service Graph PBR NX-OS: ePBR

Figure 12

Cisco ACI connectivity options and policy domain evolution



- Single Fabric with End-to-End Encapsulation
- Single Overlay domain

- Multiple Fabrics with Integrated DCI
- Integrated DCI Scaling within and between Fabrics
- Multiple Overlay domains
- End-to-End automation support by NDFC

Figure 13.

Cisco Nexus 9000 Series Switch based Fabric and Benefit

Cisco ACI Design for Azure Stack HCI Connectivity

This section explains how Azure Stack HCI can connect to Cisco ACI by using the EPG and bridge domains.

This design assumes that the customer already has the Cisco ACI fabric in place with spine switches and APICs deployed and connected through a pair of leaf switches.

Cisco ACI for Azure Stack HCI Connectivity

The figure below shows the basic traffic flow of Azure Stack HCl traffic through the Cisco ACl fabric. In this design, the Cisco ACl fabric has two pairs of leaf nodes and two spine nodes, which are controlled by an APIC cluster. A pair of border leaf switches have the L3Out configured. This provides connection to a pair of external routers and thus to the Internet and Enterprise networks. Another pair of leaf nodes are connected to the Azure Stack HCl servers and other servers.



Figure 14.

Azure Stack HCI Traffic flow via Cisco ACI Fabric

In this design, each leaf switch is connected to the Azure Stack HCI servers by using the 100GbE links. The two links between the ACI leaf switches and each Azure Stack HCI server are individual connections instead of a port-channel or vPC.

The figure below illustrates an ACI interface configuration example along with the domain and the VLAN pool configuration. Although it's possible to use different interfaces on a pair of ToR switches, this document uses the same interfaces: **node-101 (ethernet1/11 and 1/12)** and **node-102 (ethernet1/11 and 1/12)**. The figure below illustrates an ACI interface configuration example.



Figure 15.

ACI leaf interface configuration for Azure Stack HCI servers

Azure Stack HCI ACI Tenant Model Overview

The figure 16 illustrates an example of a high-level relationship between various ACI tenant elements as deployed in the design by highlighting the Azure Stack HCI tenant. In this example, Azure Stack HCI tenant (HCI_tenant1) contains Virtual Routing and Forwarding (VRF), Bridge domains (BD), and end point groups (EPGs) for tenant networks, and the common tenant contains an external connectivity (L3Out) and EPGs for storage and management networks.

For Azure Stack HCl tenant networks to be able to communicate with other data center networks and access external networks, a contract must exist between the EPG in tenant **HCl1_tenant1** and the other EPG in the same tenant and the external EPG (L3Out EPG) in the common tenant. For the EPGs in storage network A and B, the storage traffic is within its segment (BD), then there is no need to configure a contract with another EPG.



Figure 16.

ACI Tenant Overview for Azure Stack HCI

In addition to the generic ACI configuration, the following configurations are required for the Azure Stack HCI network:

- Enable the required LLDP TLVs on the interfaces that are connected to the Azure Stack HCI servers
- QoS configuration for storage and cluster communication

For more information about configuring Cisco ACI and NDFC Fabric, see Solution Deployment.

Cisco NX-OS based Fabric Design for Azure Stack HCI Connectivity

This section explains how Azure Stack HCl can connect to Cisco Nexus 9000 Series Switches in the NX-OS mode. You can use the Cisco Nexus 9000 NX-OS based VXLAN or the traditional classical LAN fabrics to connect to the Azure HCl environments. VXLAN leverages ECMP based multipathing over L3 links between the spine switches and Leaf switches and the traditional classic LAN fabric uses the L2 links (between Access and Aggregation devices) running STP. VXLAN is gaining more popularity and adoption for building data center fabrics because of its benefits over the traditional classical LAN.

VXLAN uses CLOS architecture where Leafs (also known as VTEP) are used to connect the end-host and performs origination and termination of VXLAN tunnels while Spine switches provide layer-3 connectivity between the Leaf switches.

Both these fabrics can be built and managed by Cisco NDFC. This enables faster and error-free deployment unlike the CLI-based approach that was used previously. Cisco NDFC supports various fabric templates to cater to any kind of data center fabric deployment. For the interest of Azure HCI, Data Center VLXAN EVPN and Enhanced Classic LAN fabric templates are the ones which should be used. This document describes the steps and workflows to connect Azure HCI to the VXLAN fabric.

Cisco NX-OS based Fabric for Azure Stack HCI Connectivity

The figure below illustrates the basic traffic flow of Azure Stack HCI traffic through the NX-OS based VXLAN fabric.



Figure 17.

Azure Stack HCI Traffic flow through Cisco NX-OS based VXLAN fabric

In this design, a pair of leaf switches in vPC are connected to the Azure Stack HCl servers by using the 100 Gigabit Ethernet links. The two links between the leaf switches and each Azure Stack HCl server are individual connections instead of a port-channel or vPC.

Solution Deployment

This section provides a detailed procedure to configure the Cisco ACI and Cisco NDFC fabric to use in the environment. It also explains how to add new components to an existing Cisco ACI or the Cisco NDFC fabric.

Note: After the Cisco ACI or Cisco NDFC configuration is completed as per the procedure in this document, Azure Stack HCI cluster can be installed. Before you register the Azure Stack HCI, you can use the connectivity validator (Invoke-AzStackHciConnectivityValidation) on the Azure Stack HCI nodes or any other computer in the same network where you'll deploy the Azure Stack HCI cluster. This validator checks the network connectivity that is required to register the Azure Stack HCI cluster to Azure.

Note: This document does not cover the Cisco ACI or Cisco NDFC fabric deployment and the automated installation of Azure Stack HCI.

Table 3 lists the hardware and software versions that are used in this solution.

Table 3. Hardware and Software Versions

Layer	Device	Software version	Comments
Cisco ACI	Cisco APIC	6.0 (3e)	ACI Controller
	Cisco Nexus Switches in ACI Mode	16.0(3e)	ACI Spine and Leaf switches
Cisco NX-OS	Cisco NDFC	12.1.3b	NDFC
	Cisco Nexus Switches in NX-OS mode	10.2(3F)	ToR switches
Cisco Azure Stack HCI		2022H2	Azure Stack HCI release (Includes individual releases of software for all the devices that are part of Azure Stack HCI)

Cisco ACI Configuration for Azure Stack HCI

This section explains how to configure Cisco ACI for Azure Stack HCI servers with the assumption that the ACI fabric and APICs already exists in the customer's environment. This document does not cover the configuration required to bring the initial ACI fabric online.

The following are the configuration steps to configure Cisco ACI for Azure Stack HCI Servers:

- Configuring leaf interfaces connected to Azure Stack HCI servers
- Configure QoS
- Configure EPGs

Configuring Leaf Interfaces Connected to Azure Stack HCI Servers

This section contains the following steps:

- Create VLAN Pool for Azure Stack HCI Physical Domain
- Configure Physical Domain for Azure Stack HCI
- Create Attachable Access Entity Profile for Azure Stack HCI Physical Domain
- Create LLDP policy to enable the required TLVs for Azure Stack HCI
- Create Interface Priority Flow Control Policy to enable the required TLVs for Azure Stack HCI
- Create Interface Policy Group for Interfaces connected to Azure Stack HCI servers
- Associate the Interface Policy Group to the leaf interfaces connected to Azure Stack HCI servers

Figure 18 and Table 4, summarize the topology, interface, and physical domain configuration parameters used in this section. The connection uses four 100 GbE interfaces between ACI Leaf switches and Azure Stack HCI servers.



Figure 18.

Interface and physical domain configuration for Azure Stack HCI Servers

Table 4.	Interface and	physical	domain	configuration	for Azur	e Stack	HCI Servers
	internace and	pilysioai	aomani	ooningaradion	1017 La	C Otdok	1101 001 001

Interface	Interface Policy Group	LLDP Interface Policy	Interface PFC Policy	AAEP Name	Domain Name	Domain type	VLAN Pool
Leaf1 and Leaf2 Ethernet 1/11-12	Individual-HCI	HCI_LLDP (DCBXP: IEEE 802.1)	PFC-Auto	HCI_AAEP	HCI_phys	Physical	HCI_VLAN_pool (VLAN 1600- 1699)

Tables 5 and 6 summarize the common and the user tenant configuration parameters that are used in this section. The ACI Leaf switches serve as the gateway to the Azure Stack HCI networks except storage networks that are L2 only. Although contract names are listed for your reference, the Shared L3Out configuration in common tenant and contract configuration steps are not covered in this document.



Figure 19.

Tenant configuration example

Table 5. Azure Stack HCI common tenant configuration example

Property	Name
Tenant	common
Tenant VRF	common_VRF
Bridge domains	Storage-A in common_VRF (No subnet) Storage-B in common_VRF (No subnet) Mgmt in common_VRF (10.1.1.254/24)
Leaf nodes and interfaces	Node 101 & 102 ethernet1/11 and 1/12
EPGs	EPG Mgmt in BD Mgmt (VLAN 1600) EPG Storage-A in BD Storage-A (VLAN 1601) EPG Storage-B in BD Storage-B (VLAN 1602)
External EPG (L3 Out)	Shared_L3Out in common tenant
Contract	Allow-Shared-L3Out provided by common tenant

Table 6. Azure Stack HCI tenant configuration example

Property	Name
Tenant	HCI_tenant1
Tenant VRF	VRF1

Property	Name
Bridge domain	BD1 (192.168.1.254/24) in VRF1
Leaf nodes and interfaces	Node 101 & 102 ethernet1/11 and 1/12
EPGs	Web EPG in BD1 (VLAN 1611) App EPG in BD1 (VLAN 1612)
Contract	Allow-Shared-L3Out provided by common tenant Web-App contract defined in the tenant

Create VLAN Pool for Azure Stack HCI Physical Domain

In this section, you will create a VLAN pool to enable connectivity to the Azure Stack HCI.

To configure a VLAN pool to connect the Azure Stack HCI servers to the ACI Leaf switches, follow these steps:

- 1. From the APIC top navigation menu, select **Fabric > Access Policies**.
- 2. From the left navigation pane, expand and select **Pools** > **VLAN**.
- 3. Right-click and select Create VLAN Pool.
- 4. In the **Create Pool** pop-up window, specify a Name (For example, **HCI_VLAN_pool**) and for Allocation Mode, select **Static Allocation**.
- 5. For **Encap Blocks**, use the **[+]** button on the right to add VLANs to the VLAN Pool. In the **Create Ranges** pop-up window, configure the VLANs that need to be configured from the Leaf switches to the Azure Stack HCI servers. Leave the remaining parameters as is.

	sco	APIC									
Sy	stem	Tenants	Fabric	Virtual N	letworking	Admin	Operations	Apps	Integrations		
		itory Fal	bric Policies	Access Polic	ies -						
Poli	cies			©90	Pools - VL	.AN					
0	Quick S										
	interfac	e Configuratio									
	Switche	Contiguration			Croate	VI AN De	el				
>	Module				Create	VLAN PO	01				0
> 🖿	Interfac					Description: 0	CLVLAN_pool				
	Policies						proronnan				
l -	Pools	r and External				location Mode:	Dynamic Allocat	ion Sta	tic Allocation		
	VLA					Encap Blocks:					n +
	Crea	ate Rang	jes						8	vlode Role	
			Type: VLAN						-		
		Descrip	ption: optio	nal							
							_				
		R	Inge: VLAN	1600 Integer Valu	- VLAN	integer Value					
		Allocation N	Aode: Dyr	namic Allocation	Inherit alloch	Ade from parent	Static Alloca	rtion			
			Role: Ext	emal or On the wi	re encapsulations	Internal					
										Cancel	Submit
								Cance	ОК	[600-699]	
								Cance	OK	[600-699]	

- 6. Click OK.
- 7. Click Submit.

Configure Physical Domain for Azure Stack HCI

To create a physical domain type, connect to Azure Stack HCI servers, follow these steps:

- 1. From the APIC top navigation menu, select Fabric > Access Policies.
- 2. From the top navigation menu, select **Fabric > Access Policies**.
- 3. From the left navigation pane, expand and select **Physical and External Domains > Physical Domains**.
- 4. Right-click Physical Domains and select Create Physical Domain.
- In the Create Physical Domain pop-up window, specify a Name for the domain (For example, HCI_phys). For the VLAN Pool, select the previously created VLAN Pool (For example, HCI_VLAN_pool) from the drop-down list.

alialia cisco	APIC									
System	Tenants	Fabric	Virtual N	etworking	Admin	Operations	Apps	Integrations		
Inver	ntory Fal	oric Policies	Access Polici	es						
Policies			00	Physical Do	omains					
🗘 Quick S	tart			Create P	hysical	Domain				\otimes
Switch	e Configuratio Configuration	n			Name:	HCl_phys				
> 🚞 Switche				Associated	Attachable	select a value		\sim		
> 🔚 Module					VLAN Pool:	HCI_VLAN_pool(sta	tic)	~ B		
> 💼 Interfac	es			Securit	y Domains:				O +	
Policies	l and External	Domains				Select N	lame	Description		
> 🚞 Exte	rnal Bridged D	omains								
> 🔚 Fibre	e Channel Dom	ains								
> 🖬 L3 D	omains									
> 🚞 Phys	ical Domains									
> 🖿 Pools									Cancel	ubmit

Create Attachable Access Entity Profile for Azure Stack HCI Physical Domain

To create an Attachable Access Entity Profile (AAEP), follow these steps:

- 1. From the APIC top navigation menu, select **Fabric > Access Policies.**
- 2. From the left navigation pane, expand and select **Policies > Global > Attachable Access Entity Profiles**.
- 3. Right-click and select Create Attachable Access Entity Profile.
- 4. In the **Create Attachable Access Entity Profile** pop-up window, specify a Name (For example, **HCI_AAEP**) and **uncheck "Enable Infrastructure VLAN" and "Association to Interfaces".**
- 5. For the **Domains**, click the **[+]** on the right-side of the window and select the previously created domain from the drop-down list below **Domain Profile**.
- 6. Click Update.
- 7. You should now see the selected domain and the associated VLAN Pool as shown below.
- 8. Click **Next**. This profile is not associated with any interfaces at this time because "Association to Interfaces" is unchecked at step 4 above. They can be associated once the interfaces are configured in an upcoming section.

cisco APIC											
System Tenants Fabric	Virtual I	Networking	Admin	Operations	Apps	Integrations					
Inventory Fabric Policies	Access Poli	cies				-					
Policies	നൈഭര	Create At	tachabl	e Access E	ntity Pro	ofile					\otimes
Ouick Start										Drofile	
Interface Configuration		STEP 1 > Profi	le							. FIONR	- -
Switch Configuration			Name:	HCI_AAEP							
> 🚞 Switches			Description:	optional							
> 🚞 Modules											
> 🚞 Interfaces		Enable Infrastru	icture VLAN:								
V 🚍 Policies		Association to	o Interfaces:							-	
> 🧮 Switch		External) To Be	Associated								+
> 🚞 Interface		10	o interraces:	Domain Profile	ICI phys		Encapsula	-1600 to:vlan-1699			
✓ 🗖 Global				r nysical bollain	ioi_piiya		Tomvan	1000 10.0111 1000			
> PTP User Profile											
> DHCP Relay											
Fror Disabled Recovery Pol	licy										
MCP Instance Policy default	t										
> 🗖 QOS Class		EPG DEPLOYN	MENT (All Se	ected EPGs will be de	ployed on all th	e interfaces associate	ed.)				
> 🚞 Monitoring		Application El	0.00				Frees	Drimony Encon	Mada		+
> 🚞 Troubleshooting		Application El	PGS				Encap	Primary Encap	wode		
> 🚞 Physical and External Domains											
> 🚞 Pools											
								Previous Cane	cel	Finish	
								Previous Can	cel	Finish	

9. Click **Finish**.

Create LLDP policy to Enable the Required TLVs for Azure Stack HCI

To create an LLDP policy to enable the required TLVs for Azure Stack HCl, follow these steps:

- 1. From the APIC top navigation menu, select **Fabric > Fabric Policies**.
- 2. From the left navigation pane, expand and select Policies > Global > LLDP policy by default.
- 3. Check the following optional TLVs:
 - i. **DCBX** (for storage network)
 - ii. Port Link Aggregation
 - iii. Port Maximum Frame Size
 - iv. Port VLAN Name

Note: Port VLAN, that is also required for Azure Stack HCI, is always enabled regardless LLDP policy configuration.

cisco AP	IC (MinakoSi	te)						
System Tenan	ts Fabric	Virtua	I Networking	Admin	Operations	Apps		
Inventory	Fabric Policies	Access Po	licies					
Policies	$(\mathbf{\hat{r}})$		LLDP Policy -	default				
C Quick Start								
> 🚞 Pods								
> 🚞 Switches								
> 🚞 Modules			Properties					
> 🚞 Interfaces		Hold '	Time (sec):	120	\Diamond			
✓			Initial Delay	Time (sec):	2	\Diamond		
> 🚞 Pod			Transmit Frequ	ency (sec):	30	$\hat{>}$		
> 🚞 Switch			Optional TL	V Selector:				
> 🚞 Interface				E	DCBX			
🗸 🚞 Global				le le	Port Link Aggreg Port Maximum F	jation rame Size		
> 🚞 DNS Profile	es			Ē	Port Vlan Name			
> 🚞 Fabric L2 N	ити							
🗧 Multicast T	ree Policy default							
E LLDP Polic	y default							
Fabric Wild	Icard Rogue Exceptio	n						

4. Click Submit.

Create LLDP Interface Policy

To create an LLDP policy to enable the required TLVs for Azure Stack HCl, follow these steps:

- 1. From the APIC top navigation menu, select Fabric > Access Policies.
- 2. From the left navigation pane, expand and select **Policies > Interfaces > LLDP Interfaces**.
- 3. Right-click and select Create LLDP Interface Policy.
- 4. In the Create LLDP Interface Policy pop-up window, specify a Name (For example, HCI_LLDP).
- 5. Select Enable for Transmit State
- 6. Select IEEE 802.1 for DCBXP Version.

cisco APIC					
System Tenants Fabric	virtual Networking	Admin	Operations	Apps	Integrations
Inventory Fabric Policie	s Access Policies				
Policies	1 Therface	e - LLDP Inte	rface		
 Interface Configuration Switch Configuration 	▲ Name		labe	1	Receive State
> 🖬 Switches	Crea	te LLDP In Name:	terface Pol	licy	8
> 🖿 Interfaces ~ 🖿 Policies		Description:	optional		
 Switch Interface 802.1x Port Authentication CDP Interface COPP Interface Data Plane Policing DWDM Fibre Channel Interface Fibre Channel Interface Firewall 		Alias: Receive State: Transmit State: Warning: Chan converging. Th DCBXP Version:	Disabled E Disabled E ging the DCBX ve le link may need to CEE IEEE 8	nabled nabled ersion may to be reset	prevent the port parameters from for the change to take effect.
> 🚘 L2 Interface > 🚔 Link Flap > 🚰 Link Level > 🚰 Link Level Flow Control > 🚰 LLDP Interface					Cancel Submit

Create Interface Priority Flow Control Policy

To create an interface policy group to enable PFC on leaf downlinks, follow these steps:

- 1. From the APIC top navigation menu, select **Fabric > Access Policies**.
- 2. From the left navigation pane, expand and select Policies > Interface > Priority Flow Control
- 3. Right-click and select Create Priority Flow Control Policy.
- In the Create Priority Flow Control Policy pop-up window, specify a Name (For example PFC-Auto) and select Auto. (To include PFC configuration state via DCBX protocol, it needs to be set to Auto.)

cisco	APIC									
System	Tenants	Fabric	Virtual N	etworking	Admin	Operations	Apps	Integrations		
Inver	ntory Fabr	ic Policies	Access Polic	ies						
Policies > Policies > Swit > Swit Swit > Swit > Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit Swit	ch face 02.1x Port Auth DP Interface coPP Interface lata Plane Polici WDM ibre Channel Inf irewall 2 Interface ink Flap ink Level Interface Interface Interface IdP Interface IdP Interface IdP Interface IdFlow oE ort Channel ort Channel Me ort Security	entication ng terface Control mber	ſ ¶ 6	Priority Flo	Description: State:	e Priority F PFC-Auto optional Auto Off	Iow Cor	ntrol Policy	cel Submit	

Create Interface Policy Group for Interfaces connected to Azure Stack HCI servers

To create an interface policy group to connect to external gateways outside the ACI fabric, follow these steps:

- 1. From the APIC top navigation menu, select **Fabric > Access Policies**.
- 2. From the left navigation pane, expand and select **Interfaces > Leaf Interfaces > Policy Groups >** Leaf Access Port.
- 3. Right-click and select Create Leaf Access Port Policy Group.
- 4. In the Create Leaf Access Port Policy Group pop-up window, specify a Name (For example **Individual-HCI**) and the applicable interface policies from the drop-down list for each field.
- For the Attached Entity Profile, LLDP Policy and Priority Flow Control fields, select the previously created AAEP, LLDP policy and Priority Flow Control policy (For example, HCI_AAEP, HCI_LLDP and PFC-auto).

diale APIC		Create Leaf Acces	s Port Policy Gr	oup				\times
		Name:	Individual-HCI					
System Tenants Fabric	Virtual	Description:	optional					
Inventory Fabric Policies	Access Poli							
Policies		Attached Entity Profile:	HCI_AAEP	~ 🖉	Link Level Policy: sel	ect a value	\sim	
[CDP Policy:	select a value	\sim	LLDP Policy: HC	I_LLDP	~ 🛃	
Quick Start		View Advanced Settings 🗸						
Interface Configuration		902 1v Dort Authonticatio	an: coloct o voluo		MCP	coloct o voluo		
Switch Configuration		Transpooliver poli	select a value		Monitoring Deliau	select a value	Y	
> 🧮 Switches		i ransceiver polic	cy: select a value	Y	Monitoring Policy:	select a value	\sim	
> 🧮 Modules		CoPP Polic	cy: select a value	~	PoE Interface:	select a value	×	
✓		DWD	M: select a value	~	Port Security:	select a value	×	
✓		Egress Data Plane Policir	ng: select a value	~	Priority Flow Control:	PFC-Auto	[⊻	
> 🚞 Profiles		Fibre Channel Interfac	ce: select a value	\sim	Slow Drain:	select a value	\sim	
🗸 🚞 Policy Groups		Ingress Data Plane Policir	ng: select a value	\sim	Storm Control Interface:	select a value	\sim	
> 🚞 Leaf Access Port		L2 Interfac	ce: select a value	\sim	STP Interface Policy:	select a value	\sim	
> 🚞 PC Interface		Link Flap Polic	cy: select a value	\sim	SyncE Interface Policy:	select a value	\sim	
> 🚞 VPC Interface		Link Level Flow Control Polic	cy: select a value	\sim				
> 🚞 PC/VPC Override		MACse	ec: select a value	\sim				
> 🚞 Leaf Breakout Port Group								
> 🚞 FC Interface		NetFlow Monitor Policies:						+
> 🚞 FC PC Interface			NetFlow IP Filter Type		NetFlow Monit	tor Policy		
> 🚞 Overrides								
> 🚞 Spine Interfaces								
> 🖿 Policies								
> 🚞 Physical and External Domains								
> 🖿 Pools								
						Cance	I Submi	t

Associate the Interface Policy Group to the Leaf Interfaces Connected to Azure Stack HCI servers

To configure leaf interfaces connected to Azure Stack HCI servers, follow these steps:

- 1. From the APIC top navigation menu, select **Fabric > Access Policies**.
- 2. From the left navigation pane, select Interface Configuration.
- 3. From the right pane, right-click **Actions** and select **Configure Interfaces**.

cisco	APIC		admin 🔇 🗩 👯 🛱 🥐 💷				
System	Tenants Fabric	Virtual Networking Admin Operations Apps Integrations					
Inve	ntory Fabric Policie	s Access Policies					
Policies	•	\odot					
O Quick Start Interface Configuration							
🗄 Interfac	e Configuration		_				
Switch (Configuration	Some of the interfaces are still configured using Selectors and Profiles. We can help you migrate them.					
> Switche							
> Interfac	ac						
> Policies		Filter by attributes	Actions 🔿 💠				
> 🚞 Physica	l and External Domains	P., * Node * Interface * Port Type Admin State Port Mode Policy Group Interface Desc	Configure Interfaces				
> 🚞 Pools			Convert Interfaces				

- 4. In the **Configure interfaces** window, select the following options.
 - i. Node Type: Leaf
 - ii. Port Type: Access

- iii. Interface Type: Ethernet
- iv. Interface Aggregation Type: Individual
- 5. Click **Select Node**. In the Select Nodes pop-up window, select leaf nodes to connect Azure Stack HCI servers (For example, Node 101-102) and click **OK**.
- 6. Specify the Leaf interfaces to connect Azure Stack HCl servers (For example, 1/11-12).

