Cisco APIC Container Plug-in Release 5.0(2), Release Notes

This document describes the features, bugs, and limitations for the Cisco Application Policy Infrastructure Controller (APIC) Container Plug-in.

The Cisco Application Centric Infrastructure (ACI) Container Network Interface (CNI) Plug-in provides network services to Kubernetes, Red Hat OpenShift, and Docker EE clusters on a Cisco ACI fabric. It allows the cluster pods to be treated as fabric end points in the fabric integrated overlay, as well as providing IP Address Management (IPAM), security, and load balancing services.

Release notes are sometimes updated with new information about restrictions and bugs. See the following website for the most recent version of this document:

https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html

Date	Description
2022-2-28	Updated the Supported Scale table.
2020-10-26	Adds CSCvv87025 and CSCvv90680 to the table of resolved bugs.
2020-09-10	Removed limitation saying that the Cisco ACI CNI 5.0(2) release is not tested for OpenShift 3.11 and recommending that deployments running OpenShift 3.11 continue using the Cisco ACI CNI plugin from 4.2(2) release.
2020-09-02	Revised usage guideline for upgrading a Kubernetes cluster. Added information about deploying Kubernetes flavors and corrected file to be updated.
2020-08-31	Adds open and resolved bugs.
2020-07-16	Release 5.0(2) became available.

Contents

This document includes the following sections:

Cisco ACI Virtualization Compatibility Matrix

For information about Cisco ACI supported Container Products, see the *Cisco ACI Virtualization Compatibility Matrix* at the following URL:

https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/aci/virtualization/matrix/virtmatrix.html

New and Changed Information

This section lists the new and changed features in this release.

New Software Features

Kubernetes 1.17

Cisco ACI supports Kubernetes 1.17.

Istio 1.5.2

Cisco ACI supports the deployment of Istio control plane using the upstream community supported Istio Operator.

Note: This is a preview feature and is rapidly evolving in response to upstream changes.

Cisco CNI Operator

Beginning with this release, Cisco ACI CNI deployments are managed by a Kubernetes Operator.

Destination-based Egress SNAT

Multiple SNAT policies can be associated with a Pod such that the SNAT IP is allocated based on the destination. SNAT can also be completely suppressed to known destinations on the same fabric.

Support for VMware Teaming Policy

You can configure VMware teaming policy when link aggregation groups (LAGs) are used. For more information, see Cisco ACI and Kubernetes Integration.

Mixed Form-Factor Support for Upstream Kubernetes

You can deploy a Kubernetes cluster with the Cisco ACI CNI plug-in on a mixed collection of virtual machines (VMs) for the master nodes, and bare-metal servers for the worker nodes.

Support for Adding a Kubernetes Cluster to an Existing Tenant

You can add a Kubernetes cluster to an existing tenant—such as the common tenant in Cisco APIC. You do so by modifying the configuration file. For more information, see Cisco ACI and Kubernetes Integration.

Universal Base Image (UBI) 8

Container images are now based on Red Hat Universal Base Images instead of Alpine base images which were used in earlier releases.

Changes In Behavior

- The Cisco ACI CNI deployments and daemonsets are now deployed by the Cisco ACI CNI Operator, and the lifecycle of these resources is subsequently managed by this Operator. The user workflow does not change as this change is captured in the Kubernetes deployment file that is generated by the acc-provision tool, and that deployment now deploys the Cisco ACI CNI Operator.
- The Cisco APIC resources created by acc-provision for Kubernetes clusters now contain the prefix “aci-containers-“ instead of “kube-“. Some resource names, for instance endpoint group (EPG) names, also contain the system-id. Running acc-provision will always result in this new naming convention unless it is explicitly configured to use the older naming convention when upgrading existing clusters. Please refer to the Usage Guidelines for `use_legacy_kube_naming_convention` to use older naming convention.
- The egress SNAT feature has gone through significant enhancements by decentralizing the computation previously performed by the SNAT operator container and delegating it to the host agents. This resulted in the elimination of the SNAT operator as a separate container. Also, starting in the 5.0(2) release, the `snatlocalinfos` custom resources is a purely operational resource made available to the user to get the SNAT policies and SNAT IP addresses associated with pods. Any tools or scripts written to reference the `snatlocalinfos` resource in the 4.2(2) release will