Configure Interfaces						
General						
Node Type Leaf Spine						
Port Type Access Fabric						
Interface Type Ethernet Fibre Channel						
Interface Aggregation Type						
Node* 🛈						
101-102	Select Node					
Interfaces For All Switches * 🛈						
1/11-12						

 Click Select Leaf Access Port Policy Group. In the Select Leaf Access Port Policy Group pop-up window, select the previously created Leaf Access Port Policy Group (For example, Individual-HCI) from the list, and click Select.

)			
)			
)			
hannel			
ypc			
	Select Node		
I Switches * 🛈			
t Policy Group*			
S			
Name	Interfaces	Configuration Status	
Pod1-Leaf1	1/11-12	 Configuration will be updated 	/ 1
Pod1-Leaf2	1/11-12	 Configuration will be updated 	/ 1
ji I	ation Type vPC Switches * Policy Group • Name Pod1-Leaf1 Pod1-Leaf2	ation Type vPC Select Node Switches * Policy Group Name Interfaces Pod1-Leaf1 1/11-12 Pod1-Leaf2 1/11-12	ation Type vPC Select Node Switches • Policy Group • s Name Interfaces Configuration Status Pod1-Leaf1 1/11-12 O Configuration will be updated Pod1-Leaf2 1/11-12 O Configuration will be updated

8. Click Save.

Configure QoS

The table below summarizes the host network QoS recommendation from Microsoft. Please refer to the Microsoft document for details: <u>https://learn.microsoft.com/en-us/azure-stack/hci/concepts/host-network-requirements</u>.

Table 7. Azure Stack HCI host network QoS recommendation

	Cluster Communication Traffic	Storage traffic	Default (Tenant and Management Networks)
Purpose	Bandwidth reservation for cluster heatbeats	Bandwidth reservation for lossless RDMA communication for Storage Spaces Direct	For all other traffic such as tenant networks.
Flow Control (PFC enabled)	No	Yes	No
Bandwidth reservation	1% for 25GbE or higher RDMA networks 2% for 10GbE or lower RDMA networks	50%	Default (no host configuration required)

Based on the recommendation, this document uses the following ACI QoS configurations as an example, which are the same as the bandwidth reservation and Priority configurations that are used in <u>the Cisco UCS</u> <u>C240 M6 Solution for Microsoft Azure Stack HCI</u>.

- Level1 for RDMA (storage) traffic (Traffic comes with Cos 4 marked by Azure Stack HCI)
 - $\circ \quad \text{PFC is enabled}$

- Bandwidth reservation: 50% 0
- ETS (Weighted round robin in ACI) \bigcirc
- Level2 for cluster communication (Traffic comes with Cos 5 marked by Azure Stack HCI) •
 - PFC is not enabled 0
 - Bandwidth reservation: 1% 0
 - ETS (Weighted round robin in ACI) 0
- Level3(default) for VM traffic and management traffic (Other traffic) •
 - PFC is not enabled 0
 - Bandwidth reservation: 49% 0
 - ETS (Weighted round robin in ACI) 0

The following figure illustrates an example of QoS configuration.



QoS config on ACI fabrics

- Level1: For Storage EPGs Cos 4. 50%. PFC is enabled for Cos 4.
- Level2: For Storage EPGs Cos 5. 1%.
- Level3(default): default configuration for other EPGs. 49%

QoS config on AzureStack HCI

- Storage: 50% (Priority 4 = Cos 4)
- Cluster: 1% (Priority $5 = \cos 5$)
- (Cluster traffic is also carried over Storage networks: Storage A and Storage B)

Figure 20.

ACI QoS configuration for Azure Stack HCI

The Cisco ACI fabric supports six user-configurable QoS levels (Level1-6) as well as two levels reserved for fabric control traffic, SPAN, and traceroute traffic.

Table 8.	Cisco	ACI	QoS	Levels
----------	-------	-----	-----	--------

Class of Service	QoS Group Used by DCBX (ETS configuration and ETS recommendation) *	Traffic Type	Doc1p (Cos) Marking in VXLAN Header	DEI Bit**
0	0	Level 3 (default)	0	0
1	1	Level 2	1	0
2	2	Level 1	2	0

Class of Service	QoS Group Used by DCBX (ETS configuration and ETS recommendation) *	Traffic Type	Doc1p (Cos) Marking in VXLAN Header	DEI Bit**
4	7	Level 6	2	1
5	6	Level 5	3	1
6	5	Level 4	5	1
3	3	APIC Controller	3	0
9	Not Advertised	SPAN	4	0
8 (SUP)	4	Control	5	0
8 (SUP)	4	Traceroute	6	0
7	Not Advertised	Copy Service	7	0

* In IEEE DCBX PFC configuration LLDP TLV, the Priority value is the associated Cos value regardless of which Level (Level 1-6) the PFC is enabled. The configuration section below includes an example.

**The Drop Eligible Indicator (DEI) bit is a 1-bit field that is used to indicate frames that are eligible to be dropped during traffic congestion. The CoS value (3 bits) + DEI value (1 bit) represents the QoS class.

Configure QoS Classes

To configure Cisco ACI QoS classes, follow these steps:

- 1. From the APIC top navigation menu, select Fabric > Access Policies.
- From the left navigation pane, expand Policies > Global > QoS Class and select one of the levels. (For example, level1 for storage traffic).
- 3. In the **Scheduling algorithm** field, from the drop-down list, choose **Weighted round robin.** This is the default configuration.
- 4. In the Bandwidth allocation (in %) field, specify a number. (For example, **50** for storage traffic).
- 5. If PFC is not required in the class, leave PFC Admin State field unchecked.
- 6. If PFC is required in the class,
 - a. Check PFC Admin State field
 - b. In the No Drop-Cos field, select Cos value (For example, Cos 4 for storage traffic)
 - c. In the scope field, select **Fabric-wide PFC**. (If the traffic is within the same leaf, IntraTor PFC is also fine)

System	Tenar	nts	Fabric	Virtu	al N	etworking	Admin	Operations	Apps	Integration	าร	
Inve	entory	Fab	ric Policies	Access	Polic	ies						
Policies			(00		QOS Class	Policy - Lev	vel1				0
C Quick S	Start ce Config Configura	uration ation			l					Policy	Histo Ö	ry +
> 🖬 Switch	es					Properties	QoS Class:	: Level1				
> 📩 Interfac	ces						Admin State: MTU:	Enabled	\sim			
	tch rface					M	linimum buffers:	0		dom opriv doto	otion	
✓ 🚍 Gioi ✓ 🚍 Gioi	bal PTP User	Profile			4	Queue	control method:	: Dynamic		dom early dete		
> 🗖 (> 🗖 (OHCP Rela	ay e Acce:	ss Entity Prof	ïles		Bandwidth	allocated (in %): FC Admin State:	50				
	Error Disa MCP Insta	bled Re ince Po	ecovery Polic licy default	у			No-Drop-CoS	Cos 4 When PFC Admin	State is unchecked	, this field value wi	ll be set to	emp
	QOS Class						Scope:	Fabric-wide	e PFC Intra			
	Level2 Level3	(Defau	ilt)									
	Level4 Level5 Level6							Show U	Jsage F	Reset	Submit	

With this QoS configuration and LLDP IEEE DCBX configuration, the following values are set in LLDP.

- IEEE ETS Configuration and IEEE ETS Recommendation
 - PGID for Prio 4: 2 (because Cos 4 is selected and level1 is QoS group 2)
 - Bandwidth for PGID 2: 50 (level1 is QoS group 2)
 - TSA for Traffic Class 2: Enhanced Transmission Selection (level1 is QoS group 2)
- IEEE Priority Flow Control Configuration
 - PFC for Priority 4: Enabled (because Cos 4 is selected, and PFC is enabled)


By default, all "PGID for Pri 0" to "PGID for Pri 7" are set to 0 and all "PFC for Priority 0" to "PFC for Priority 7" are set to Disabled. If PFC is enabled, the value for the specific priority (Cos value) is updated. ("PGID for Pri 4: 2" and "PFC for Priority 4" in the example above.)

- 8. Repeat step 2 -7 for the level for cluster communication traffic. For example, **level2** for cluster communication traffic with **1%** bandwidth reservation configuration is the following:
- QoS Class: Level2
- Scheduling algorithm: Weighted round robin (default configuration)
- Bandwidth allocation (in %): 1
- PFC Admin State: unchecked

With this QoS configuration and LLDP IEEE DCBX configuration, the following values are set in LLDP. There is no change on PGID and PFC for Priority 0-3 and 5-7.

- IEEE ETS Configuration and IEEE ETS Recommendation
 - a. Bandwidth for PGID 1: 1 (because level2 is QoS group 1 based on table 8)
 - b. TSA for Traffic Class 1: Enhanced Transmission Selection
- 9. Repeat step 2 -7 for the level other traffic. For example, **level3(Default)** for VM traffic with **49%** bandwidth reservation configuration is the following:
- QoS Class: level3(Default)
- Scheduling algorithm: Weighted round robin (default configuration)
- Bandwidth allocation (in %): 49
- PFC Admin State: unchecked

With this QoS configuration and LLDP IEEE DCBX configuration, the following values are set in LLDP. There is no change on PGID and PFC for Priority 0-3 and 5-7.

- IEEE ETS Configuration and IEEE ETS Recommendation
 - a. Bandwidth for PGID 0: 10 (because level3 is QoS group 0 based on table 8)
 - b. TSA for Traffic Class 0: Enhanced Transmission Selection



Configure Custom QoS Policy

ACI has multiple QoS classification options that are illustrated in the figure below.



Figure 21. ACI QoS configuration priority

This document uses QoS Class configuration at EPGs for tenant and management networks (default level3), and uses the custom QoS policy configuration at EPG for storage and cluster communication network (level1 for storage with Cos 4 and level2 for cluster communication with Cos 5).



Figure 22.



To configure a Custom QoS policy, follow these steps:

- 1. From the APIC top navigation menu, select **Tenants > common** (or select an existing tenant where you want to configure EPGs).
- 2. From the left navigation pane, expand and select Policies > Protocol > Custom QoS.
- 3. Right-click and select **Create Custom QoS Policy** to open the **Create Custom QOS Policy** pop-up window.
- 4. In the Name field, specify a Name (For example, Storage_and_Cluster).
- 5. In the **Dot1P Classifiers** field, click + and configure the followings:
 - a. Priority (In this example, select level2 from the drop-down list for storage traffic)
 - b. Dot1P Range From and To (In this example, specify 4 for storage traffic)
- 6. Click Update.
- 7. Repeat step 5-6 for cluster communication traffic. (In this example, **level1 with 5** for cluster communication traffic.)

APIC (172.31.184.201)	Create Custom	QOS Polic	у				8
cisco	Name:	Storage_and_Clust	br .				
System Tenants Fabric Virtual N	Description:	optional					
ALL TENANTS Add Tenant Tenant Search:							
	DSCP to priority map:						+
This object was created by the Nexus Das		Priority	DSCP Range From	DSCP Range To	DSCP Target	Target CoS	
common (P)()()							
Quick Start							
- 🛄 common							
> 🚍 Application Profiles	Part Constant						
> 🚍 Networking	Dottip Classifiers:					1	+
> 🔤 Contracta		Priority	Dot1P Range From	Dot1P Range To	DSCP Target	Target CoS	
v 📰 Policies		Level2	4	4	Unspecified	Unspecified	
🗸 🚞 Protocol		Level1	5	5	Unspecified	Unspecified	
) 🚍 8FD							
) 🚍 BFD Multikop							
) 🧱 ND RA Prefix							
) 🚍 SGP							
Connectivity Instrumentation Policy							
default Creater Custom DoS Policy					Cance	Submit	
) 🧰 Data Plane Policing					Gance	Jucinic	

8. Click Submit.

This Custom QoS Policy is referred to in the next step (Configuring EPGs)

Configure EPGs

The following EPGs are created in this section.

- Tenant EPGs for VMs
- Management EPG for management network
- Storage EPGs for storage networks
- Configure contracts
- Add consumer and provider EPGs to the contract

Configure Tenant EPGs

To configure a tenant EPG for Azure Stack HCI VMs, follow these steps:

- 1. From the APIC top navigation menu, select Tenants > Add Tenant
- 2. In the **Create Tenant** dialog box, specify a Name (For example, **HCI_tenant1**).
- 3. In the **VRF Name** field, enter the VRF name (For example, **VRF1**).
- 4. Check Create A Bridge Domain and click Next.

ADIC (172 31184	Create VRF	0
cisco AFIC (1/2.01.104.	STEP 1× VRF	1. VRF 2. Bridge Domain
System Tenants Fabric	Name: VRF1	
ALL TEMANTO I AND THE TOPOLO	Allas:	
ALL TENANTS Add Tenant Tenant	Description: optional	
HCI_tenant1	Annotations: 🚭 Click to add a new annotation	
> 🕩 Quick Start	Policy Control Enforcement Preference: Enforced Unenforced	
✓ Ⅲ HCI_tenant1	Policy Control Enforcement Direction: Egress Ingress Mixed policy	
Application Profiles	BD Enforcement Status:	
✓ ➡ Networking	Endpoint Ketention Policy: Select a value V This policy only applies to remote L3 entries	
🔚 Bridge Domains	Monitoring Policy: select a value	
VRFs	DNS Labels: enter names separated by	
> L2Outs Create VRF	Transit Route Tag Policy: select a value	
> 🖿 L3Outs	IP Data-plane Learning: Disabled Enabled	
	Create A Bridge Domain: 🗹	
	Configure BGP Policies:	
	Configure OSPF Policies:	
	Configure EIGRP Policies:	
		Previous Cancel Next

5. In the Name field, specify a Name (For example, BD1) and click Finish.

Create VRF	8
STEP 2 > Bridge Domain	1. VRF 2. Bridge Domain
Name: BD1	
Alias:	
Description: optional	
Type: fc regular	
Forwarding: Optimize	
IP Data-plane Learning: Yes No	
Limit Local IP Learning To BD/EPG Subnet(s): 🗹	
(i) Info: This option is not available when "Enforce Subnet Check" is enabled from	"System Settings" \rightarrow "Fabric-Wide Settings Policy".
Config BD MAC Address: MAC Address: 00:22:BD:F8:19:FF	
	Previous Cancel Finish

- 6. To create an anycast gateway IP address on the bridge domain, in the Navigation pane, expand the created bridge domain (**BD1**) under **Networking > Bridge Domains**.
- 7. Right-click Subnets and choose Create Subnet.

8. In the **Gateway IP** field, configure the anycast gateway IP address (In this example, **192.168.1.254/24**), and click **Submit**.

	Create Subnet	\otimes
cisco AFIO (172.01.104.201)	Gateway IP: 192.168.1.254/24	
System Tenants Fabric Virtu	Treat as virtual IP address:	
ALL TENANTS Add Tenant Tenant Search:	Make this IP address primary: 🔲	
HCI_tenant1	Scope: Advertised Externally	
C Quick Start	Description: optional	
✓ I HCl_tenant1		
Application Profiles	Subnet Control: No Default SVI Gateway	
✓		
🗸 🚞 Bridge Domains	IP Data-plane Learning: Disabled Enabled	
∨ (1) вD1	L3 Out for Route Profile: select a value	
> 🖿 DHCP Relay Labels	ND RA Prefix Policy: select a value	
> 🧮 ND Proxy Subnets	Policy Tags: 🕕 Click to add a new tag	
> 🖬 Subnets		
VRFs Create Subnet		
> 🐽 VRF1		
		Cancel Submit

- 9. To create an Application Profile, from the left navigation pane, right-click **Application Profiles** and choose **Create Application Profile**.
- 10.In the Name field, specify a Name (For example, AP1) and click Submit.
- 11.To create an EPG, from the left navigation pane, expand the created Application Profile, right-click **Application EPGs** and choose **Create Application EPG.**
- 12.In the Name field, specify a Name (For example, Web).
- 13.In the **QoS class** field, from the drop-down list, choose a Level. (For example, **Level3 (Default)** for VM traffic, which is the default configuration)
- 14.In the **Bridge Domain** field, from the drop-down list, choose the BD we created (In this example, **BD1**).
- 15.Check Statically Link with Leaves/Paths and click Next.

APIC (172.31.184.201)	Create Application EPG	0
	STEP 1 - Identity	1. identity 2. Leaves/Paths
System Tenants Fabric Virtual Ne	Name: Wob	Concernance and the second
ALL TENANTS Add Tenant Tenant Search: nam	Allas:	
	Description: optional	
HCI_tenant1		
A Oulick Start	Annotations: CSck to add a new envolution	
	Contract Exception Tag:	
✓	QoS class: Level3 (Default)	
🗸 🚞 Application Profiles	Custom GoS: select a value -	
∼ 🖓 AP1	Data-Plane Policer: select a value	
Application EPGs	Intra EPG Isolation: Enforced Unenforced	
> uSeg EPGs Create Application EPG	Preferred Group Member: Exclude Include	
> 🖿 Endpoint Security Groups	Flood in Encapsulation Disabled Enabled	
	Bridge Domwin: BD1 🗸 🚱	
	Monitoring Policy: select a value	
Bridge Domains	FHS Trust Control Policy: select a value 🔤	
∨ ()) вD1	EPO Admin State: Admin Up Admin Shut	
> 🚞 DHCP Relay Labels	Associate to VM Domain Profiles:	
> 🧮 ND Proxy Subnets	Statically Link with Leaves/Paths: 🛃	
V 🚍 Subnets	EPG Contract Master	12 +
	Application EPGs	
- 192.108.1.234/24		
V 🗖 VRFs		
> 🕂 VRF1		
		Cancel Next

Note: QoS class is Level3 (Default) for the tenant EPG, which doesn't enable PFC by default.

- 16.In the Physical Domain field, from the drop-down list, choose the physical domain we created (In this example, **HCI_phys**).
- 17.In the **Paths** field, click + and select a Path and configure Port Encap. (In this example, **Pod-1/Node-101/eth1/11** and **vlan-1611** for **Web**).
- 18.Repeat step 17 to add all the interfaces that are connected to Azure Stack HCl servers in the cluster. (In this example, Node-101/eth1/11-12 and Node-102/eth1/11-12 with vlan-1611 for Web).
- 19.Repeat step 11-18 for other tenant EPGs (For example, EPG App with vlan-1612).

Configure a Management EPG

To configure Azure Stack HCI storage networking, follow these steps:

- 1. From the APIC top navigation menu, select **Tenants > common** (or select an existing tenant where you want to configure a management EPG).
- 2. From the left navigation pane, expand and select Networking > Bridge Domains.
- 3. Right-click and select Create Bridge Domain.
- 4. In the **Name** field, specify a Name (For example, **Mgmt**) and select a VRF name (In this example, **common-VRF**).
- 5. Click Next.
- 6. In the **Subnets** field, click + to create subnet.
- 7. In the Gateway IP field, specify an IP (For example, 10.1.1.254/24).
- 8. Click **OK**.

- 9. To create an EPG, from the left navigation pane, expand **Application Profiles** and select an existing Application Profile (or create a new Application Profile).
- 10.Right-click Application EPGs and select Create Application EPG.
- 11.In the Name field, specify a Name (For example, Mgmt).
- 12.the **QoS class** field, from the drop-down list, choose a Level. (For example, **Level3(Default)** for management traffic).
- 13.In the **Bridge Domain** field, from the drop-down list, choose the BD we created (In this example, **Mgmt**).
- 14.Check Statically Link with Leaves/Paths and click Next.
- 15.In the **Physical Domain** field, from the drop-down list, choose the physical domain we created (In this example, **HCI_phys**).
- 16.In the Paths field, click + and select a Path and configure Port Encap (In this example, Pod-1/Node-101/eth1/11 and vlan-1600 for Mgmt). If native VLAN (untagged) is used for management network, select Trunk (Native) in the Mode field.
- 17.Repeat step 16 for other Azure Stack HCl server interfaces in the cluster. (In this example, **Node-101/eth1/11-12 and Node-102/eth1/11-12** with **vlan-1600** for **Mgmt**).

Configure Storage EPGs

To configure Azure Stack HCI storage networking, follow these steps:

- 1. From the APIC top navigation menu, select **Tenants > common** (or select an existing tenant where you want to configure storage EPGs).
- 2. From the left navigation pane, expand and select Networking > Bridge Domains.
- 3. Right-click and select Create Bridge Domain.
- 4. In the **Name** field, specify a Name (For example, **Storage-A**) and select a VRF name (In this example, **common-VRF**).
- 5. In the Forwarding field, from the drop-down list, choose Custom.
- 6. In the L2 Unknown Unicast field, from the drop-down list, choose Flood.
- 7. Click Next.
- 8. Uncheck Unicast Routing checkbox to disable Unicast Routing and click Next.
- 9. Click Finish.
- 10.To create an EPG, from the left navigation pane, expand **Application Profiles** and select an existing Application Profile (or create a new Application Profile).
- 11.Right-click Application EPGs and select Create Application EPG.
- 12.In the Name field, specify a Name (For example, Storage-A).
- 13.In the **Custom QoS** field, from the drop-down list, choose the Custom QOS Policy we created (In this example, **Storage_and_Cluster**).

14.In the **Bridge Domain** field, from the drop-down list, choose the BD we created (In this example, **Storage-A**).

	PG
TEP 1 > Identity	1. Identity 2. Leaves/Paths
Name:	Storage-A
Alias:	
Description	optional
Annotations:	Click to add a new annotation
Contract Exception Tag:	
QoS class:	Level3 (Default)
Custom QoS	Storage_and_Cluster 🗸 🔁
Data-Plane Policer:	select a value
Intra EPG Isolation:	Enforced Unenforced
Preferred Group Member:	Exclude Include
Flood in Encapsulation:	Disabled Enabled
Bridge Domain:	Storage-A 🗸 🔁
Monitoring Policy:	select a value
FHS Trust Control Policy:	select a value
EPG Admin State:	Admin Up Admin Shut
Associate to VM Domain Profiles:	
Statically Link with Leaves/Paths:	
EPG Contract Master:	· · · · · · · · · · · · · · · · · · ·

15.Check Statically Link with Leaves/Paths and click Next.

- 16.In the **Physical Domain** field, from the drop-down list, choose the physical domain we created (In this example, **HCI_phys**).
- 17.In the **Paths** field, click + and select a Path and configure Port Encap (In this example, **Pod-1/Node-101/eth1/11** and **vlan-107** for **Storage-A**).
- 18.Repeat step 17 for other Azure Stack HCl servers in the cluster (In this example, **Pod-1/Node-102/eth1/11** and **vlan-107** for **Storage-A**).
- 19.Repeat step 2-21 for the second storage EPG (For example, Storage-B and EPG Storage-B using the created Custom QoS Storage_and_Cluster, physical domain HCl_phys and Path Pod-1/Node-101/eth1/12 and Pod-1/Node-102/eth1/12 with vlan-207).

Configure Contracts

To configure a contract, follow these steps:

- 1. From the APIC top navigation menu, select **Tenants** and select a tenant where the provider EPG resides. For example, select tenant **HCI_tenant1** for a contract between Web and App EPGs.
- 2. From the left navigation pane, expand and select Contracts.
- 3. Right-click and select Create Contract.
- 4. In the Name field, specify a Name (For example, Web-to-App).

- 5. In the **Scope** field, from the drop-down list, choose a Scope (In this example, **VRF.** If it's intertenant contract, select **Global**.)
- 6. In the Subjects field, click + and specify a contract subject name. (For example, Subject1.)
- 7. In the **Filter** field, click **+** and choose an existing filter (or create a new filter from the drop-down list).
- 8. Click **Update** and repeat step 7, if you have another filter.
- 9. Click **OK**.

	Create Contrac	t			\times
cisco APIC (1/2.51.164	Name:	Web-to-App			
System Topants Eabric	Alias:				
System Tenants Fabric	Scope:	VRF	\sim		
ALL TENANTS Add Tenant Tenant	QoS Class:	Unspecified	\sim		
HCI tenant1	Target DSCP:	Unspecified	\sim		
	Description:	optional			
> 🕩 Quick Start					
✓	Annotations:	Click to add a	new annotation		
> En Application Profiles	Subjects:			Ť	+
		Name	Description		
		Subject1			
Create Contract					
> 🖬 Standard					
> Taboos					
> 🛅 Imported Create Filter					
			Ca	ncel Submit	

10.Click Submit.

11.Repeat step 1-10 if you have another contract.

Add Consumer/Provider EPGs to the contract

To add an EPG to a contract, follow these steps:

- 1. From the APIC top navigation menu, select **Tenants** and select a tenant where the EPG resides. For example, select tenant **HCI_tenant1** for a contract between Web and App EPGs.
- 2. From the left navigation pane, expand **Application Profiles** and expand the Application Profile where the EPG resides.
- 3. Expand Application EPGs and expand the EPG. (For example, Web).
- 4. Right-click **Contracts** and select **Add Provided Contract** or **Add Consumed Contract** depending on whether the EPG is the provider or the consumer. (In this example, Web EPG is the consumer to the contract).
- 5. In the **Contract** field, from the drop-down list, choose the contract we created (In this example, **Web-to-App**).



- 6. Click Submit.
- 7. Repeat step 1-6 for other EPGs.

Cisco NX-OS based Fabric configuration for Azure Stack HCI

This section explains how to configure Cisco NX-OS based VXLAN fabric for Azure Stack HCI servers with the assumption that the VXLAN fabric managed by Cisco NDFC already exists in the customer's environment. This document does not cover the configuration required to bring the initial VXLAN fabric. For building IGP based Underlay and iBGP based Overlay (BGP EVPN), **Data Center VXLAN EVPN** fabric template should be used.

This document does not cover NX-OS based traditional classical LAN fabric however, the same workflow can be followed for traditional classical LAN fabrics. NDFC comes with **Enhanced Classic LAN (ECL)** fabric template for building NX-OS based traditional classical LAN fabrics.

The overall configuration can be categorized as below:

- Configure QoS
- LLDP configuration
- · Configuring leaf interfaces connected to Azure Stack HCI servers
- Configuration of Networks and VRFs
- Configuring External connectivity

Configure QoS

The QoS requirement for Azure Atack HCI host is same for both ACI and NX-OS based fabrics. For more details, please refer <u>Table 7 Azure Stack HCI host network QoS recommendation</u>.

Only the switches connected to Azure Stack HCl servers need to have the required QoS configurations as shown below:

Create Class-maps to classify RDMA and cluster communication traffic on ingress interface based on CoS markings set by the Azure Stack HCl servers -

```
class-map type qos match-all RDMA
  match cos 4
class-map type qos match-all CLUSTER-COMM
  match cos 5
```

Once the traffic is classified (based on CoS value set by the Server) it needs to be mapped to the respective QoS Groups -

```
policy-map type qos AzS_HCI_QoS
  class RDMA
   set qos-group 4
  class CLUSTER-COMM
   set qos-group 5
```

Define Network QoS classes and match traffic based on the QoS Groups -

```
class-map type network-qos RDMA_CL_Map_NetQos
  match qos-group 4
  class-map type network-qos Cluster-Comm_CL_Map_NetQos
  match qos-group 5
```

Create Network QoS policy to enable PFC for RDMA traffic and set Jumbo MTU -

```
policy-map type network-qos QOS_NETWORK
class type network-qos RDMA_CL_Map_NetQos
    pause pfc-cos 4
    mtu 9216
class type network-qos Cluster-Comm_CL_Map_NetQos
    mtu 9216
class type network-qos class-default
    mtu 9216
```

Configure Queuing policy to enable ECN for RDMA traffic and bandwidth allocation for other classes -

```
policy-map type queuing QOS EGRESS PORT
 class type queuing c-out-8q-q-default
   bandwidth remaining percent 49
 class type queuing c-out-8q-q1
   bandwidth remaining percent 0
 class type queuing c-out-8q-q2
   bandwidth remaining percent 0
 class type queuing c-out-8q-q3
   bandwidth remaining percent 0
 class type queuing c-out-8q-q4
   bandwidth remaining percent 50
   random-detect minimum-threshold 300 kbytes maximum-threshold 300 kbytes drop-probability 100
weight 0 ecn
 class type queuing c-out-8q-q5
   bandwidth percent 1
 class type queuing c-out-8q-q6
   bandwidth remaining percent 0
 class type queuing c-out-8q-q7
   bandwidth remaining percent 0
```

Apply the Queuing and Network QoS policies to System QoS -

```
system qos
service-policy type queuing output QOS_EGRESS_PORT
service-policy type network-qos QOS_NETWORK
```

The above QoS configuration is only required on the Leaf switches that are used to connect Azure Stack HCI servers. There is no requirement of fabric-wide QoS configuration as long as all the Azure Stack HCI servers of same cluster are connected to same vPC pair of Leafs.

The steps to configure the QoS policies through NDFC are as follows:

Step 1: Select both the Leaf switches (connecting to Azure Stack HCI) and create a Group Policy using **switch_freefrom** policy template and paste all the QoS related configuration (shown above) in Switch Freeform Config box.

To create a policy, go to Fabric **Detailed View > Policies** Tab.

Switch List: i inst2 Priority* 500 0 10 10 10 10 10 10 10 10 1	Switch Lis: i i 2000 i 1 2000 Description Cost configuration for Azure HCI Cluster Description Cost configuration for Azure HCI Cluster Cost configuration for Azure HCI Cluster Description Cost configuration for Azure HCI Cluster Description Cost configuration for Azure HCI Cluster Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Description Descript	e Policy		? – \times
Switch List: verticity* 500 500 Casc configuration for Azure HCI Cluster Casc configuration for Azure HCI Cluster Select Template* switch Lifetion 5 Corp Croup Switch Freeform Config* Lides-map type entition(-sigs RDMA_CL_MAD_NetCos match cos 4 Class-map type entit entit-sigs RDMA_CL_MAD_NetCos policy-map type entit entit -sig	Switch Fuedom Config" Switch Fuedom Config Switch Config Swit			
List1 List2 Priority* So 500 Image: Solution for Acure HCI Cluster Description Code configuration for Acure HCI Cluster Select Template* Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster Solution for Acure HCI Cluster <th>test test Priority* 00 90 0 1 12000 Description 0 Gos configuration for Azure HCI Cluster Setuct Template* witch freedom > Core Switch Freedom Config* Class-map type forderdage SRMA_CL_Maps_NetDogs match ogs groups 6 policy parts type forderdage SRMA_CL_Maps_NetDogs match ogs groups 6 policy parts type forderdage SRMA_CL_Maps_NetDogs policy parts type forderdage SRMA_CL_Maps_NetDogs policy forderdage SRMA_CL_Maps_NetDogs policy forderdage SRMA_CL_Maps_NetDogs policy forderdage SRMA_CL_Maps_NetDogs <t< th=""><th>Switch List:</th><th></th><th></th></t<></th>	test test Priority* 00 90 0 1 12000 Description 0 Gos configuration for Azure HCI Cluster Setuct Template* witch freedom > Core Switch Freedom Config* Class-map type forderdage SRMA_CL_Maps_NetDogs match ogs groups 6 policy parts type forderdage SRMA_CL_Maps_NetDogs match ogs groups 6 policy parts type forderdage SRMA_CL_Maps_NetDogs policy parts type forderdage SRMA_CL_Maps_NetDogs policy forderdage SRMA_CL_Maps_NetDogs policy forderdage SRMA_CL_Maps_NetDogs policy forderdage SRMA_CL_Maps_NetDogs <t< th=""><th>Switch List:</th><th></th><th></th></t<>	Switch List:		
Priority* 500 12000 Description CoS configuration for Azure HCI Cluster Select Template* switch.freeform 2 Select Template* switch.freeform 2 CoS configuration for Azure HCI Cluster Select Template* switch.freeform 2 CoS configuration for Azure HCI Cluster Select Template* switch.freeform 2 CoS configuration for Azure HCI Cluster Select Template* switch.freeform 2 CoS configuration for Azure HCI Cluster Select Template* switch.freeform 2 CoS configuration for Azure HCI Cluster Select Template* switch.freeform 2 CoS configuration for Azure HCI Cluster Communication for Azure HCI Cluster Select Template* switch.freeform 2 CoS configuration for Azure HCI Cluster Communication for Azure HCI Cluster Select Template* Select	Priority* S0 1-2000 Description Or Ginguration for Azure HCI Cluster Select Template* Selec	Leaf-1 Leaf-2		
500 Image: Control of the second	90 IP 1 1:2000 Description Ocs configuration for Azure HCI Cluster Select Template* which, freeform > Image: Configuration for Azure HCI Cluster Select Template* which, freeform > Image: Configuration for Azure HCI Cluster Select Template* which, freeform > Image: Configuration for Azure HCI Cluster Select Template* which, freeform Config* Class-map type network: ogo BDMA_CL_Map_MetQos match ogo: group 4 class-map type network: ogo Cluster-Comm, CL_Map_MetQos match ogo: group 3 page Select-Freeform Config* Class-map type ogg match-all BDMA match ogo: group 4 class map type ogg match-all BDMA match ogo: group 5 page Select-Freeform page Select-Freeform class map type ogg match-all BDMA match ogo: group 3 class map type ogg match-all BDMA match ogo: group 4 class type type ogg match-all BDMA page Strip page Strip page Strip class type ontextrese BDMA -	Priority*		
Benerghan Conservation Cons	I - 2000 Description Cos configuration for Azure HCI Cluster Select Template* switch_freeform > Comp Costend Freeform Config* Class-map type network:nego RBMA_CL_Map_NetQos match nos: group get match nos: group get class-map type network:nego RBMA_CL_Map_NetQos match nos: group RBMA_CL_Map_NetQos match nos: group RBMA_CL_Map_NetQos match nos: group RBMA_CL_Map_NetQos policy-map type network:nego Cluster-Comm_CL_Map_NetQos match nos: group RBMA_CL_Map_NetQos match nos: group RBMA_CL_Map_NetQos match nos: group RBMA_CL_Map_NetQos match nos: group RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_MAD_RBMA_RL_	500 0		
Description Descr	Description Ges configuration for Azure HCI Cluster Select Template* which freeform > Comp Comp Comp Comp Comp Comp Comp Com	1-2000		
Ges configuration for Azure HCI Cluster Salect Template* switch.freeform > Corosp Switch Freeform Config* class-map type network-ogs RDMA_CL_Map_MetQos match ogs group 4 class-map type gg match-all RDMA match ogs 4 policy-map type traiters-class type traiter-B policy-map type traiters-class type t	Ges configuration for Azure HCI Cluster Select Template* which freeform > Corp Select Template* Select Template* Corp Select Template* Class-map type network: one RDMA.CL_Map.MetQos match dos: snopp 4 Class-map type options: Select-Comm_CL_Map.MetQos match dos: snopp 4 Class-map type options: Selector-Comm_CL_Map.MetQos match dos: Snopp 4 Class-map type options: Selector-Comm_CL_Map.MetQos match dos: Snopp 4 Class-map type options: Selector-Site Cluster-Comm_CL_Map.MetQos match dos: Snopp 4 Class-map type options: Selector-Site Cluster-Comm_CL_Map.MetQos match dos: Snopp 4 Class-map type options: Selector-Site Cluster-Comm_CL_Map.MetQos match dos: Snopp 6 policy-map type retwork: selector-Site SNopA.CL_Map.MetQos match dos: Snopp 6 policy-map type retwork: selector-Site SNopA.CL_Map.MetQos match dos: Snopp 6 class type retwork: selector-Site SNopA.CL_Map.MetQos match dos: Snopp 6 policy class-default match dos: Snopp 6 class type retwork: selector-Site SNopA.CL_Map.MetQos	Description		
Select Template* switch_freedom > Select Template* switch_freedom > Select Template* switch_freedom Config* Select Template* Select Templat	Select Template* sextch_freedom > Settch Freedom Config* Calss-map type network:ops RDMA_CL_Map_NetQos match ops_group 4 class-map type network:ops RDMA_CL_Map_NetQos match ops_group 3 class-map type ogg match-all RDMA class-type network:ogg Class-class-default match ods: group diverse: gog All-all-all-RDMA class-type network:ogg class-default	QoS configuration for Azure HCI Cluster		
switch_freedom > Switch_freedom Config* Cass-map type nettype:sgn REMA_CL_Mep_NetQos match ope_group 4 Cass-map type nettype:sgn REMA_SL_Map_NetQos match ope_group 4 Cass-map type nettype:sgn REMA-all RDMA Cass type n	suitch_freedom > Settich_freedom Config* Celss-map type nativaris-aga_RDMA_CL_Map_NetIQog match ogs_group 5 class-map type nativaris-aga_RDMA_CL_Map_NetIQog match ogs_group 5 class-map type nativaris-aga_RDMA_CL_Map_NetIQog match ogs_group 5 policy-map type nativaris-aga_RDMA_CL_Map_NetIQog pass ag match-all CUSTER-COMM index for 5 policy-map type nativaris-aga_RDMA_CL_Map_NetIQog pass ag fictors 4 index for 5 policy-map type nativaris-aga_RDMA_CL_Map_NetIQog match ogs_group 5 class-type pass and nativaris-all CUSTER-COMM index for 5 policy-map type nativaris-aga_RDMA_CL_Map_NetIQog pass ag fictors 4 index for 5 policy-map type nativaris-aga_RDMA_CL_Map_NetIQog index type nativaris-aga	Select Template*		
Switch Freeform Config* descreape type optimized:coge REMA_CL_Map_NetCoge match ogg=rouge 4 class-map type optimized:coge REMA_CL_Map_NetCoge match ogg=rouge 4 class-map type optimized:coge REMA_CL_Map_NetCoge match ogg=rouge 4 class-map type optimized:coge REMA_CL_Map_NetCoge match ogg=rouge 4 class-type rettype:coge REMA_CL_Map_NetCoge mitil State or class-type rettype:coge Class-default Address static onfig Note At config should static match odg in the class state in the state or class state mitil State or class state mitil State or class state mitil State or class-default	Switch Freeform Config* Class-map type network-gage Match-LiMep_NetQog match dog-group 4 class-map type network-gage Cluster-Comm_CL_Mep_NetQog match dog-group 4 class-map type network-gage Match-all CLUSTER-COMM class-map type network-gage Match-all CLUSTER-COMM policy-map type network-gage Cluster-Comm_CL_Map_NetQogs policy-map type network-gage Cluster-Comm_CL_Map_NetQogs policy-map type network-gage Cluster-Comm_CL_Map_NetQogs class type policy-map type network-gage Cluster-Comm_CL_Map_NetQogs policy-map type network-gage Cluster-Comm_CL_Map_NetQogs policy-map type network-gage Cluster-Comm_CL_Map_NetQogs policy-map type network-gage Cluster-Comm_CL_Map_NetQogs class type policy-map type network-gage Cluster-Comm_CL_Map_NetQogs class type policy-figge Cluster-Comm_CL_Map_NetQogs class type network-gage class-default match dota	switch_freeform >		
Switch Freeform Config* Cates-map type retwork-coge Cluster-Comm_CL_Map_NetQos match oge=20004 class-map type oge match-all CDMA class-type network-coge Cluster-Comm_CL_Map_NetQos mit bare class-type network-coge Class-defaut class-type network-coge Class-de	Switch Freedom Config* class-map type network-gog Cluster-Comm_CL_Map_NetQog match gog grant-hall CLUSTER-COMM match gog grant-hall match gog grant-hall CLUSTER-COMM match gog	Group		
Switch Freeform Config*	Switch Freeform Config* Class-map type network-gog RDMA.CL_Map.MetQog match dog-group 4 class-map type network-gog RDMA.CL_Map.MetQog match dog-group 4 class-map type gg match-all RDMA cl_Map.MetQog class-map type gg match-all RDMA match dog-group cl_Map.MetQog class-map type gg match-all RDMA cl_Map.MetQog class-map type gg match-all RDMA match dog-group cl_Map.MetQog class-type network-ggg Class-default			
class-map type network-gog RDMA_CL_Map_NetQog match ogs_group 4 class-map type network-gog Cluster-Comm_CL_Map_NetQog match ogs_group 4 class-map type ogtsmatch-all RDMA match ogs 4 class-map type ogtsmatch-all CLUSTER-COMM match ogs 5 policy-map type fettroft-gog CQS_NETWORK uwspected dth-aurug express mats 210 class-map type network-gog Cluster-Comm_CL_Map_NetCos mits 2216 class type network-gog Cluster-Comm_CL_Map_NetCos	Citas-mat byte network: opg RUMA.CL_Mep_NetQog match ops_strong 4 citas-mat byte network: opg RUMA: citag RUMA citag opsilon citas-mat byte network: opg RUMA: citag RUMA citag opsilon citas-mat byte network: opg RUMA: citag RUMA citas-type network: opg RUMA: opg RUMA: citag RUMA citas-type network: opg RUMA: opg Citas-default mitu-RUMA citas-type network: opg RUMA: opg Citas-default mitu-RUMA citas-type network: opg RUMA: opg RUM	Switch Freeform Confin*		
match ogs-group 4 Address freg brogs class-map type ogs match-all CUSTER-COMM Address freg brogs class-map type ogs match-all CUSTER-COMM match ogs-group 5 class-map type ogs match-all CUSTER-COMM match ogs-group 6 class-type type ogs match-all CUSTER-COMM match ogs-group 6 class-type type ogs/match-all CUSTER-COMM match ogs-group 6 policy-map type ogtymetric-all CUSTER-COMM matogs-group 6 polic	match ogs-group 4 Addiend selor comm_CL_Map_MetQogs class-map type ogn match-all RDMA Addiend selor comfg, boet All confgs should strictly class-map type ogn match-all RDMA Addiend selor confg, boet All confgs should strictly class-map type ogn match-all RDMA match confg shoet All confgs should strictly class-map type ogn match-all RDMA match confg shoet All confgs should strictly policy-map type retwork.cgs RDMA. Confgs thould strictly match confg shoet All confgs shoet All confgs thould strictly policy-map type retwork.cgs RDMA. Confgs thould strictly match confg shoet All confgs thould strictly policy-map type retwork.cgs RDMA. Confgs thould strictly match confg shoet All confgs thould strictly pass and newlos. Sy manutation all public strictly match confg shoet All confgs thould strictly pass and previows.cgs RDMA. Confgs thould strictly mass thous and methors. Sy manutation all public strictly pass and previows.cgs RDMA. Confgs thould strictly mass thous and methors. Sy manutation all public strictly pass and previows.cgs RDMA. Confgs thould strictly mass thous and methods. Sy manutation all public strictly class type retwork.cgs RDMA. Confgs thould strictly mass thous and methods. Sy manutation all public strictly class type retwork.cgs RDMA. Confgs thould strictly masy method strictly	class-map type network-gos RDMA.CL_Map_NetQos		
Class-map bype optimics.gog Cluster-Comm_CL_Map_MetQog mithur Q3-Class-Default Address Lastin Config Note At configs should strictly match class class type restruction: gog match-all RDMA Address Lastin Config Note At configs should strictly match class type restruction: gog match-all RDMA Class-map type gog match-all RDMA Class type restruction: gog Class-default	Class-map type detwork-gos Cluster-Comm_CL_Map_MetX88 Additional waith coeffigs board at coeffics board at coeffigs board at c	match gos-group 4		
class-map type ogs match-sil BDMA Additional sells norm All configs troub All configs trou	class-map type ggs match-all BDMA Additional within config forer All configs from All configs match cost All costs Additional within config forer All configs from All config	class-map type network-gos Cluster-Comm_CL_Map_retubs		
match cos 4 Adstruit within certils Nort AL cortige used struty class-map type gogs match-all CLUSTER-COMM match how or dog ready should struty match cos 5 match-all CLUSTER-COMM match cos 5 gazes and readres. Any minutches will yield pclicy-map type instructives BLOALA_CLMap_MetQos unsepteted after auring depart match cos 5 gazes and readres. Any minutches will yield class type instructives BLOALA_CLMap_MetQos unsepteted after auring depart mittal 2716 class-type instructives:gos Class-default	match cos 4 Addisore atoto confg. Notel A confga handa storiy classe mo by a gog match-all CLUSTER-COMM Addisore atoto confg. Notel A confga handa storiy classe mo by a gog match-all CLUSTER-COMM match confg. Notel A confga handa storiy policy-map type network.cogo QOS_NETWORK spose and newlos. Any meantons at yead policy-map type network.cogo QOS_NETWORK spose and newlos. Any meantons at yead policy-map type network.cogo QOS_NETWORK unequested afth averag deety. policy-map type network.cogo Classe-Comm_CL_Map_NetGogs unequested afth averag deety. policy-map type network.cogo class-default match coals.cogo network.cogo class-default	class-map type gos match-all RDMA		
match dos 5 mitch bitward dolp, 4th respect to see, lawing policy-map type network-gos QOS_NETWORK space and reviews. A mitch bitward dolp, 4th respect to see, lawing class type network-gos QOS_NETWORK space and reviews. A mitch bitward dolp, 4th respect to see, lawing class type network-gos QOS_NETWORK unequested dth, during depty. class type network-gos QOS_NETWORK space pic-cos 4 class type network-gos QOS_NETWORK unequested dth, during depty. class type network-gos Class-default class type network-gos Class-default	match os 5	match cos 4 class-man tune nos match-all CLUSTER-COMM	Additional switch config. Note! All configs should strictly	
policy-map type <u>retwork; cop</u> 00,5.NETWORK update task type <u>retwork; cop</u> 00,5.NETWORK update page 2016 clask type <u>retwork; cop</u> 0,0.1.4.1.4.0.0.4.1.4.0.0.0.4.1.4.0.0.0.4.0.0.0.0	policy-map type network.cog QQS_NETWORK pass type network.cog QQS_NETWORK	match cos 5	match 'show run' output, with respect to case, leading	
mty 9218 Class type network:gos Cluster-Comm_CL_Msp_NetQos mty 9218 Class type network:gos class-default	mtu g216 Class type retetvolt: gos Cluster-Comm_CL_Map_NetQos mtu g216 Class type retetvolt: gos class-default	policy-map type network-gos QOS, NETWORK class type network-gos RDMA_CL_Map_NgQos pause pro-cos 4	spaces and newlines. Any manuacties we yeed unexpected diffs during deploy.	
mty 9216 Class type network-gos class-default	mtu 2016 disa type politikoti: gos class-default mtu 2018	mtu 9216 class type network-gos Cluster-Comm_CL_Map_NetQos		
mtu 0216	mth: 0218	mtu 9216 class type network-gos class-default		
		mti 9216		

Click on **Save** and you would be returned to **Policy** tab. From Policy tab page select the policy just created and click on **Push** button from **Actions** drop-down to deploy generated config to the Leaf switches

Step 2: Apply the QoS policy on the Peer-link of Leaf switches (connecting to Azure HCI).

This is required to apply QoS on any traffic which may pass over the peer-link.

From Fabric **Overview** > **Interfaces** tab, select the peer-link port-channel interfaces for Leaf-1 and Leaf-2 and click on **Edit** from **Actions** drop-down.

Fa	bric C	verview - Azure	e-HCI													Actions	• 🔿	?	_	\times
Ov	erviev	/ Switches I	Links	Interfaces	Interface Grou	os Policies	Networks	VRFs	Services	Event Analytics	History	Resources	Virtual	Infrastructure						
	Interf	ace contains 500 >	×													Edit	Clear All		ctions	^
	~	Device Name		Interface		Admin Status	Oper. Status	Reason		Policies		Overlay N	etwork	Sync Status	Interface Group	Por Cha	Create In Create St	terfac ubinte	e rface	
		Leaf-1		Port-chann	el500	↑ Up	↑ Up	ok		int_vpc_peer_link_p	0	NA		In-Sync			Edit			
		Leaf-2		Port-chann	el500	↑ Up	1 Up	ok		int_vpc_peer_link_pe	0	NA		In-Sync			Normalize Multi-Atte	e ach		

Edit interface(s)			
	1 of 2 Coloridad Interference) -		
	1 of 2 Selected Interface(s) :		
	Interface		
	Leaf-1 : Port-channel500		
	Policy*		
	int_vpc_peer_link_po >		
	Policy Options		
	VPC Peer-Link Port-Channel Member Interfaces		
	Ethernet1/39,Ethernet1/40	A list of member interfaces (e.g. et/5,eth//7-9)	
	vPC Peer-link Trunk Allowed Vians		
	Select an Option	VPC Peer-link Allowed Vian list (empty=all or none)	
	Native Vlan	VLAN ID to set as the interface native vian	
	Port Channel Description		
		Add description to the port-channel (Max Size 254)	
	Members Description		
		Add description, if members don't have any (same for all members, Max Size 254)	
	Port Channel Admin State*	Admin state of the port-channel	
	Freeform Config		
	service-policy type gos input AzS_HCI_QoS		
			Additional CLI for the interface

Click on Save button for Leaf-1.

Click on **Next** button and repeat the same step for vPC peer-link of Leaf-2.

Verify the pending configuration and deploy.

Pending config	Р	ending config
Azure-HCI > Leaf-1 > Port-channel500	Az	zure-HCl > Leaf-2 > Port-channel500
<pre>1 interface port-channel500 2 switchport 3 switchport mode trunk 4 spanning-tree port type network 5 description "vpc-peer-link Leaf-1Leaf 6 no shutdown 7 service-policy type qos input AzS_HCI_C 8 configure terminal 9 </pre>	:-2" IOS	<pre>1 interface port-channel500 2 switchport 3 switchport mode trunk 4 spanning-tree port type network 5 description "vpc-peer-link Leaf-2Leaf-1" 6 no shutdown 7 service-policy type qos input AzS_HCI_QoS 8 configure terminal 9 </pre>

Step 3: Apply the QoS policy on Leaf switch interfaces which are used to connect to Azure HCI.

Cisco NDFC allows grouping the interfaces using Interface Groups. All the interfaces which require identical configuration can be grouped together using an Interface Group and all the required configuration is applied only to the Interface Group.

Although Leaf-1 and Leaf-2 interfaces connecting to Azure Stack HCl server require same QoS configuration, they would be carrying different VLANs for RDMA traffic (Leaf-1 for Storage-A and Leaf-2 for Storage-B) therefore two separate Interface Groups are required.

	erview - Azure-HCI								Action	••• • • • • • • • • • • • • • • • • •
erview	Switches Links I	nterfaces Interface Gro	ups Policies	Networks	VRFs Services	Event Analytics History	Resources Virtual	Infrastructure		
Descripti	tion contains AzS $ imes$								Edi	t Clear All Actions ^
🚍 De	evice Name	Interface	Admin Status	Oper. Status	Reason	Policies	Overlay Network	Sync Status	Interface Port Group Channel I	Create Interface Create Subinterface
🔽 Les	af-1	Ethernet1/11	↑ Up	U Down	XCVR not inserted	int_trunk_host	NA	In-Sync		Edit
🔽 Lei	af-1	Ethernet1/12	↑ Up	U Down	XCVR not inserted	int_trunk_host	NA	In-Sync		Normalize Multi-Attach
Lei	af-2	Ethernet1/11	↑ Up	🕹 Down	XCVR not inserted	int_trunk_host	NA	In-Sync		Multi-Detach
Les	af-2	Ethernet1/12	↑ Up	🕹 Down	XCVR not inserted	int_trunk_host	NA	In-Sync		Deploy
										No Shutdown
								Add to	o Interface Group	More >

Ports Eth1/11-12 are added to Leaf-1_Azure_HCl_Server_ports Interface Group with following settings:

- Set Interface Type: Ethernet
- Policy: int_ethernet_trunk_host
- Enable BPDU Guard: True
- Enable Port Type Fast: Yes
- MTU: Jumbo (9216 bytes)
- Native VLAN: Can be set to Mgmt Vlan (Optional)
- Freefrom Config: Provide service-policy CLI command to apply QoS and Queuing policies and CLI command to enable Policy Flow Control to the interfaces

ate Interface Group	
abric Name* zure-HCI teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_ports teaf-Lazure,HCL_Server_p	
Enable BPDU Guard* true	Enable spendag-twe tipdogoant trave leader, telectificater, not etime to default settings'
IG for Fex Ports*	Shared group for fee parts
Enable Port Type Fast*	Fasilis epsaving-isas odga part bekaver
MTU*	MTU for the interface
SPEED*	
	Interface Speed
on v	Auto Negotiate mode for speed
Trunk Allowed Vlans* none	Allowed values: hone, 'wil, or vian ranges (ec. 1-200,500-2000,3000)
Native Vlan	
	Set native VLAN for the Hisrface
Enable W/G Urphan Port	If enabled, configure the interface as a VPC orphan port to be suspended by the secondary pear in VPC failures
Freeform Config	
priority-flow-control made on service-policy type gos input <u>ArS_HOL_QoS</u> service-policy type queuing output <u>QOS</u> _EGRESS_PORT	J

Repeat the above steps for adding Leaf-2 ports Eth1/11-12 to Leaf-2_Azure_HCI_Server_ports Interface Group -

	abric Overview - Azure-HCI											Actio	ns v 🔿 (? — ×
0	verview Switches Links	Interfaces Interface G	roups Policie	es Network	s VRFs Service	s Event Analytics History	Resources Virtua	I Infrastructure						
ľ	Description contains AzS \times											E	lit Clear All	Actions v
	Device Name	Interface	Admin Status	Oper. Status	Reason	Policies	Overlay Network	Sync Status	Interface Group	Port Channel ID	vPC ld	Speed	MTU	Mode
	Leaf-1	Ethernet1/11	↑ Up	🕹 Down	XCVR not inserted	int_shared_trunk_host	NA	In-Sync	Leaf-1_Azure_HCI_Server_ports			25Gb	9216	trunk
	Loaf-1	Ethernet1/12	↑ up	U Down	XCVR not inserted	int_shared_trunk_host	NA	In-Sync	Leaf-1_Azure_HCI_Server_ports	J		25Gb	9216	trunk
	Leaf-2	Ethernet1/11	↑ Up	J Down	XCVR not inserted	int_shared_trunk_host	NA	In-Sync	Leaf-2_Azure_HCI_Server_ports			25Gb	9216	trunk
	Leaf-2	Ethernet1/12	↑ Up	U Down	XCVR not inserted	int_shared_trunk_host	NA	In-Sync	Leaf-2_Azure_HCI_Server_ports)		25Gb	9216	trunk
1														

Now we have enabled PFC and applied QoS and Queuing policies on Leaf-1 & Leaf-2 respective interfaces. We'll now create the networks (Vlans) required for Azure Stack HCl in next section.

Configure LLDP

Cisco NDFC enables the LLDP feature on all the devices in the VXLAN fabric and LLDP is enabled on all the interfaces on all devices. However, LLDP is not enabled by Cisco NDFC for traditional classic LAN fabrics. For traditional classic LAN fabrics, the _IIdp policy feature must be associated to the Leaf switches for LLDP support.

Configure Networks for Azure Stack HCI

Following are the network requirements for Azure Stack HCI:

- Two Layer-3 networks with Anycast Gateway configured on the leafs
- Two Layer-2 networks for Storage (one for each leaf)



Figure 23. Cisco NX-OS based networks for Azure Stack HCI

On VXLAN fabric all the Layer-3 networks need to be mapped to a VRF which provides isolation between any two tenants. All the networks pertaining to a tenant are mapped to the respective tenant VRF. Layer-2 networks do not need to be mapped to VRF.

To create VRF, go to **Fabric Detailed View > VRF > Actions** and choose **Create VRF** and provide following parameters:

- VRF Name: Azure_Tenant_VRF_50000
- VRF ID: provide VNI for VRF
- VLAN ID: provide Vlan for VRF
- VRF VLAN Name: provide name for the VLAN (optional)

reate VRF	
VRF Name*	
Azure_Tenant_VRF_50000	
VRF ID• 50000	
VLAN ID	use VI AN
Prop	USS VLAN
VRF Template*	
Default_VRF_Universal >	
VRF Extension Template*	
Default VRF Extension Universal N	
· · · · · · · · · · · · · · · · · · ·	
General Parameters Advanced Route Target	
VRF VLAN Name	
Azure_Tenant_VRF_Vlan	If > 32 chars, enable 'system vian long-name' for N
VRF Interface Description	
VRF Description	

Once the VRF is created, Networks can be created. To create Networks, go to **Fabric Detailed View >>** Network >> Actions and choose **Create Network**.

Let's create Layer-3 network used for management of Azure HCI Stack recourses with following parameters:

Network Name - Azure_Mgmt_Network_30000

- VRF Name provide Azure_Tenant_VRF_50000
- Network ID 30000
- VLAN ID 2300
- IPv4 Gateway/Netmask 172.16.10.1/24
- VLAN Name Azure_Mgmt Vlan
- MTU for L3 Interface 9216 bytes

Network Name*	
Azure_Mgmt_Network_30000	
Layer 2 Only	
VPE Nome*	
Azure_Tenant_VRF_50000 X V	Create VRF
Network ID*	
30000	
VLAN ID	
2300	Propose VLAN
Network Template*	
Default Network Universal N	
Delution Control of y	
Network Extension Template*	
Default_Network_Extension_Universal >	
Generate Multicast IP Please click only to gene	rate a New Multicast Group address and override the default value!
General Parameters Advanced	
IPv4 Gateway/NetMask	
IPv4 Gateway/NetMask 172.16.10.1/24	example 192.0.2.7/2.6
IPv4 Gateway/NetMask 172.16.10.1/24 IPv6 Gateway/Prefix List	example 192.0.2.724
IPv4 Gateway/NetMask 172.16.10.1/24 IPv6 Gateway/Prefix List	example 192.0.2.1/24
IPv4 Gateway/NetMask 172.16.10.1/24 IPv6 Gateway/Prefix List	example 192.0.2.1/24 example 2001.db8:1/64.2001.db8:1/64
IPv4 Gateway/NetMask 172.16.10.1/24 IPv6 Gateway/Prefix List VLAN Name	example 192.0.2.1/24 example 2001.db8:1/64.2001.db8:1/64
IPv4 Gateway/NetMask 172.16.10.1/24 IPv6 Gateway/Prefix List VLAN Name Azure_Mgmt_Vlan	example 192.0.2.1/24 example 2001.db8:1/64.2001.db8:1/64 If > 32 chars, enzide 'system vian long-name' for NK-OS, disable VTPv1 and VTPv2 or switch to VT
IPv4 Gateway/NetMask 172.16.10.1/24 IPv6 Gateway/Prefix List VLAN Name Azure_Mgmt_Vlan Interface Description	example 192.0.2.1/24 example 2001.db8:1/64.2001.db8:1/64 If > 32 chers, enetide 'system vian long-name' for NK-OG, disable VTPv1 and VTPv2 or switch to VT
IPv4 Gateway/NetMask 172.16.10.1/24 IPv6 Gateway/Prefix List VLAN Name Azure_Mgmt_Vlan Interface Description	example 192.0.2.1/24 example 2001.db8:1/64.2001.db8:1/64 If > 32 chars, energie 'system vian long-name' for NK-OG, disable VTPv1 and VTPv2 or switch to VT
IPv4 Gateway/NetMask 172.16.10.1/24 IPv6 Gateway/Prefix List VLAN Name Azure_Mgmt_Vlan Interface Description	example 192.0.2.1/24 example 2001.db8:1/64.2001.db8:1/64 If > 32 chars, energie 'system vian long-name' for NK-OG, disable VTPv1 and VTPv2 or switch to VT
IPv4 Gateway/NetMask 172.16.10.1/24 IPv8 Gateway/Prefix List VLAN Name Azure_Mgmt_Vlan Interface Description MTU for L3 interface	example 192.0.2.5/24 example 2001.db8:1/64.2001.db8:1/64 If > 32 chers, enable 'system vian long-name' for NK-OS, disable VTPv1 and VTPv2 or switch to VT

Let's create second Layer-3 network used for Azure HCI Stack Tenants:

- Network Name: Tenant_Network_30001
- VRF Name: Azure_Tenant_VRF_50000
- Network ID: 30001
- VLAN ID: 2301
- IPv4 Gateway/Netmask: 172.16.20.1/24
- VLAN Name: Tenant_Network_Vlan
- MTU for L3 Interface: 9216 bytes

ate Network	
Network Name*	
Tenant_Network_30001	
aver 2 Only	
VRF Name*	
Azure_Tenant_VRF_50000 × V	Create VRF
Natwork ID*	
30001	
2201	Dranozo VI AN
2301	PTOPOSE VLAN
Default_Network_Universal >	
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Generate Multicast IP Please click only to generate a	New Multicast Group address and override the default value!
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Generate Multicast IP Please click only to generate a General Parameters Advanced	New Multicast Group address and override the default value!
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Generate Multicast IP Please click only to generate a General Parameters Advanced IPv4 Gatewaw/MetMask	New Multicast Group address and override the default value!
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Generate Multicast IP Please click only to generate a General Parameters Advanced IPv4 Gateway/NetMask 172.16.20.1/24	New Multicast Group address and override the default value!
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Cenerate Multicess IP Please click only to generate a General Parameters Advanced IPv4 Gateway/NetMask 172.16.20.1/24	New Multicast Group address and override the default value!
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Generate Multicast IP Please click only to generate a General Parameters Advanced IPv4 Gateway/NetMask 172.16.20.1/24 IPv6 Gateway/Prefix List	example 192.0.2.1/24
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Cenerate Multicest IP Please click only to generate a General Parameters Advanced IPv4 Gateway/NetMask 172.16.20.1/24 IPv6 Gateway/Prefix List	New Multicast Group address and override the default valuel example 192.0.2.1/24 example 2001-ab8::1/64.2001-db9::1/64
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Cenerate Multicast IP Please click only to generate a General Parameters Advanced IPv4 Gateway/NetMask IT22.16.20.1/24 IPv6 Gateway/Prefix List VLAN Name	New Multicast Group address and override the default valuel example 192.0.2.1/24 example 2001-stell::1/64.2001-db9::1/64
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Cenerate Multicest IP Please click only to generate a General Parameters Advanced IPv4 Gateway/NetMask 172.16.20.1/24 IPv6 Gateway/Prefix List VLAN Name Tenant_Network_Vlan	New Multicast Group address and override the default value! example 192.0.2.1/24 example 2001.db8::1/84,2001.db9::V64 If > 32 chars, enable 'system vian long-name' for NX-OS, disable VTPv1 and VTPv2 or switch to VTPv3 for IOS X
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Cenerate Multicest IP Please click only to generate a General Parameters Advanced IPv4 Gateway/NetMask 172.16.20.1/24 IPv6 Gateway/Prefix List VLAN Name Tenant_Network_Vian Interface Description	New Multicast Group address and override the default value! example 192.0.2.124 example 2001:db8::1/84,2001:db9::1/84 if > 32 chars, enable 'system vian long name' for NX-OS, disable VTPv1 and VTPv2 or switch to VTPv3 for IOS X
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Cenerate Multicest IP Please click only to generate a General Parameters Advanced IPv4 Gateway/NetMask 172.16.20.1/24 IPv6 Gateway/Prefix List ULAN Name Tenant_Network_Vian Interface Description	New Multicast Group address and override the default valuel example 192.0.2.1/24 example 2001:stb8::1/84,2001:stb9::1/84 if > 32 chara, enable 'system vian long-name' for NK-DS, disable VTPv1 and VTPv2 or switch to VTPv3 for IOS X
Default_Network_Universal > Network Extension Template* Default_Network_Extension_Universal > Cenerate Multicast IP Please click only to generate a Ceneral Parameters Advanced IPv4 Gateway/NetMask 172.16.20.1/24 IPv6 Gateway/Prefix List ULAN Name Tenant_Network_Vlan Interface Description MTU for L3 interface	New Multicast Group address and override the default valuel example 192.0.2.1/24 example 2001:stb8::1/84,2001:stb9::1/84 If > 32 chars, enable "system vian long-name" for NX-DS, disable VTPv1 and VTPv2 or switch to VTPv3 for IOS X

Now, we will create Layer-2 networks for Storage. Unlike the L3 networks, L2 networks don't have any SVI and does not require mapping to VRF. To create L2 network, check **Layer 2 Only** check box.

Create L2 network for Storage-A with the following parameters:

- Network Name: Storage-A_30100
- Network ID: 30100
- VLAN ID: 2400
- VLAN Name: Storage-A_Vlan

Network Name*		
Storage-A_Network,	.30100	
Laura A Calu		
Layer 2 Only		
VRF Name*		
NA	\sim	
Network ID*		
30100	U U	
VLAN ID		
2400	0	Propose VLAN
Default_Network_Univ Network Extension Ter Default_Network_Exter	ersal > nplate* nsion Universal >	
Defauit_Network_Univ Network Extension Ter Defauit_Network_Exter Generate Multicast IP	ersal > mplate* nsion_Universal > Please click only to generat	e a New Multicast Group address and override the default value!
Default_Network_Univ Network Extension Ter Default_Network_Exter Generate Multicast IP	ersal > mplate* nsion_Universal > Please click only to generat	e a New Multicast Group address and override the default value!
Default_Network_Univ Network Extension Ter Default_Network_Exter Generate Multicast IP General Parameter IPv4 Gateway/NetM	ersal > mplate* nsion_Universal > Please click only to generat Advanced ask	e a New Multicast Group address and override the default value!
Default_Network_Univ Network Extension Ter Default_Network_Exter Generate Multicast IP General Parameters IPv4 Gateway/NetM	ersal > mplate* nsion_Universal > Please click only to generat Advanced ask	e a New Multicast Group address and override the default value!
Default_Network_Univ Network Extension Ter Default_Network_Exter Generate Multicast IP General Parameter IPv4 Gateway/NetM IPv6 Gateway/Prefit	ersal > mplate* nsion_Universal > Please click only to generat Advanced ask List	e a New Multicast Group address and override the default value!
Default_Network_Univ Network Extension Ter Default_Network_Exter Generate Multicast IP General Parameter IPv4 Gateway/NetM IPv6 Gateway/Prefix	ersal > mplate* nsion_Universal > Please click only to generat Advanced ask List	e a New Multicast Group address and override the default value! example 192.0.2.1/24 example 192.0.2.1/24 example 2001ub8=1/64,2001ub8=1/64
Default_Network_Univ Network Extension Ter Default_Network_Exter General Parameter IPv4 Gateway/NetM	ersal > mplate* mplate* Please click only to generat Advanced ask List	e a New Multicast Group address and override the default value! example 192.0.2.1/24 example 2001.stb8-1/64,2001.stb8-1/64
Default_Network_Univ Network Extension Ter Default_Network_Exter Generate Multicast IP General Parameter IPv4 Gateway/NetM IPv6 Gateway/Prefix VLAN Name	ersal > mplate* sion_Universal > Please click only to generat Advanced ask List	e a New Multicast Group address and override the default value! example 192.0.2.1/24 example 2001 ub8-1/64,2001.ub8-1/64
Default_Network_Univ Network Extension Ter Default_Network_Exten General Parameter: IPv4 Gateway/NetM IPv6 Gateway/Prefix VLAN Name Storage-A_Vlan	ersal > mplate* msion_Universal > Please click only to generat Advanced ask List	e a New Multicast Group address and override the default value! example 182.0.2.1/24 example 2001.scb=1/e4,2001.scb=1/e4 If > 32 chars, enable isystem vian long-name' for NK-OS, disable VTPv1 and VTPv2 or switch
Default_Network_Univ Network Extension Ter Default_Network_Exten General Parameter: IPv4 Gateway/NetM IPv6 Gateway/Prefix VLAN Name Storage-A_Vlan Interface Descriptio	rsai > nplate* nsion_Universal > Please click only to generat Advanced ask List n	e a New Multicast Group address and override the default value! example 182.0.2.1/24 example 2001.scb=1/e4,2001.scb=1/e4 If > 32 chars, enable 'system vian long-name' for NX-OS, disable VTPv1 and VTPv2 or switc
Default_Network_Univ Network Extension Ter Default_Network_Exter General Parameter: IPv4 Gateway/NetM IPv6 Gateway/Prefix VLAN Name Storage-A_Vlan Interface Descriptio	rsai > nplate* nsion_Universal > Please click only to generat Advanced ask List n	e a New Multicast Group address and override the default value! example 182.0.2.1/24 example 2001.sb8=1/64,2001.sb8=1/64 If > 32 chars, enable 'system vian long-name' for NK-OS, disable VTPv1 and VTPv2 or switc

Create L2 network for Storage-B with the following parameters:

- Network Name Storage-B_30101
- Network ID 30101
- VLAN ID 2401
- VLAN Name Storage-B_Vlan

Network Name*			
Storage-B_Networ	k_30101		
Value 2 Only			
VPE Name*			
NA		Create VPE	
	· · · · · · · · · · · · · · · · · · ·) Create Vill	
Network ID*			
30101	0		
VLAN ID			
2401	0	Propose V	ILAN
Network Template*			
Default_Network_Un	iversal >		
Default_Network_Un	versal >		
Default_Network_Un Network Extension T	versal >		
Default_Network_Un Network Extension T Default_Network_Ext	iversal > emplate* ension_Universal >		
Default_Network_Un Network Extension T Default_Network_Ext Generate Multicast IF	iversal > emplate* ension_Universal > Please click only to gene	rate a New Mul	ticast Group address and override the default value!
Default_Network_Un Network Extension T Default_Network_Ext Generate Multicast If	versal > emplate* ension_Universal > Please click only to gene	rate a New Mul	ticast Group address and override the default value!
Default_Network_Un Network Extension 1 Default_Network_Ext Generate Multicast If General Parameter	versal > emplate* ension_Universal > Please click only to gene prs Advanced	rate a New Mul	ticast Group address and override the default value!
Default_Network_Un Network Extension 1 Default_Network_Ext Generate Multicast If General Parameter	versal > emplate* ension_Universal > Please click only to gene prs Advanced	rate a New Mul	ticast Group address and override the default value!
Default_Network_Un Network Extension 1 Default_Network_Ext Generate Multicest If General Paramete	versal > emplate* ension_Universal > Please click only to gene pre Advanced Mask	rate a New Mul	ticast Group address and override the default value!
Default_Network_Un Network Extension 1 Default_Network_Ext Generato Multicast IF General Paramete IPv4 Gateway/Net	versal > emplate* ension_Universal > Please click only to gene ers Advanced Mask	rate a New Mu	ticast Group address and override the default value!
Default_Network_Un Network Extension 1 Default_Network_Exc General Multicast IF General Parameter IPv4 Gateway/Net	versal > emplate* ension_Universal > Please click only to gene ers Advanced Mask fix List	rate a New Mul	ticast Group address and override the default value! example 192.0.2.1/24
Default_Network_Un Network Extension 1 Default_Network_Exc Generate Multicast IF General Parameter IPv4 Gateway/Net IPv6 Gateway/Pre	versal > emplate* ension_Universal > Please click only to gene ers Advanced Mask fix List	rate a New Mul	example 192.0.2.1/24
Default_Network_Un Network Extension 1 Default_Network_Ext General Paramete IPv4 Gateway/Net IPv6 Gateway/Pre	versal > emplate* ension_Universal > Please click only to gene ers Advanced Mask fix List	rate a New Mu	ticast Group address and override the default value! example 192.0.2.1/24 example 2001.dtb8::1/64.2001.dtb9::1/64
Default_Network_Un Network_Extension 1 Default_Network_Ext General Parameter IPv4 Gateway/Net IPv6 Gateway/Pre VLAN Name	versal > emplate* ension_Universal > Please click only to gene ors Advanced Mask	rate a New Mu	ticast Group address and override the default value! example 192.0.2.1/24 example 2001.db8=1/64,2001.db9=1/64
Default_Network_Un Network Extension 1 Default_Network_Ex General Paramete IPv4 Gateway/Net IPv6 Gateway/Pre VLAN Name Storage-B_Vlan	versal > emplate* ension_Universal > Please click only to gene ors Advanced Mask fix List	rate a New Mu	ticast Group address and override the default value! example 192.0.2.1/24 example 2001:db8:1/84,2001:db9:1/64 # > 32 chars, enable 'system vian long-name' for NX-OS, disable VTPv1 and VTPv2 or selich to V
Default_Network_Un Network_Exc General Paramete IPv4 Gateway/Net IPv6 Gateway/Pre VLAN Name Storage-B_Vlan Interface Decision	versal > emplate* ension_Universal > Please click only to gene rs Advanced Mask fix List	rate a New Mu	ticast Group address and override the default value! example 192.0.2.1/24 example 2001db8::V64,2001db9::1/64 if > 32 chars, enable 'system vian long-name' for NK-OS, disable VTPv1 and VTPv2 or switch to V
Default_Network_Un Network_Exc Generate Multicast II Generate Multicast II IPv4 Gateway/Net IPv6 Gateway/Pre VLAN Name Storage-B_Vlan Interface Descript	versal > emplate* ension_Universal > Please click only to gene rs Advanced Mask fix List ion	rate a New Mu	ticast Group address and override the default value! example 192.0.2.1/24 example 2001.db8::1/64,2001.db9::1/64 If > 32 chars, enable 'system vian long-name' for NX-OS, disable VTPv1 and VTPv2 or switch to 1
Default_Network_Un Network Extension 1 Default_Network_Exc General Multicast I General Parameto IPv4 Gateway/Net IPv6 Gateway/Pre VLAN Name Storage-B_Vlan Interface Descript	versa > emplate* ension_Universal > Please click only to gene re Advanced Mask fix List ion	rate a New Mu	ticast Group address and override the default value! example 192.0.2.1/24 example 2001.eb8=1/64,2001.db9=1/64 if > 32 chars, enable 'system vian long-name' for NK-OS, disable VTPv1 and VTPv2 or switch to V
Default_Network_Un Network Extension 1 Default_Network_Exc Generale Multicast II Generale Multicast II IPv4 Gateway/Net IPv4 Gateway/Net VLAN Name Storage-B_Vlan Interface Descript MTU for L3 Interfa	versal > emplate* ension_Universal > Please click only to gene re Advanced Mask fix List ion ice	rate a New Mul	ticast Group address and override the default value! example 192.0.2.1/24 example 2001.db8=1/64,2001.db9=1/64 If > 32 chars, enable 'system vian long-name' for NK-OS, disable VTPv1 and VTPv2 or switch to V

We can verify all the networks from Networks tab of the fabric -

Fabri	Fabric Overview - Azure-HCI										
Overv	Dverview Switches Links Interfaces Interface Groups Policies Networks VRFs Services Event Analytics History Resources Virtual Infrastructure										
	Etrac bu shtibutas										
<u> </u>	Network Name	Network ID .	VRF Name		IPv4 Gateway/Prefix	IPv6 Gateway/Prefix	Network Status	VLANID			
C	Azure_Mgmt_Network_30000	30000	Azure_Tenant_VR	F_50000	172.16.10.1/24		INA INA	2300			
C	Tenant_Network_30001	30001	Azure_Tenant_VR	F_50000	172.16.20.1/24		NA	2301			
C	Storage-A_Network_30100	30100	NA				INA INA	2400			
C	Storage-B_Network_30101	30101	NA				INA INA	2401			

Next, we attach the networks to the interfaces, select the networks to be attached and click **Actions** >> **Attach to Interface Group**. We have attached Azure_Mgmt and Tenant networks to both the Leafs however Storage networks are attached to the respective switches.

Fabric Overview - Azure-HCI						Actions - 🔿 ? —			
Drevriew Switches Links Interfaces Interface Groups Policies Metworks VRFs Services Event Analytics History Resources Virtual Infrastructure									
Filter by attributes						Actions			
Network Name	Network ID 🙏	VRF Name	IPv4 Gateway/Prefix IPv8 Gateway/Prefix	Network Status	VLAN ID	Interface Oroup			
Azure_Mgmt_Network_30000	30000	Azure_Tenant_VRF_50000	172.10.10.1/24	O DE PLOYED	2300	Leaf-1_Azure_HCI_Server_ports,Leaf-2_Azure_HCI_Server_ports			
Tenant_Network_30001	30001	Azure_Tenant_VRF_50000	172.16.20.1/24	DEPLOYED	2301	Leaf-1_Azure_HCI_Server_ports,Leaf-2_Azure_HCI_Server_ports			
Storage-A_Network_30100	30100	NA		DEPLOYED	2400	Leaf-1_Azure_HCLServer_ports			
Storage-B_Network_30101	30101	NA		DEPLOYED	2401	Leaf-2_Azure_HCLServer_ports			

Once all the networks are attached, select the networks and click on **Actions** > **Deploy** for NDFC to generate and push the config to the devices.

Build External Connectivity for Azure Stack HCI servers

Any network outside of VXLAN fabric is referred as external, to provide connectivity to such networks VRF_Lite (MPLS Option A) is used. Cisco NDFC supports full automation for extending connectivity to external networks from a VXLAN or Traditional Classical LAN fabric.

VXLAN devices which perform IPv4/IPv6 handoff are referred as Border devices this role is also supported in Cisco NDFC. Once the Tenant VRF is deployed on the border devices it can be further extended towards external networks.

Following NDFC settings are required under **Resources** tab of the fabric template for setting up external connectivity for VXLAN fabric.

VRF Lite Deployment*	
Back2Back&ToExternal ~	VRF Lite Inter-Fabric Connection Deployment Options. If 'Back2Back&ToExternal' is selected, VRF Lite IFCs are auto
	created between border devices of two Easy Fabrics, and between border devices in Easy Fabric and edge routers in
	External Fabric. The IP address is taken from the VRF Lite Subnet IP Range' pool.
Auto Deploy for Peer	Whether to auto generate VRF LITE sub-interface and 8GP peering configuration on managed neighbor devices. If set, auto created VRF LITe IFC links with have 'Auto Depioy for Peer' enabled.
Auto Deploy Default VRF	
	Whether to auto generate Default VRF interface and BGP peering configuration on VRF LITE IFC auto deployment. If
	set, auto created VRF Lite IFC links will have 'Auto Deploy Default VRF' enabled.
Auto Deploy Default VRF for Peer	Whether to auto generate Default VRF Interface and BGP peering configuration on managed neighbor devices. If set, auto created VRF Life IFC links will have 'Auto Deploy Default VRF for Peer' enabled.
Redistribute BGP Route-map Name	Route Map used to redistribute BOP routes to IOP in default vrf in auto created VRF Lite IFC links
VRF Lite Subnet IP Range* 10.33.0.0/16 VDE Lite Subnet Markt*	Address range to assign P2P Interfabric Connections
30	(Mer.8, Max.31)

Change VRF Lite IP Subnet range and subnet mask (if required), if required.

Before you start make sure, border devices have the VRF deployed. If not, attach the VRF to the border devices.

To configure the VRF_Lite extension, select the required VRF and go to the VRF detailed view from VXLAN fabric. Under **VRF Attachments** tab, select the border devices and click on **Edit** from **Actions** drop-down -

RF Overview - Azure_Tenant_VRF_50000									Actions ~	Refresh —
verview VRF Attachments Networks										
Filter by attributes										Actions
VRF Name VRF ID	VLAN ID	Switch	Status	Attachment	Switch Role	Fabric Name	Loopback ID	Loopback IPV4 Addres	s Loopba	History
Azure_Tenant_VR 50000	2000	Leaf-2	DEPLOYED	Attached	leaf	Azure-HCI				Edit
Azure_Tenant_VR 50000	2000	Leaf-1	DEPLOYED	Attached	leaf	Azure-HCI				Preview
Azure_Tenant_VR 50000	2000	Leaf-5	DEPLOYED	Attached	border	Azure-HCI				Import
Azure_Tenant_VR 50000	2000	Leaf-6	O DEPLOYED	Attached	border	Azure-HCI				Export
										Quick Attach
										Quick Detach

For each border device select **VRF_LITE** from drop-down under **Extend** and click on **Attach-All** button. Additional parameters can be provided by clicking on **Exit** link under **Action**.

Edit VBF Attachment - Azure_Tenant_VBF_50000	? – ×
1 of 2 - Azura_Tenant_VRF.50000 - Leaf-5(FD027280TNL)	
Lust-Appoorzerhaj Desan Carlos X. An 2000 B Desart Verdure X.	
CL Prevelow Config Easy 2 All origing already already match the thorn our output, including cases and new line Any symmethic our output output of the any shorty Longetact VI Longetact VI Longetact VI dolores	
Ingent DVM Mont Target	
Filter by attributes Reinit-All Statut-Sall	
Action Attended Severe Betten Type F_MAME overs. DOTIQ_D P_MARK P_TAG HEGHIGO. HEGHIGO. PVE_MA. PVE_MA. <t< td=""><td></td></t<>	
	CCE STAR & ECHI NO.41

Repeat the same steps and any additional border devices and click on Save.

Now we are back to VRF Attachment tab, to deploy the configuration to devices click on **Deploy** from **Actions** (at top) drop-down.

Owner WH Allachmedian Interest Annexes Image: State of the state	VRF Overview - Azure_Tenant_VRF_50000								C	icliere - Refreah — 🗙
Million Mark Market Market </td <td>Overview VRF Attachments Networks</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>	Overview VRF Attachments Networks									
Vit/How Vit/D <	Filter by attributes									Actiere ~
Annu, Transult, 19 5000 200 Lad 5 IP Nome Annumber Logar HD Annu, Transult, 19 5000 200 Lad 4 IP Nomes Annumber of Lagar HD Annumber of Lagar HD	VRFNere VRFID	VLANID	Switch	Status	Attachment	Switch Role	Fabric Name	Loopback ID	Loopback IPV4 Address	Loopback IPV8 Address
Azere_Timan(1/2 5000 2000 Les/-6 Pencies Attached bordsr Azere-HC	Ature,Tenant,Vik 50000	2000	Leaf-5	PENDING	Attached	border	Azure-HCI			
	Azure,Tenant,VR 50000	2000	Leaf-6	PENDING	Attached	border	Azure-HCI			
Attached and Attached and Attached	Azure_Tenant_VR 50000	2000	Leaf-2	OBRICHED	Attached	leaf	Azure-HCI			
Asset.Tenent,V3 50000 Lesh1 Oliverity Assored inst Asure HO	Azure, Tenant, VR 50000	2000	Leaf-1	OEPLOYED	Attached	leaf	Azure-HCI			

Cisco NDFC will push the required configuration to the border devices in the VXLAN fabrics.

If the external network is also managed by NDFC, perform **Recalculate and Deploy** in External fabric too for Cisco NDFC to push configuration to the device which is being used as other end for VRF_Lite extension.

This allows VXLAN networks to be advertised to external and vice-versa for any outside communication to take place.

Appendix

Design Example with Microsoft Software Defined Networking (SDN) in Azure Stack HCI

In addition to VLAN based tenant network, Azure Stack HCI has a network design option with Microsoft SDN, which includes VXLAN termination in the server side. This section provides design examples of Cisco ACI and Nexus 9000 for Microsoft SDN in Azure Stack HCI. This section does not cover the configuration required on Azure Stack HCI side. The physical architecture of the Microsoft Azure HCI connectivity to Cisco Nexus Switches remains the same as the one explained in <u>Physical Architecture</u> section.

Microsoft Azure SDN Components

Microsoft Azure SDN introduces additional features, such as Software Load Balancer, Firewalls, Site-to-Site IPsec-VPN, and Site-to-Site GRE tunnels. The Software Load Balancer and Firewalls provide load balancing and firewalling services for the virtual machines hosted in the Azure Stack HCI cluster. Site-to-Site IPsec VPN and Site-to-Site GRE tunnels enable connectivity between virtual machines hosted in Azure Stack HCI cluster and external networks outside the Azure Stack HCI. The following VMs are the major components of Microsoft Azure SDN in Azure Stack HCI:

- Network Controller VMs: Network Controller VMs provide a centralized point to create and manage virtual network infrastructure inside the Azure Stack HCI. Network Controller VMs act as the control plane for the Azure Stack HCI SDN and do not carry actual data traffic. Microsoft recommends a minimum of three Network Controller VMs for redundancy.
- Software Load Balancer VMs: Software Load Balancer (SLB) VMs provide Layer 4 load balancing services for north-south and east-west TCP/UDP traffic. The Software Load Balancer VMs are installed on the Azure Stack HCI servers to provide load balancing services in the Azure Stack HCI Cluster. Microsoft uses the terminology Software Load Balancer Multiplexer VMs or SLB MUX VMs instead of SLB VMs. Henceforth, this document will use SLB MUX VMs to describe the Software Load Balancer VMs. A minimum of one SLB MUX VM is required per Azure Stack HCI Cluster, and the count can be increased based on the scale. More on the Software Load Balancer will be discussed later in this document.
- Gateway VMs: Gateway VMs create layer 3 connections between Microsoft Azure SDN virtual networks (VNETs) inside the Azure Stack HCI and external networks outside the Azure Stack HCI. Features such as IPsec VPNs and GRE tunnels are handled by the Gateway VMs. Microsoft recommends a minimum of two Gateway VMs per Azure Stack HCI Cluster and the count can be increased based on the scale.

Note: Please contact Microsoft for official scalability guidelines for the deployment of Network Controller VMs, SLB MUX VMs, and Gateway VMs.

Logical Architecture

Apart from the <u>Management Network</u> and <u>Storage Network</u> described earlier in this document, the following networks are to be used in Microsoft Azure SDN within Azure Stack HCI:

- HNV PA Network (Hyper-V Network Virtualization Provider Address Network)
- Logical Network

HNV PA Network

The Hyper-V Network Virtualization (HNV) Provider Address (PA) network is deployed when multi-tenancy is required in Microsoft Azure SDN within the Azure Stack HCI. The PA Network uses VXLAN encapsulation to achieve multi-tenancy. The PA network Address is similar to VTEP IP address in Nexus Switches. It serves as the underlay physical network for east-west VM-to-VM communication within an Azure Stack HCI cluster. The PA network requires a VLAN to be assigned on the physical network, which is passed as trunk on the data interfaces of all the servers in the cluster.

Each server in an Azure Stack HCI cluster has two PA network IP addresses, while each SLB MUX VM and Gateway VM has one IP address from the PA network. Thus, for a 16-node cluster, a /26 or larger subnet may be required because multiple SLB MUX VMs and Gateway VMs are required based on the scale.

Logical Network

A Logical Network is a network segment between the Azure Stack HCI servers and top-of-rack switches such as Cisco ACI leaf switches. Each Logical Network consists of a Logical subnet that requires a VLAN ID and an address prefix. The VLAN ID needs to be unique in the Azure Stack HCI cluster. The address prefix requires at least four IP addresses: one for the Azure Stack HCI cluster, one for each VLAN interface of each top-of-rack switch, and one for the virtual IP address that is shared by the pair of top-of-rack switches. The Logical Network acts as a physical path to carry traffic between the Azure Stack HCI VNET

and the top-of-rack switches. VNET is a virtual network in Azure Stack HCl and is equivalent to VRF in Cisco ACl and Nexus 9000 in NX-OS mode.

PA Network and SLB MUX VMs Connectivity

This section describes how to connect the PA network and SLB MUX VMs to a Cisco ACI and Cisco NX-OS based fabric.

Software Load Balancer (SLB)

An important consideration before designing the PA network connectivity in a Cisco ACI and Cisco NX-OS-based fabric is to understand the Software Load Balancer functionality and its connectivity requirements because SLB MUX VMs are mandatory in the Microsoft Azure SDN installation. SLB MUX VMs can be used for public access to a pool of load balanced VMs inside the VNET in Azure Stack HCI as well as load balancing network traffic within the VNETs.

This document uses an example with three SLB MUX VMs deployed in an Azure Stack HCI cluster. Each SLB MUX VM has one unique IP address from the PA network. An SLB MUX VM can be hosted on any of the Azure Stack HCI servers that are part of the Azure Stack HCI Cluster.

SLB MUX VMs need to have eBGP peering configured with the IPs of external routers (Cisco ACI Leaf switches in this case) for external network reachability.

Two additional IP Pools (Public VIP Pool and Private VIP pool) are required for the SLB MUX VMs deployment. The Public VIP Pool and Private VIP Pool are allocated to the SLB MUX VMs for assigning Virtual IPs. These Virtual IPs are used by applications or services that are hosted inside the Azure Stack HCI cluster that require the load balancing feature. These IP Pools are provisioned on top of the SLB MUX VMs.

Note: The SLB MUX VMs do not use an IP address to be assigned to themselves from these IP pools. SLB MUX VMs use IP addresses assigned from thePA network.

- Public VIP Pool: It must use IP subnet prefixes that are routable outside the Azure Stack HCI cluster (not necessarily an Internet Routable Public IP). These are front-end IP addresses that external clients use to access VMs in the VNETs, including front-end VIP for Site-to-Site VPN. The Public VIP is used to reach a load balanced application or a service from outside of the Azure Stack HCI cluster.
- Private VIP Pool: This IP subnet prefix is not required to be routable outside of the Azure Stack HCI cluster. These VIPs are meant to be accessed by internal client's that are part of the VNET in the Azure Stack HCI Cluster. The Private VIP is used if the load-balanced application or service does not require reachability from outside the Azure Stack HCI cluster.

Cisco ACI Design for PA Network and SLB Connectivity

SLB MUX VMs are part of PA network and need to have eBGP peerings with the leaf switches for communication with other networks. Therefore, L3Out needs to be configured with an encap VLAN that is same as the PA network VLAN ID configured inside the Azure Stack HCI.

The figure below illustrates a logical design example of eBGP peering of SLB MUX with Cisco ACI leaf switches.

Tenant: common



Figure 24.

eBGP peering of SLB MUX & ACI in PA Network

The figure above also illustrates an example of a high-level relationship between Cisco ACI tenant elements as deployed in the design for the Azure Stack HCI underlay connectivity. In this example, Cisco ACI common tenant contains a VRF called Common_VRF EPGs for storage and management networks.

This tenant also contains an L3Out named Cluster_01_PA_L3Out that is dedicated for the PA network connectivity for the specific cluster. eBGP will be the routing protocol configured in the L3Out, while the encap vlan used in the L3Out will be the same VLAN configured as the PA network VLAN in the Azure Stack HCI Cluster.

As this example has three SLB MUX VMs deployed per cluster, each Cisco ACI leaf will have three eBGP peers. Therefore, a total of six The eBGP peerings are established between the Azure Stack HCI cluster and the pair of Cisco ACI leaf switches. In this example, 10.2.1.0/24 is the IP subnet, and 401 is the VLAN ID assigned to the PA network. The SVI interface configured on Cisco ACI leaf switch will be 10.2.1.2/24 and 10.2.1.3/24 for Leaf 01 and Leaf 02 respectively. The three SLB MUX VMs will have IP addresses as 10.2.1.4/24, 10.2.1.5/24, and 10.2.1.6/24 respectively. The eBGP peering with a loopback IP address or an IP address that is not directly connected is NOT supported. Therefore, eBGP peering is formed with an L3Out SVI interface of the Cisco ACI leaf switches.

Note: Each Azure Stack HCI Cluster requires one dedicated EPG for storage, one dedicated EPG for management, and one dedicated L3Out and its external EPG for the PA network.

Azure Stack HCI VNET Connectivity (Logical Network and Gateway VMs connectivity)

VNET is a virtual network in the Azure Stack HCI. It is created with an address prefix. Multiple smaller subnets can be created from the VNET address prefix for the purpose of IP assignment to workload VMs.

One of the subnets is used as a gateway subnet. The gateway subnet is required to communicate outside the Azure Stack HCI VNET. An IP address from this subnet is automatically provisioned on the gateway VM. This subnet can be configured with a /28, /29, or /30 prefix. The /28 or /29 subnet prefix is required if an IPsec or GRE tunnel is needed in the gateway subnet because additional IP addresses from the subnet are provisioned on the gateway VMs whenever an IPsec or GRE tunnel is required. This document doesn't cover IPsec or GRE tunnel.

Cisco ACI Design for Azure Stack HCI VNET Connectivity

The gateway VM establishes two eBGP peerings with the loopback IP address configured on the pair of ACI leaf switches. A static route is required in the Azure Stack HCI VNET for reachability to the loopback IP address. The next hop IP address for the static route is the virtual IP address configured on the pair of Cisco ACI leaf switches from the Logical Network.

Note: The next hop IP address for static route used for eBGP peering is called the L3 Peer IP in Azure Stack HCI and the virtual IP address configured on VLAN interface in Azure Stack HCI is called the secondary IPv4 address in Cisco ACI.

An L3Out is configured on the Cisco ACI fabric for the connection towards the VNET in Azure Stack HCI Cluster. The Cisco ACI leaf switches establish two eBGP peerings (one from each ACI leaf switch) with the IP address assigned to the gateway VM. This IP address can be found in the BGP router IP address under the Gateway connections section in Azure Stack HCI. A static route is configured on the Cisco ACI leaf switches for reachability to the gateway VM IP address. The next hop for this static route is the IP address from the Logical Network configured on the Azure Stack HCI Cluster.

The figure below shows an example of the Cisco ACI L3Out with the Azure Stack HCI VNET connectivity.



Tenant: HCI1_tenant1

Figure 25.

EBGP peering of Azure Gateway VM with Cisco ACI Leaf switches

The design example has a 3-node Azure Stack HCI Cluster connected to a pair of ACI leaf switches, which contains the following network configurations in Azure Stack HCI:

- A VNET named VNET01 is created in Azure Stack HCI with an address prefix 192.168.1.0/24. The gateway subnet is 192.168.1.0/29.
- A Logical Network in the Azure Stack HCl uses the IP subnet 10.10.1.0/29 and VLAN ID 501. 10.10.1.6/29 is used for the gateway connection towards the Cisco ACl leaf switch. In this example, eBGP Multihop is used, and 65201 is the BGP ASN of the Gateway VM.
- Static routes (10.10.10.10/32 and 10.10.10.20/32 via 10.10.1.1) are configured to reach the loopback IP addresses of the pair of ACI leaf switches. The IP address 10.10.1.1 is configured as the virtual IP address (Secondary IPv4 address) on the VLAN interface of both ACI leaf switches.
- The Web and App VM that are also part of the VNET01 will always send traffic to the gateway VM if the destination IP address is outside the VNET_01.

To establish the connection with Azure Stack HCI, the Cisco ACI fabric contains the following configurations:

- An ACI tenant named HCI1_tenant1 and a VRF named VRF1 are created, which correspond to the VNET_01 in Azure Stack HCI.
- An L3Out named VNET01_L3Out is created for eBGP peering with the gateway VMs in VNET01.
 - Leaf01 has the loopback IP 10.10.10/32 and Leaf02 has the loopback IP 10.10.10.20/32.
 - The Logical Interface profile inside the L3Out is configured with VLAN interfaces. The VLAN interfaces are assigned IP addresses from the subnet 10.10.1.0/29, and the encap VLAN ID is 501 (which is same as the one defined in the Azure Stack HCI Logical Network).
 - A static route (192.168.1.0/29) is configured to reach the gateway VM (192.168.1.2) under the Logical Node profile inside the L3Out, and the next hop is 10.10.1.6.
 - o eBGP multihop with a value two or greater is required to build the eBGP peering.
- Another L3Out named EXT_L3Out is used for communication outside the Cisco ACI fabric.

Solution Deployment

This section provides a detailed procedure to configure Cisco ACI and Azure Stack HCI with SDN enabled. It is assumed that the ACI fabric and APICs already exist in the customer's environment. This document does not cover the configuration required to bring the initial ACI fabric online.

Table 3 lists the hardware and software versions used in this solution.

The figure below and Table 9 summarize the topology, interface, and L3 domain configuration parameters used in this section. The connection uses six 100 GbE interfaces between ACI leaf switches and Azure Stack HCI servers.



Figure 26.

Interface and L3 domain configuration for Azure Stack HCI servers with SDN

Table 9. Interface and L3 Domain configuration for Azure Stack HCI Ser	vers
------------------------------------------------------------------------	------

Interface	Interface Policy Group	LLDP Interface Policy	Interface PFC Policy	AAEP Name	Domain Name	Domain type	VLAN Pool
Leaf1 and Leaf2 Ethernet 1/11-13	Individual-HCI	HCI_LLDP (DCBXP: IEEE 802.1)	PFC-Auto	HCI_AAEP	HCI_EXT_L3DOM	L3	HCI_VLAN_pool (VLAN 400- 600)

Interface and L3 Domain configuration for Azure Stack HCI Servers

Tables 10 and 11 summarize the ACI common and the user tenant configuration parameters used in this section. The ACI leaf switches serve as the gateway to the Azure Stack HCI networks except for storage networks that are L2 only. Although contract names are listed for your reference, the shared L3Out configuration in the common tenant and contract configuration steps are not covered in this document.



Figure 27.

ACI Tenant Overview for Azure Stack HCI with Microsoft SDN

Table 10.	ACI common tenant c	onfiguration ex	xample for SLB MUX	connectivity
-----------	---------------------	-----------------	--------------------	--------------

Property	Name
Tenant	Common
Tenant VRF	common_VRF
Bridge domains	Storage-A in common_VRF (No subnet) Storage-B in common_VRF (No subnet) Mgmt in common_VRF (10.1.1.254/24)
Leaf nodes and interfaces	Node 101 & 102 ethernet1/11, 1/12 and 1/13
EPGs	EPG Mgmt in BD Mgmt EPG Storage-A in BD Storage-A EPG Storage-B in BD Storage-B
Contract	Contract_1_to-external
L3Out	Cluster_01_PA_L3Out (BGP) in common tenant
Logical Node Profiles	Cluster_01_PA_101_NP (Node-101) - Router ID: 1.1.1.1 Cluster_01_PA_102_NP (Node-102) Router ID: 2.2.2.2
Logical Interface Profile	Cluster_01_PA_101_IFP (eth1/11, eth1/12 and eth1/13) - Interface Type: SVI - Primary IP: 10.2.1.2/24 - Secondary IP: 10.2.1.1/24 - Encap: 401

Property	Name
	- BGP Peer: 10.2.1.4, 10.2.1.5, 10.2.1.6
	- Remote AS: 65002
	Cluster_01_PA_102_IFP (eth1/11, eth1/12 and eth1/13)
	- Interface Type: SVI
	- Primary IP: 10.2.1.3/24
	- Secondary IP: 10.2.1.1/24
	- Encap: 401
	- BGP Peer: 10.2.1.4, 10.2.1.5, 10.2.1.6
	Remote AS: 65002
External EPGs	Cluster_01_PA_EXT_EPG
	Export Route Control Subnet (0.0.0.0)

Table 11.	ACI user tenant	configuration	example for	Gateway VM	connectivity
-----------	-----------------	---------------	-------------	-------------------	--------------

Property	Name
Tenant	HCI1_tenant1
Tenant VRF	VRF1
Leaf nodes and interfaces	Node 101 & 102 ethernet1/11, 1/12 and 1/13
Contract	Contract_2_to-external
L3Out	VNET01_L3Out (BGP) in HCl1_tenant1
Logical Node Profiles	 VNET01_101_NP (Node-101) Loopback IP: 10.10.10.10 Router ID: 1.1.1.1 Static route: 192.168.1.0/29, Next Hop: 10.10.1.6 BGP Peer: 192.168.1.2, Source Interface: loopback Remote AS: 65201 VNET02_102_NP (Node-102) Loopback IP: 10.10.10.20 Router ID: 2.2.2.2 Static route: 192.168.1.0/29, Next Hop: 10.10.1.6 BGP Peer: 192.168.1.2, Source Interface: loopback
Logical Interface Profile	 VNET01_101_IFP (eth1/11, 1/12 and 1/13) Interface Type: SVI Primary IP: 10.10.1.2/29 Secondary IP: 10.10.1.1/29 VLAN Encap: 501 VNET01_102_IFP (eth1/11, 1/12 and 1/13) Interface Type: SVI Primary IP: 10.10.1.3/29 Secondary IP: 10.10.1.1/29

Property	Name
	VLAN Encap: 501
External EPGs	VNET01_EXT_EPG - Export Route Control Subnet (0.0.0.0)
	External Subnet for External EPG (192.168.1.0/24)

Create VLAN Pool for Azure Stack HCI L3 Domain

In this section, you will create a VLAN pool to enable connectivity to Azure Stack HCI.

To configure a VLAN pool to connect the Azure Stack HCI servers to the ACI leaf switches, follow these steps:

- 1. From the APIC top navigation menu, select **Fabric > Access Policies**.
- 2. From the left navigation pane, expand and select **Pools** > **VLAN**.
- 3. Right-click and select Create VLAN Pool.
- 4. In the **Create Pool** pop-up window, specify a Name (For example, **HCI_VLAN_pool**) and for Allocation Mode, select **Static Allocation**.
- For Encap Blocks, use the [+] button on the right to add VLANs to the VLAN pool. In the Create Ranges pop-up window, configure the VLANs that need to be configured from the leaf switches to the Azure Stack HCI servers. Leave the remaining parameters as is.
- 6. Click OK.
- 7. Click Submit.

cisc	APIC										
Syste	em Tenants	Fabric	Virtual Networking	Adm	in Ope	rations Ap	os Integrations				
	Inventory Fa	bric Policies	Access Policies	_							
Policie			()	\bigcirc	Pools - V	LAN					
O• a	uick Start										
🗄 Ir	terface Configuratio										
> 🚍 s					Creat	e VLAN Po	bol			(28
	lodules					Name:	HCI_VLAN_POOL				
→ 🖬 P	olicies					Description:	optional				
> 🚞 P	hysical and External										
~ 🖿 P				_		Allocation Mode:	Dynamic Allocation	Static Allocatio	n		
~	VLAN					Encap Blocks:					1 +
	Create Rar	nges						? X n	Allocation Mode	Role	
		Type: VLAN									
	Des	cription: optio	nal								
> 6		Range: VLAN	400 - 1		600						
>			Integer Value	- DAN	Int & Yalue						
	Allocation	n Mode: (Dyr	namic Allocation Inherit	allocMode	from parent	Static Allocation					
		Role: Ext	ernal or On the wire encapsu	ations	Internal						
									Car	cel Subm	it
							Cancel	ок		ocationy	
									[4007] (Static All [4014] (Static All	ocation)	

Configure L3 Domain for Azure Stack HCI

To create an L3 domain type and connect to Azure Stack HCl servers, follow these steps:

- 1. From the APIC top navigation menu, select Fabric > Access Policies.
- 2. From the left navigation pane, expand and select Physical and External Domains > L3 Domains.
- 3. Right-click L3 Domains and select Create L3 Domain.
- In the Create L3 Domain pop-up window, specify a Name for the domain (For example, HCI_EXT_L3DOM). For the VLAN pool, select the previously created VLAN pool (For example, HCI_VLAN_pool) from the drop-down list.

cisco APIC								
System Tenants Fabric	Virtual Networking Ad	min Operations	Apps	Integrations				
Inventory Fabric Policies	Access Policies							
Policies		L3 Domains						
Quick Start		Create L3 Dom	ain					\otimes
Interface Configuration		Name:	HCI_EXT_L	3DOM				1.
> Switches		Associated Attachable Entity Profile:	select a va	lue	\sim			t
> 🖬 Modules		VLAN Pool:	HCI_VLAN	POOL(static)	~ 🗗			
> 🖬 Interfaces		Security Domains:					Ö +	
Policies Physical and External Domains			Select	Name		Description		ji I
> 🚞 External Bridged Domains								
> 🚞 Fibre Channel Domains								
> 🖬 L3 Domains								
Physical Domains	1							
						Cance	Sul	omit

5. Click Submit.

Create Attachable Access Entity Profile for Azure Stack HCI L3 Domain

To create an Attachable Access Entity Profile (AAEP), follow these steps:

- 1. From the APIC top navigation menu, select Fabric > Access Policies.
- From the left navigation pane, expand and select Policies > Global > Attachable Access Entity Profiles.
- 3. Right-click and select Create Attachable Access Entity Profile.
- 4. In the Create Attachable Access Entity Profile pop-up window, specify a Name (For example, HCI_AAEP) and uncheck "Enable Infrastructure VLAN" and "Association to Interfaces".
- 5. For the **Domains**, click the [+] on the right side of the window and select the previously created domain from the drop-down list below **Domain Profile**.
- 6. Click Update.

- 7. You should now see the selected domain and the associated VLAN pool as shown below.
- 8. Click **Next**. This profile is not associated with any interfaces at this time because "Association to Interfaces" was unchecked in step 4 above. They can be associated once the interfaces are configured in an upcoming section.

cisco APIC								
System Tenants Fabric	Virtual Networking A	Create Attachable	e Access Entity Profile				0	\otimes
Inventory Fabric Policies	Access Policies	STEP 1 > Profile		1. Profile	2. Association	n To Interfac	ces	
Policies	$\bigcirc \bigcirc \bigcirc \bigcirc$	Name:	HCI_AAEP					
C Quick Start		Description:	optional					
F Interface Configuration								
> 🚞 Switches		Enable Infrastructure VLAN:						
> 🚞 Modules		Domains (VMM, Physical or External) To Be Associated					1	+
> 🚞 Interfaces		To Interfaces:	Domain Profile	Encapsulation				
✓			Physical Domain - HCI_phys	from:vlan-1600 to	o:vlan-1699			
> 🚞 Switch			L3 External Domain - HCI_EXT_L3DOM	from:vlan-400 to:	vlan-600			
> 🚞 Interface								
🗸 🚞 Global								
> 🚞 PTP User Profile								
> 🧮 DHCP Relay		FRO DEDLONA (ENT.						
	s	EPG DEPLOYMENT (All Sele	acted EPGs will be deployed on all the interfaces associated.)					
efault								+
RW_AEP		Application EPGs		Encap Pr	imary Encap	Mode		
Error Disabled Recovery Policy								
MCP Instance Policy default								
> 🧮 QOS Class								
> 🧮 Monitoring								
> 🧮 Troubleshooting								
> 🚞 Physical and External Domains		-						
> 🖿 Pools								
					Canc	el	vext	

9. Click Finish.

Perform the following configurations that are common for VLAN-based tenant network and Microsoft SDNbased network in Azure Stack HCI:

- <u>Create LLDP policy</u>
- <u>Create LLDP Interface Policy</u>
- <u>Create Interface Priority Flow Control Policy</u>
- <u>Create Interface Policy Group for Interfaces connected to Azure Stack HCI servers</u>
- Associate Interface Policy Group for Interfaces connected to Azure Stack HCI servers
- <u>Configure QoS</u>

The Management VLAN, Storage VLANs, and the PA VLAN are the VLAN-based networks for the Azure Stack HCI with SDN. The next sub-section covers an L3Out configuration example for PA network deployment. For the deployment of Management EPGs corresponding to the Management VLAN and Storage EPGs corresponding to the Storage VLANs, please refer "Configure EPGs" section" of this document.

Cisco ACI Configuration for PA Network and SLB Connectivity

This section explains how to configure L3Out in Cisco ACI to enable PA Network and SLB MUX VMs connectivity. To create an L3Out, follow these steps:

- 1. From the APIC top navigation menu, select **Tenants > common** (or select an existing tenant where you want to configure the PA L3Out).
- 2. From the left navigation pane, expand and select Networking > L3Outs.
- 3. Right-click and select Create L3Out.
- In the Name field, specify a Name (For example, Cluster_01_PA_L3Out), select a VRF name (In this example, Common_VRF), select a previously created L3 domain from the drop-down list (In this example, HCI_EXT_L3DOM).
- 5. Check the BGP checkbox and click Next.

cisco APIC	ac
System Tenants Eabric Virt	Create L3Out
ALL TENANTS Add Tenant Tenant Sea	1. Identity 2. Nodes And Interfaces 3. Protocols 4. External EPG
common	
	Protocol-
> C Quick Start	
Common	
> Application Profiles	Revite
V 🗖 Networking	
> 🚞 Bridge Domains	Leaf Router
> VRFs	
> L2Outs	
✓ 🗖 L3Outs	Identity
> 🚹 default	A Layer 3 Outside (L3Out) network configuration defines how the ACI fabric connects to external layer 3 networks. The L3Out supports connecting to external
> 🔤 SR-MPLS VRF L3Outs	networks using static routing and dynamic routing protocols (BGP, OSPF, and EIGRP).
> Dot1Q Tunnels	Prozenujeitae
> 🚞 Contracts	 Configure an L3 Domain and Fabric Access Policies for interfaces used in the L3Out (AAEP, VLAN pool, Interface selectors).
> E Policies	Configure a BGP Route Reflector Policy for the fabric infra MP-BGP.
> 🚞 Services	
Security	
> 🧰 IP Address Pools	Name: Cluster_01_PA_L3Out BGP GSP OSPF
	VRF: COMMON_VRF 🛛 🖓 🗗
	L3 Domain: HCLEXT_L3DOM V
	Use for GOLF:
	Previous Cancel Next

Uncheck the Use Defaults checkbox to manually specify a name in the Node Profile Name field (In this example, Cluster_01_PA_101_NP) and Interface Profile Name field (In this example, Cluster_01_PA_101_IFP).

ate L3Out						
			1. Identity	2. Nodes And Interfaces	3. Protocols	4. External EPC
Use Defaults:						
Node Profile Name: Clus	ster_01_PA_101_NP					
terface Types						
Layer 3:	terface Sub-Interface	SVI Floating SVI				
Layer 2: P	ort Virtual Port Channel	Direct Port Channel				
ades						
0003						
Node ID	Router ID	Loopback	Address			
LEAF1 (Node-101)	∨ 1.1.1.1			+ Hide Interfaces		
		Leave empty any Loopbac	to not configure k			
Interface	Interface Profile Name	Encap	MTU (bytes)	IP Address		
eth1/11 🗸	Cluster_01_PA_101_IFP	VLAN \checkmark 401	9216	10.2.1.2/24	+	
Ex: eth1/1 or topology/pod- 1/paths-101/pathep-[eth1/23]		Integer Va	lue	address/mask		
Interface	Interface Profile Name	Encap	MTU (bytes)	IP Address		
eth1/12 🗸	Cluster_01_PA_101_IFP	VLAN ~ 401	9216	10.2.1.2/24	1 +	
Ex: eth1/1 or topology/pod- 1/paths-101/pathep-[eth1/23]		Integer Va	lue	address/mask		
Interface	Interface Profile Name	Encap	MTU (bytes)	IP Address		
eth1/13	Cluster_01_PA_101_IFP	VLAN VLAN	9216	10.2.1.2/24		
1/paths-101/pathep-[eth1/23]		integer va	ue	address/mask		

- 7. In the Interface Types section, select SVI for Layer 3 and Port for Layer 2.
- 8. In the **Nodes** section, input all the details related to the first leaf switch (In this example, **Node ID** as **Node-101** and **Router ID** as **1.1.1.1**, leave the **Loopback Address** field blank).
- 9. Click + in the second row to add additional interfaces on the same Node (In this example, there are three servers connecting on three interfaces of one leaf switch, **eth1/11, 1/12** and **1/13**).
- 10. From the drop-down list, select the interfaces connecting to the servers, specify the Interface Profile Name, Encap, Encap value, MTU and IP address. The Azure Stack HCI servers uses maximum MTU size as 9174, hence the MTU configured on the TOR switches must be same or more than 9174 (In this example, Interface Profile Name is Cluster_01_PA_101_IFP, Encap is VLAN, Encap value is 401, MTU is 9216 and IP address is 10.2.1.2/24).
- 11. Enter the same values for all the interfaces and click **Next.** The equivalent configurations for the second leaf will be added later though it is also possible to add them through this wizard.
- 12. Click on Next without entering any BGP related information on this page.
| Crea | te L3Ou | it | | | | | | | | \otimes |
|--------|---------------|-------------------------|-------------------|----|-------------|-------------|---------------|--------------|-----------------|-----------|
| | | | | | 1. Identity | 2. Nodes An | d Interfaces | 3. Protocols | 4. External EPG | |
| Protoc | col Associa | ations | | | | | | | | |
| | BGP | | | | | | | | | |
| | Loopback P | olicies | | | | | | | | |
| | Node Profil | e: Cluster_01_PA_101_NP | | | | | | | | |
| | Neder | Deer Address | FDCD Multiber TT | De | mote AChi | | Hide Policy 🔲 | | | |
| | 101 | Peel Address | | | mote ASN | \Diamond | | | | |
| | Interface Pr | licies | | | | | | | | |
| | interface r c | nicies | | | | | | | | |
| | Node ID: 10 | 1 | | | | | Hide Policy | | | |
| | Interface | Peer Address | EBGP Multihop TTL | R | emote ASN | | | | | |
| | 1/11 | | | | | \bigcirc | | | | |
| | 1/12 | | | | | \Diamond | | | | |
| | 1/13 | | | | | \Diamond | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | Previous | Cancel Next | |
| | | | | | | | | | HOAT | |

13. Click on **Finish** on the **External EPG** page without making any changes at this moment. The External EPG will be created at a later stage.

Create L3Out				\otimes
	1. Identity	2. Nodes And Interfaces	3. Protocols	4. External EPG
External EPG				
The LOOUT Network on External EDC is used for traffic algoritisation	n contract acception	no and route control policies. O	accification is matching	outomol potucarico to this

The L3Out Network or External EPG is used for traffic classification, contract associations, and route control policies. Classification is matching external networks to this EPG for applying contracts. Route control policies are used for filtering dynamic routes exchanged between the ACI fabric and external devices, and leaked into other VRFs in the fabric.

Name:		
Provided Contract:	select a value	~
Consumed Contract:	select a value	~
Default EPG for all external networks:	✓	

Previous	Cancel	Finish
----------	--------	--------

14. From the APIC top navigation menu, select Tenants > common > Networking > L3Outs > L3Out Name (in this example, Cluster_01_PA_L3Out) > Logical Node Profiles (in this example, Cluster_01_PA_101_NP) > Logical Interface Profiles (in this example, Cluster_01_PA_101_IFP) > SVI.

common PAA									
	Logical Interface Profile - Clus	ter_01_PA_101_IFP							0
								Policy	Faults History
Application Profiles					Canaral	Douted Cub Interfa	Deuted	atorfocos	CV/I Election CV/I
V Networking					General	Routeu Sub-Interna	ces Routed	Internaces	SVI Floating SVI
Bridge Domains									O <u>+</u>
									<u> </u>
> 🖬 L20uts	▲ Path	Side A IP	Side B IP	Secondary IP	IP Address	MAC Address	MTU (bytes)	Encap	Encap Scope
✓ ➡ L3Outs	Ped-1/Nede-101/ath1/11			Address	10.21.2/24	00-22-PD-E9-10-EE	0216	vian-401	Local
✓	Pod-l/Node-Iol/ethi/11				10.2.1.2/24	00-22-80-26-19-22	9216	Vian=401	Local
Logical Node Profiles	Pod-1/Node-101/eth1/12				10.2.1.2/24	00:22:BD:F8:19:FF	9216	vlan-401	Local
Cluster_01_PA_101_NP	 Pod-1/Node-101/eth1/13 				10.2.1.2/24	00:22:BD:F8:19:FF	9216	vlan-401	Local
> 🧰 Configured Nodes									
Logical Interface Profiles									
Cluster_01_PA_101_IFP									
> 🚞 External EPGs									
> 🚞 Route map for import and export route control									
> 🚹 L3DOM_F5_01									
> 🚹 L3out-osp_shared									
> 合 default									
> 🚞 SR-MPLS VRF L3Outs									
> 🧮 Dot1Q Tunnels							Showrl	leaga Dr	
> 🧰 Contracts								Jage Ne	

15. Double-click on the first interface and click + to add the **IPV4 Secondary Addresses.** This will work as a virtual IP address, and it will be common across both the leaf switches (In this example, double click on **eth1/11** and enter secondary IP address as **10.2.1.1/24**).

SVI									
						Policy	Faults	Hi	story
8 👽 📣 🕔							Õ	+	***
Properties									
Path:	topology/pod-1/paths-10	01/pathe	p-[eth1/11]						
Path Description:									
Description:	optional								
Encap:	VLAN V 401								
Encap Scope:	VRF Local	2							
Auto State:	disabled enable	ed							
Mode:	Trunk (Native)	Trunk	Access (Untagged)	\bigcirc					
IPv4 Primary / IPv6 Preferred Address:	10.2.1.2/24 address/mask								
IPv6 DAD:	disabled enable	ed							
IPv4 Secondary / IPv6 Additional Addresses:			1 +						
	 Address IPv6 	6 DAD	Enable for DHCP Relay						
	10.2.1.1/24 ena	abled	Disabled						
Link-Local Address:	::								
					Show Usa	ige	Close		

- 16. Scroll down and click + to add the **BGP Peer Connectivity Profiles.** The BGP peer address will be the SLB MUX VMs IP address.
- 17. Enter the **Peer Address** and **Remote AS** while keeping all the values to their default and click on **Submit** (In this example, Peer Address is **10.2.1.4** and Remote AS is **65002**).

Create Peer Connectivity Profile	8
Peer Address: 10.2.1.4 address	
Description: optional	
Remote AS: 65002	
Admin State: Disabled Et	nabled
BGP Controls: 🕑 🔳	_
Allow Self AS	
AS override	
Disable Peer AS	Check
Next-hop Self	
Send Community	
Send Extended C	Community
Send Domain Pat	th
Capability: 🗌 Receive Addition	al Paths
Password:	
Confirm Password:	
Allowed Self AS Count: 3	
Peer Controls: Didirectional Forv	warding Detection ed Check
Address Type Controls: ☐ AF Mcast ☑ AF Ucast	
EBGP Multihop TTL: 1	
Weight for routes from this neighbor:	
	Cancel Submit

18. Repeat step 16 and step 17 to add multiple BGP peers and click **Close** (In this example, **10.2.1.5** and **10.2.1.6**).

SVI $\mathbf{\Omega} \otimes$ Policy Faults History Õ + **-Properties Link-Local Address: :: MAC Address: 00:22:BD:F8:19:FF MTU (bytes): 9216 Target DSCP: Unspecified External Bridge Group Profile: select an option BGP Peer Connectivity Profiles: 前 Peer IP Address Peer Controls 10.2.1.4 10.2.1.5 10.2.1.6 Rogue Exception MAC Group: select an option Exclude all MACs from Rogue EP Control: Show Usage

19. Repeat step 15 to step 18 for the remaining interfaces (In this example, eth1/12 and eth1/13).

common	090	Logical Interface Profile - Cluster	_01_PA_101_IFP							Q
V 🖿 L3Outs									Policy	Faults History
✓										
Logical Node Profiles						General	Routed Sub-Interfac	es Routed I	nterfaces	SVI Floating SVI
Cluster_01_PA_101_NP										
> 🧮 Configured Nodes										
🗸 🧮 Logical Interface Profiles										₩ +
✓		 Path 	Side A IP	Side B IP	Secondary IP Address	IP Address	MAC Address	MTU (bytes)	Encap	Encap Scope
BGP Peer 10.2.1.4- Node-10	01/1/11	Pod-1/Node-101/eth1/11			10.2.1.1/24	10.2.1.2/24	00:22:BD:F8:19:FF	9216	vlan-401	Local
BGP Peer 10.2.1.4- Node-10	01/1/12	Pod-1/Node-101/eth1/12			10.2.11/24	10.21.2/24	00:22:8D:E8:19:EE	9216	vlan-401	Local
BGP Peer 10.2.1.4- Node-10	01/1/13	Ded 10lade 101/stb102			10.011/04	10.01.0/04	00-22-80-58-30-55	0.216	ulan 401	Leeel
BGP Peer 10.2.1.5- Node-10	01/1/11	Pod-1/Node-101/eth1/13			10.2.1.1/24	10.2.1.2/24	00.22.00.10.19.11	9210	vian-401	Local
BGP Peer 10.2.1.5- Node-10	01/1/12									
BGP Peer 10.2.1.5- Node-10	01/1/13									
BGP Peer 10.2.1.6- Node-10	01/1/11									
BGP Peer 10.2.1.6- Node-10	01/1/12									
BGP Peer 10.2.1.6- Node-10	01/1/13									
> 🚞 External EPGs										
> The Route map for Import and export route of the Rou	control									
> 📤 L3DOM_F5_01										
> 🐽 L3out-osp_shared										
> 合 default								Show L	Isage Re	
> 🚞 SR-MPLS VRF L3Outs										

- 20. Notice all the **BGP Connectivity Profiles** are seen on the leaf side under **Logical Interface Profile** (In this example, there are nine BGP Connectivity Profiles considering three BGP peers on per interface).
- 21. Navigate to **Tenants > common > Networking > L3Outs > L3Out Name** (In this example, **Cluster_01_PA_L3Out) > Logical Node Profiles.**
- 22. Right-click and select **Create Node Profile.** This will create a Node Profile for the second leaf switch.
- 23. Specify the **Name** and click + to add the **Nodes** details (In this example, the name will be **Cluster_01_PA_102_NP**).

Create Node Pi	rofile				\mathbf{X}
Name:	Cluster_01_PA_102_NP				
Description:	optional				
Target DSCP:	Unspecified V				
BGP Timers:	select a value				
Nodes:					+
	Node ID	Router ID	Static Routes		
BGP Peer Connectivity Profiles:					+
	Peer IP Address	Peer Controls			
			Cancel	Submit	

24. Specify the **Node ID** and **Router ID**. Uncheck the **Use Router ID as Loopback Address** checkbox and click **OK** lin this example, Node ID is **102**, Router ID is **2.2.2.2**).

Select Node					(\times
Node ID	: LEAF2 (Node-102)	\sim				
Router ID	: 2.2.2.2					
Use Router ID as Loopback Address	:					
Loopback Addresses:						+
	IP					•
Static Routes:						+
	IP Address	Description	Next Hop IP	Track Policy		

- 25. Click **Submit** on the Node Profile page.
- 26. Navigate to **Tenants > common > Networking > L3Outs > L3Out Name** (In this example, Cluster_01_PA_L3Out) > Logical Node Profiles (in this example, Cluster_01_PA_102_NP) > Logical Interface Profiles.
- 27. Right-click and select Create Interface Profile.
- 28. Specify the Name, select the SVI tab (In this example, the name is Cluster_01_PA_102_IFP).

Create Interface Profile

CTED	1	d	ant	i+.,
JIEP	-	u	51 I L	ILV

STEP 1 > Identity				1. Identity
Name:	Cluster_01_PA_102_IFP			
Description:	optional			
		Routed Sub-Interfaces	Routed Interfaces SVI	Floating SVI
SVI Interfaces				1 +
Path	IP Address	MAC Address	MTU (bytes)	

Config Protocol Profiles: Config Advance Protocol: 🗌

29. Click + to create the SVI interface.

 \times

Select SVI

Path Type:	Port D	irect Port Channel	Virtual Port Chan	nel		
Node: L	EAF2 (Node-10)2) 🗸				
Path:	topology/pod-1/r	node-1				
ex.	: topology/pod-1/p	paths-101/pathep-[eth1	/23]			
Description:	optional					
Encap:	VLAN 🗸 4	01				
	Int	teger Value				
Encap Scope:	VRF Lo	ocal				
Auto State:	disabled	enabled				
Mode:	Trunk (Nativ	/e) Trunk	Access (Untagged))		
IPv4 Primary / IPv6 Preferred Address:	10.2.1.3/24					
Link-Local Address:						
IPv4 Secondary / IPv6 Additional Addresses:			1 +			
	Address	IPv6 DAD	Enable for			
	_		DHCP Relay			
	10.2.1.1/24	enabled	Disabled			
MAC Address:	00:22:BD:F8:1	9:FF				
MTIL (bytes)	9216					
Tarmat DCOD:	5210					
Target DSCP.	Unspecified					
External Bridge Group Profile:	select an optic	on 🗸				
BGP Peer Connectivity Profiles:					1	+
	Peer IP Addre	ess F	Peer Controls			
	10.2.1.4					
	10.2.1.5					
	10.2.1.6					
Rogue Exception MAC Group:	select an optic	on 🗸				

- 30. Select the Path Type, specify the Node, Path, Encap Vlan id, IPV4 Primary Address, IPV4 Secondary Addresses, MTU, and BGP Peer Connectivity Profiles, and click OK at the bottom of the page (In this example, Path type is Port, Node is 102, Path is eth1/11, Encap Vlan id is 401, IPV4 Primary Address is 10.2.1.3/24, IPV4 Secondary Address is 10.2.1.1/24, MTU is 9216 bytes, BGP Peer IPs are 10.2.1.4, 10.2.1.5 and 10.2.1.6, and BGP AS number is 65002).
- 31. Repeat step 29 and step 30 for the remaining interfaces and click **Finish** (In this example, interface **eth1/12** and **eth1/13**).

 \times

common (*) (=) (©)	Logical Interface Profile - Clust	ter_01_PA_102_IFP							0
✓ ➡ L3Outs	Ĵ.								w
✓								Policy	Faults History
V 🚞 Logical Node Profiles					General	Routed Sub-Interfac	ces Routed	Interfaces	SVI Floating SVI
Cluster_01_PA_101_NP									
Cluster_01_PA_102_NP									0 +
> 🧮 Configured Nodes									☆ +
Logical Interface Profiles	 Path 	Side A IP	Side B IP	Secondary IP Address	IP Address	MAC Address	MTU (bytes)	Encap	Encap Scope
Cluster_01_PA_102_IFP	Pod-1/Node-102/eth1/11			10.2.1.1/24	10.2.1.3/24	00:22:BD:F8:19:FF	9216	vlan-401	Local
BGP Peer 10.2.1.4- Node-102/1/11	Pod-1/Node-102/eth1/12			10.2.1.1/24	10.2.1.3/24	00:22:BD:F8:19:FF	9216	vlan-401	Local
BGP Peer 10.2.1.4- Node-102/1/13	Pod-1/Node-102/eth1/13			10.2.1.1/24	10.2.1.3/24	00:22:BD:F8:19:FF	9216	vlan-401	Local
BGP Peer 10.2.1.5- Node-102/1/11									
BGP Peer 10.2.1.5- Node-102/1/12									
BGP Peer 10.2.1.5- Node-102/1/13									
BGP Peer 10.2.1.6- Node-102/1/11									
BGP Peer 10.2.1.6- Node-102/1/12									
BGP Peer 10.2.1.6- Node-102/1/13									
> 🛅 External EPGs									
> Route map for import and export route control									
> 🕂 L3DOM_F5_01									
> 🛧 L3out-osp_shared							Show	Usage	
> 🚹 default							Conom		

- 32. From the APIC top navigation menu, select **Tenants > common > Networking > L3Outs > L3Out Name** (In this example, **Cluster_01_PA_L3Out**) **> External EPGs.**
- 33. Right-click and select **Create External EPG.** Specify the **Name** (In this example, **Cluster_01_PA_EXT_EPG**).
- 34. Click + and add **Subnet** which is to be advertised by ACI Leaf (or received) to the SLB MUX VMs via this L3Out (In this example, IP subnet **0.0.0.0/0** is advertised by ACI Leaf and hence marked as **Export Route Control Subnet**).

common	$\bigcirc \bigcirc \bigcirc \bigcirc$	External EPG - Cl	uster_01_F	PA_EXT_EPO	3								0
> Application Profiles													w
V 🖿 Networking								Policy	Operational	Health	Faults	Histo	ory
> 🧮 Bridge Domains							General Contracts	Inherit	ted Contracts	Subject La	ihels F	PGLab	als
> 🖿 VRFs								inition	eu contracto	Subject Lu	0010	I O LUD	010
> 🔚 L2Outs											Ó	+	*~~
✓ ➡ L3Outs		Properties											
✓		Target DSCP.	Unspecified										
Logical Node Profiles		Configuration Status:	applied										
> 🗧 Cluster_01_PA_101_NP		Configuration issues:											
Cluster_01_PA_102_NP		Preferred Group Member:	Exclude	Include									
✓		Intra Ext-EPG Isolation:	Enforced	Unenforced									
Cluster_01_PA_EXT_EPG		Subnets:	-									₫ +	11
> The Route map for import and export route co	ntrol		 IP Address 	5	Scope	Name	Aggregate		Route Control Profile	Route	e Summariza	tion	1
> 🚹 L3DOM_F5_01										Policy	у		
> 🚹 L3out-osp_shared			0.0.0.0/0		Export Route Control Subnet								
> 合 default			172.16.1.0/23	3	External Subnets for the External EPG	Public_VIP_F	Pool						
> 🚞 SR-MPLS VRF L3Outs													
> 🧮 Dot1Q Tunnels													
> 🧮 Contracts													
> 🚍 Policies												1	
> 🖿 Services									External EPG	- Cluster_01_F	W_EXT_EPG	J	
E Security									Show Usage	e Re			
> 🥅 IP Address Pools													

Subnets which are advertised by the SLB MUX VMs such as Public VIP pool can be added in the **Subnet** section of the External EPG and marked as **External Subnet for External EPG** (In this example, IP subnet **172.16.1.0/23** is configured as Public VIP Pool on SLB MUX VMs and hence marked on Cisco ACI leaf as **External Subnet**).

<u>Configure Contracts</u> as discussed in the previous sections. A contract is necessary to permit traffic between the L3Out external EPG and other L3Out External EPGs or EPGs part of the ACI fabric.

Contracts can be added to the External EPGs from the following path - Tenants > common > Networking > L3Outs > L3Out Name (In this example, Cluster_01_PA_L3Out) > External EPGs > External EPG Name (In this example, Cluster_01_EXT_EPG) > Policy > Contracts > Add Provided Contract or Add Consumed Contract.

Cisco ACI Configuration for Azure Stack HCI VNET & Gateway VM Connectivity

The previous section covered the deployment of EPGs and L3Out to build the Azure Stack HCl underlay network. This section explains how to configure Cisco ACl to support customer's workload deployed in Azure Stack HCl. In this example, a Cisco ACl tenant, a VRF, and an L3Out that connects to the Azure HCl VNET are configured. The following are the configuration steps:

- 1. From the APIC top navigation menu, select Tenants > Add Tenant.
- 2. In the Create Tenant dialog box, specify a Name (For example, HCl1_tenant1).
- 3. In the VRF Name field, enter the VRF name and click Finish (For example, VRF1).

CISCO APIC (LAB3-S2)		
System Tenants Fabric Virtual Na	tworking Admin Operation	ne Anne Integratione
ALL TENANTS Add Tenant Tenant Search: name	e or descr	11.tenant1 momt infra
HCI1 tenant1	Create VRF	8
	STEP 1 > VRF	1. VRF
> E Application Profiles	Name:	VRF1 1
V 🖬 Networking	Alias:	
> 📩 Bridge Domains	Description:	optional
VRFs		
> 🖿 L2Outs	Annotations:	Click to add a new annotation
> 🚞 L3Outs	Policy Control Enforcement Preference:	Enforced Unenforced
> 🚞 SR-MPLS VRF L3Outs	Deliay Cantrol Enforcement Direction	Farace Ingrase Mixed policy
> 🚞 Dot1Q Tunnels	Policy Control Enforcement Direction:	Egress ingress invixed policy
> 🧮 Contracts	BD Enforcement Status:	
> 🧮 Policies	Endpoint Retention Policy.	This policy only applies to remote
> 🧰 Services	Monitoring Policy:	select a value
Security	DNS Labels:	
		enter names separated by comma
	Transit Route Tag Policy:	select a value
	IP Data-plane Learning:	Disabled Enabled
	Create A Bridge Domain:	
	Configure BGP Policies:	
	Configure OSPF Policies:	
	Configure EIGRP Policies:	
		Previous Cancel Finish

- 4. From the left navigation pane, expand and select **Networking > L3Outs.**
- 5. Right-click and select Create L3Out.
- In the Name field, specify a Name (For example, VNET01_L3Out), select a VRF name (In this example, VRF1), and select a previously created L3 domain from the drop-down list (In this example, HCI_EXT_L3DOM).
- 7. Check the BGP checkbox and click Next.

CISCO APIC (LAB3-S2)	pratnik	ka (
System Tenants Fabric V	Create L3Out	8
HCI1_tenant1	Leiner 2. House And interfaces U. Frotocols H. External Dro	
 > O Quick Start > I HCILtenant1 > Application Profiles > Networking > Networking > I Minga Domains > I ViFs 	Leaf Route Route	
 220uts 20uts Sh-MPLS VRF L30uts Dottl0 Tunnels Contracts Policies Services 	Identity A Layer 3 Outside (L3Out) network configuration defines how the ACI fabric connects to external layer 3 networks. The L3Out supports connecting to external networks using static routing and dynamic routing protocols (BGP, OSPF, and EIGRP). Prerequisites: • Configure an L3 Domain and Fabric Access Policies for interfaces used in the L3Out (AAEP, VLAN pool, Interface selectors). • Configure a BGP Route Reflector Policy for the fabric infra MP-BGP.	
	Name: VNET01_L3Out SGP GGP OSPF	
	Previous Cancel Next)

8. Uncheck the **Use Defaults** checkbox to manually specify a name in the **Node Profile Name** field (In this example, **VNET01_NP**).

cisco APIC (LAB3-S2)								pratnika
System Tenants Fabric V	Create L3Out							8
ALL TENANTS Add Tenant Tenant Sea				1	. Identity	2. Nodes And Interf	aces 3. Protocols	4. External EPG
HCI1_tenant1								
 > O Quick Start > ■ HCI1_tenant1 > ■ Application Profiles > ■ Bridge Domains > ■ Bridge Domains > ■ L2Outs > ■ L3Outs > ■ SR-MPLS VRF L3Outs 	Use Defaults: Node Profile Name: VNE Interface Types Layer 3: Layer 2: P Nodes Node D	T01_101_NP terface Sub-Interface ort Virtual Port Channe Router ID	SVI Floa Direct Port C	ting SVI Channel	55			
> 🖿 Dot1Q Tunnels	LEAF1 (Node-101)	 ✓ 1.1.1.1 		10.10.10.10 Leave empty to not	configure	+ Hide Interfaces		
> Contracts				any Loopback				
> E Policies	Interface	Interface Profile Name	Encap		MTU (bytes)	IP Address		
> 🚍 Services	eth1/11 ~	VNET01_101_IFP	VLAN ~	501	9216	10.10.1.2/29	(1) (+)	
Security	Ex: eth//1 or topology/pod- 1/paths-101/pathep-[eth1/23] Interface	Interface Profile Name	Encap VLAN	Integer Value	MTU (bytes)	IP Address	a a	
	Ex: eth1/1 or topology/pod- 1/paths-101/pathep-[eth1/23]			Integer Value		address/mask		
	Interface	Interface Profile Name	Encap	1	MTU (bytes)	IP Address		
	eth1/13 Ex: eth1/1 or topology/pod- 1/paths-101/pathep-[eth1/23]	VNET01_101_IFP	VLAN	Integer Value	9216	address/mask		
							Previous	Cancel Next

- 9. In the Interface Types section, select SVI for Layer 3 and Port for Layer 2.
- 10. In the **Nodes** section, input all the details related to the first leaf switch (In this example, **Node ID** as **Node-101**, Router ID as **1.1.1.**, and **Loopback Address** as **10.10.10.10**).
- 11. Click + in the second row to add additional interfaces on the same Node (In this example, there are three servers connecting on three interfaces of one leaf switch, **eth1/11, 1/12** and **1/13**).

- 12. From the drop-down list, select the interfaces connecting to the servers, specify the Interface Profile Name, Encap, Encap value, MTU, and IP address. The Azure Stack HCI servers uses maximum MTU size as 9174, hence the MTU configured on the TOR switches must be same or more than 9174 (In this example, values are VNET01_101_IFP, VLAN, 501, 9216 and 10.10.1.2/29).
- 13. Enter the same values for all the interfaces belonging to the first Node.
- Click + in the first row to add additional Node and input all the details regarding the second leaf switch (In this example, Node ID as 102, Router ID as 2.2.2.2, and Loopback Address as 10.10.10.20).
- 15. Click + to add additional interfaces below the second Node (In this example, there are three interfaces eth1/11, eth1/12, and eth1/13 on second leaf connecting to Azure Stack HCI servers).
- From the drop-down list, select the interfaces connecting to the servers, specify the Interface Profile Name, Encap, Encap value, MTU, and IP address (In this example, values are VNET01_102_IFP, VLAN, 501, 9216 and 10.10.1.3/29).

Node ID LEAF2 (Node-102)	Router ID		Loopback Address 10.10.10.20 Leave empty to not co any Loopback	nfigure	+ Hide Interfaces						
Interface eth1/11 Ex: eth1/1 or topology/pod- 1/paths-101/pathep-[eth1/23]	Interface Profile Name VNET01_102_IFP	Encap VLAN 🗸	501 Integer Value	MTU (bytes) 9216	IP Address 10.10.1.3/29 address/mask	+					
Interface eth1/12 Ex: eth1/1 or topology/pod- 1/paths-101/pathep-[eth1/23]	Interface Profile Name VNET01_102_IFP	Encap VLAN 🗸	501 Integer Value	MTU (bytes) 9216	IP Address 10.10.1.3/29 address/mask	÷					
Interface eth1/13 Ex: eth1/1 or topology/pod- 1/paths-101/pathep-[eth1/23]	Interface Profile Name VNET01_102_IFP	Encap VLAN 🗸	501 Integer Value	MTU (bytes) 9216	IP Address 10.10.1.3/29 address/mask	۲					
						Previous	Cancel	Cancel	Cancel	Cancel Nex	Cancel

- 17. Click Next.
- 18. Enter the BGP-related information in the **Loopback Policies** section and leave the **Interface Policies** section blank.
- 19. Enter **Peer Address,** which is the IP address assigned to Gateway VM from the Gateway subnet inside the VNET (In this example, **192.168.1.2).**
- 20. Enter the **EBGP Multihop TTL.** This value must be greater than one as eBGP peer is not directly connected (It needs to be more than 1 because the peerings are not between directly connected IP addresses. In this example, it is configured as **4**).
- 21. Enter the **Remote ASN.** This will be the BGP ASN value configured on Azure Stack HCI VNET (In this example, it is configured as **65201).**
- 22. Click Next.

CISCO APIC (LAB3-S	32)							pra
System Tenants Fabric	Create L3Ou	t						\otimes
ALL TENANTS Add Tenant Te				1. Identity	2. Nodes And	I Interfaces	3. Protocols	4. External EPG
HCI1_tenant1	Protocol Associa	itions						II.
> 🕩 Quick Start	BGP							
✓ ₩ HCI1_tenant1	Loopback Pe	olicies						
> Application Profiles								
Networking Reidee Demains	Node Profile	: VNET01_NP						
> WRFs	Nodes	Peer Address	EBGP Multihop TTL	Remote ASN		Hide Policy		
> 🚞 L2Outs	101,102	192.168.1.2	4	65201	\Diamond			
🖿 L3Outs								
> 🚞 SR-MPLS VRF L3Outs	Interface Po	licies						
> 📩 Dot1Q Tunnels	Node ID: 10	1						
Contracts						Hide Policy 🗌		
> Services	Interface	Peer Address	EBGP Multihop TTL	Remote ASN				
Security	1/11							
	1/12			\Diamond	\bigcirc			
	1/13				^			
	1/13							
	Node ID: 10	2						
						Hide Policy 🗌		
	Interface	Peer Address	EBGP Multihop TTL	Remote ASN				
	1/11							
							-	
							Previous	Cancel Next

23. In the Name field, enter the name of the External EPG (In this example, VNET01_EXT_EPG).

24. Click + to add the subnets which are advertised or received via this L3Out. After the VNET's eBGP peering with the top of rack switches, the gateway VMs advertise the entire VNET subnet to the top of rack switches (In this example, **192.168.1.0/24** is the VNET subnet that is received by the ACI leaf switches and hence marked as **External Subnets for External EPG.** The ACI leaf switches are the only exit path for the Azure Stack HCI VNET to reach the external networks outside of Azure Stack HCI, hence **0.0.0.0/0** is advertised to Azure Stack HCI VNET and marked as **Export Route Control Subnet**).

cisco APIC (LAB3-S	52)						pra
System Tenants Fabric	Create L3Out						8
ALL TENANTS Add Tenant Te				1. Identity	2. Nodes And Interfaces	3. Protocols	4. External EPG
HCI1_tenant1	External EPG						
O Quick Start UI1_tenant1 E Application Profiles	The L3Out Networ EPG for applying c in the fabric.	k or External EPG is used ontracts. Route control p	d for traffic classification, co policies are used for filtering	ontract associatior g dynamic routes e	ns, and route control policies xchanged between the ACI f	. Classification is matchin abric and external device	g external networks to this s, and leaked into other VRFs
V I Networking		Name: VNET01_EXT_	EPG				
> 🚞 Bridge Domains	Provid	led Contract: select a value					
> VRFs	Consum	ed Contract: select a value					
> L2Outs	Default EPG for all extern	nal networks: 📃					
	Subnets						
> SR-MPLS VRF L3Outs							T +
> Dotro Tunneis	IP Address	Scope	Name	Agg	pregate Ro	oute Control Profile	Route Summarization Policy
	192.168.1.0/24	External Subnets for	the Exter				
	0.0.0/0	Export Route Contro	I Subnet				
						Previous	Cancel Finish

- 25. Click **Finish.** The contracts can be added at a later stage based on the traffic flow.
- 26. Navigate to Tenants > HCl1_tenant1 > Networking > L3Outs > L3Out Name (In this example, VNET01_L3Out) > Logical Node Profiles (In this example, VNET01_NP) > Logical Interface Profiles > Interface Profile Name (In this example, VNET01_101_IFP) > Policy > SVI.

HCI1_tenant1	00	Logical Interface Profile - VNE	T01_101_IFP								Q
> C+ Quick Start									Policy	Faults	History
V III HCI1_tenant1									Toney		Thistory
> Application Profiles						General	Routed Sub-Interfa	ces Route	d Interfaces	SVI Flor	ating SVI
V Networking											0 +
> 🧰 Bridge Domains											0 -
> 🖿 VRFs											☆ +
> 🖿 L2Outs		 Path 	Side A IP	Side B IP	Secondary IP Address	IP Address	MAC Address	MTU (bytes)	Encap	Encap S	Scope
V 📥 L3Outs		Pod-1/Node-101/eth1/11				10.10.1.2/29	00:22:BD:F8:19:FF	9216	vlan-501	Local	
VNET01_L3Out		Pod-1/Node-101/eth1/12				10.10.1.2/29	00:22:BD:F8:19:FF	9216	vlan-501	Local	
		Pod-1/Node-101/eth1/13				10.10.1.2/29	00:22:BD:F8:19:FF	9216	vlan-501	Local	
BGP Peer 192 168 1 2											
> Configured Nodes	1										
✓											
VNET01_101_JFP											
> = VNET01_102_IFP											
> 🚞 External EPGs											
> E Route map for import and export route control											
> 🚞 SR-MPLS VRF L3Outs											
> 🚞 Dot1Q Tunnels											
> E Contracts											
> 🧰 Policies											
> 🚞 Services											
Security								Show	v Usage		

- 27. Double-click on the first interface (in this case, interface eth1/11).
- 28. Scroll down and click + to add **IPV4 Secondary / IPv6 Additional Addresses** (in this case, **10.10.1.1/29**).

SVI						G	
			_	Policy	Faults	His	tory
8 👽 🛆 🕔					Õ	+	***
Properties Path: Path Description: Description: Encap Encap Scope: Auto State: Mode:	topology/pod-1/paths-101 optional VLAN S01 integer Value VRF Local disabled enabled Trunk (Native) Tr	/pathep-[eth1/11]	\supset				
IPv4 Primary / IPv6 Preferred Address: IPv6 DAD: IPv4 Secondary / IPv6 Additional Addresses: Link-Local Address:	10.10.1.2/29 address/mask disabled enabled Address IPv6 [10.10.1.1/29 enabl	DAD Enable for DHCP Relay led Disabled					
NAO Addenna	0000000000000		Show Usage	Clo	ose		

29. Click **Close** at the bottom of the page.

30. Repeat steps 27 to 29 for other interfaces (In this example, eth1/12 and eth1/13).

System Tenants Fabric Virtual Networking	Admin Operations Apps	Integrations							
ALL TENANTS Add Tenant Tenant Search: name or descr	common HCI1_tenant1 user1	tn-hshahane test	BR012						
HCl1_tenant1 (*) ()	Logical Interface Profile - VNET	01_101_IFP							0
> C Quick Start	-								•
✓ III HCI1_tenant1								Policy	Faults History
> 🚞 Application Profiles					General	Routed Sub-Interfa	ces Routed	Interfaces	SVI Floating SVI
V 🚞 Networking	0000								<u> </u>
> 🧮 Bridge Domains									0 -
> 🚞 VRFs									☆ +
> 🚞 L2Outs	▲ Path	Side A IP	Side B IP	Secondary IP Address	IP Address	MAC Address	MTU (bytes)	Encap	Encap Scope
✓ ➡ L3Outs	Pod-1/Node-101/eth1/11			10.10.1.1/29	10.10.1.2/29	00:22:8D:F8:19:FF	9216	vlan-501	Local
VNET01_L3Out	Pod-1/Node-101/ath1/12			101011/29	10 10 1 2/29	00:22-RD-ER-19-EE	9216	vlan-501	Local
V 🔤 Logical Node Profiles				10.10.1.1/2.0	10.10.1.2/20	00/22/00/10/10/17	0010	vier 501	Local
VNET01_NP	Pod-1/Node-101/ethi/13			10.10.1.1/29	10.10.1.2/29	00-22-60-16-19-FF	9216	vian-501	LOCAI
BGP Peer 192.168.1.2									
Configured Nodes									
Cogical Interface Profiles									
VNETOT_TOZ_HPP									
CAUSE and the second seco									
SP-MPI S VEF I 30ute									
> Contracts									
> Policies									
> E Services									
Security									
							Show	Usage	

- 31. Navigate to second Logical Interface Profile via Tenants > HCl1_tenant1 > Networking > L3Outs > L3Out Name (In this example, VNET01_L3Out) > Logical Node Profiles (In this example, VNET01_NP) > Logical Interface Profiles > Interface Profile Name (In this example, VNET01_102_IFP) > Policy > SVI.
- 32. Repeat step 27 to step 30 for the Node-102. (In this example, **eth1/11, eth1/12** and **eth1/13**, and **10.10.1.3** is the primary IP Address for Node-102).

System Tenants Fabric Virtual Networking	Admin Operations Apps	Integrations							
ALL TENANTS Add Tenant Tenant Search: name or descr	common HCI1_tenant1 use	er1 tn-hshahane test	BR012						
HCI1_tenant1 (D) (C) (C)	Logical Interface Profile - VNB	ET01_102_IFP							0
> 🕞 Quick Start								Dellow	Equita History
✓ III HCI1_tenant1								Policy	Faults History
> E Application Profiles					General	Routed Sub-Interfa	ces Routed	i Interfaces	SVI Floating SVI
V 🖿 Networking									0 +
> 🧮 Bridge Domains									0 _
> 🖿 VRFs									Π +
> 🚞 L2Outs	 Path 	Side A IP	Side B IP	Address	IP Address	MAC Address	MTU (bytes)	Encap	Encap Scope
V 🔤 L3Outs	Pod-1/Node-102/eth1/11			10.10.1.1/29	10.10.1.3/29	00:22:8D:F8:19:FF	9216	vlan-501	Local
VNET01_L3Out	Pod-1/Node-102/eth1/12			10.10.1.1/29	10.10.1.3/29	00:22:BD:F8:19:FF	9216	vlan-501	Local
Cogical Node Profiles	Pod-1/Node-102/eth1/13			10.10.1.1/29	10.10.1.3/29	00:22:8D:E8:19:EE	9216	vlan-501	Local
					10110110120	00.22.00.10.11	02.0		
Configured Noder									
Source Profiles									
VNET01_101_IFP									
> VNET01_102_IFP									
> 🧮 External EPGs									
> E Route map for import and export route control									
> 🧮 SR-MPLS VRF L3Outs									
> 🚞 Dot1Q Tunnels									
> E Contracts									
> 🚞 Policies									
> 🚍 Services									
Carlos Security							Show	Usage R	

- 33. Navigate to Tenants > HCl1_tenant1 > Networking > L3Outs > L3Out Name (In this example, VNET01_L3Out) > Logical Node Profiles (In this example, VNET01_NP) > Configured Nodes > Node path (In this example, topology/pod-1/node-101).
- 34. Click + to add Static Routes.

System Tenants Fabric	virtual Networking	Admin Operations Apps	Integrations						
ALL TENANTS Add Tenant	Tenant Search: name or descr	common HCl1_tenant1 u	iser1 tn-hshahane testB	3R012					
HCI1_tenant1	00	Node Association							0
> C+ Quick Start	1					Dellas	E		w
✓ HCI1_tenant1						Policy	Faults	HISTO	Jry
> 🚞 Application Profiles							Ō	+	***-
V 🚞 Networking		Properties							
> 🚞 Bridge Domains		Node ID:	topology/pod-1/node-101						
> 🚞 VRFs		Router ID:	1.1.1.1						
> 🚞 L2Outs		Use Router ID as Loopback Address:	This setting will be ignored if loopback ad	dresses are defined in the table below.					
V 🚞 L3Outs		Loopback Addresses:						11 -	£1.
VNET01_L3Out			▲ IP						п.
🗸 🚞 Logical Node Profile	25		10.10.10.10						
VNET01_NP									н.
E BGP Peer 192	2.168.1.2								
Configured N	lodes								
✓	pod-1/node-101	Static Routes:						÷ 1	+
F ARP fo	r VRF-HCI1_tenant1:VRF1		 IP Address 	Description	Track Policy	Next Hop IP			
> 🗧 BGP fo	r VRF-HCI1_tenant1:VRF1				No items have been found.				
> \Xi ND for	VRF- HCI1_tenant1:VRF1				Select Actions to create a new item.				
> 🗾 topology/	pod-1/node-102								
> 🚞 Logical Interf	ace Profiles								
— External EPGs —— —— —— —— —— —— —— —— —— —— —— —— ——									
VNET01_EXT_EP	G								
> 🚞 Route map for impo	rt and export route control		C Rago 0 0f0		Objects Per Pages 15		No Object	te Four	
SR-MPLS VRF L3Outs									
> 📩 Dot1Q Tunnels						Show Usage Re			
> 🚞 Contracts									

- 35. Add the **Gateway Subnet** in the **Prefix** field (In this example, **192.168.1.0/29** is the gateway subnet. Please note that the gateway subnet is part of the VNET subnet).
- 36. Click + to add the **Logical IP address** of the Azure Stack HCI VNET in the **Next Hop Addresses** field (In this example, **10.10.1.6**).

Create Static R	oute		\times	
Prefix:	192.168.1.0/29			
Description:	optional			
Fallback Preference:	1	\Diamond		
Nexthop Type:	Static Route			
Route Control:	BFD			
Track Policy:	select an option	\sim		
Next Hop Addresses:			1 +	
	Next Hop IP	Preference		
	10.10.1.6	0		
	If there is no next hop address added, will be automatically created.	a NULL interface Cancel Subr	mit	

- 37. Click Submit.
- 38. Navigate to Tenants > HCl1_tenant1 > Networking > L3Outs > L3Out Name (In this example, VNET01_L3Out) > Logical Node Profiles (In this example, VNET01_NP) > Configured Nodes > Node path (In this example, topology/pod-1/node-102).

- 39. Repeat steps 34 to step 37 to add a static route on the second node.
- 40. External EPG can be created via wizard as shown in step 23. It can also be created from the following path Tenants > HCl1_tenant1 > Networking > L3Outs > L3Out Name (In this example, VNET01_L3Out) > External EPGs > External EPG Name (In this example, VNET01_EXT_EPG).

HCl1_tenant1	External EPG - V	NET01_EXT_EPG						0
> C Quick Start						Delieu	Liestin Coulte	Linter
V III HCI1_tenant1						Policy Operational	Health Faults	History
> 🧮 Application Profiles				General	Contracts	Inherited Contracts	Subject Labels EP	G Labels
V 🖿 Networking					•		¢	1 44
> 🔚 Bridge Domains							0	- ×*
> 🔤 VRFs	Properties pcTag:	32770						
> 🔤 L2Outs	Contract Exception Tag:							
V 🖿 L3Outs	Configured VRF Name:	VRF1						
✓	Resolved VRF:	uni/tn-HCI1_tenant1/ctx-V	RF1					
> 🚞 Logical Node Profiles	QoS Class:	Unspecified						
🗸 🚞 External EPGs	Target DSCP:	Unspecified						
VNET01_EXT_EPG	Configuration Status:	applied						
> Route map for import and export route control	Configuration Issues:							
> 🔤 SR-MPLS VRF L3Outs	Preferred Group Member:	Exclude Include)					
> 🚞 Dot1Q Tunnels	Intra Ext-EPG Isolation:	Enforced Unenfor	rced					
> 🧮 Contracts	Subnets							
> E Policies		+ ID Address	Soone	Name	Aggregate	Pouto Control Bro	lle Poute Summarizati	-T
> 🧰 Services		- IP Address	acope	Name	Aggregate	Route Control Pro	Policy	011
E Security		0.0.0.0/0	Export Route Control Subnet					
		192.168.1.0/2.4	External Subnets for the External EPG					
						Show Usag	e Reset S	

<u>Configure Contracts</u> as discussed in the previous sections. Contracts are necessary to permit traffic between the L3Out external EPG and other L3Out External EPGs or EPGs that are part of the ACI fabric.

Contracts can be added to the External EPGs from the following path - Tenants > HCl1_tenant1 > Networking > L3Outs > L3Out Name (In this example, VNET01_L3Out) > External EPGs > External EPG Name (In this example, VNET01_EXT_EPG) > Policy > Contracts > Add Provided Contract or Add Consumed Contract.

HCl1_tenant1	00	External EPC	G - VNET01_EXT	_EPG								0
> O• Quick Start								Policy	Operational	Health Fa	ults History	
> Application Profiles							General	Contracts Inh	erited Contracts	Subject Labels	EPG Labels	_
✓	- L .									,		
> 🚞 Bridge Domains		Healthy 🙁 🕚								_	0 ± %•	*
> 🖿 VRFs		Name	 Tenant 	Tenant Alias	Contract Type	Provided / Consumed	QoS Clas	ss State	Label	Add Provide	ed Contract	
> 🖿 L2Outs						No items have been fr	hund			Add Consur	ned Contract	
V 🖿 L3Outs						Select Actions to create a	new item.			Add Consur	med Contract Interface	28
✓ VNET01_L3Out										Taboo Cont	ract	
> 🚞 Logical Node Profiles										Add Intra Ex	.t-EPG Contract	
V 🚞 External EPGs										Delete		
VNET01_EXT_EPG	4											
Route map for import and export route control												
> SR-MPLS VRF L3Outs												
> Contraction Dot 10 Tunnels												
> Contracts												
> Policies												
> Services												
Security												
Ear more informat	ion											

For more information http://www.cisco.com/go/aci

Revision history

Revision	Coverage	Date		
Initial version	 Microsoft Azure Stack HCI 22H2 Cisco ACI Release 6.0(3e) Cisco NX-OS Release 12.1.3b 	12/19/2023		
Added Appendix <u>Design Example with</u> <u>Microsoft Software Defined Networking</u> (SDN) in Azure Stack HCI	Microsoft Azure Stack HCI 22H2Cisco ACI Release 6.0(3e)	07/12/2024		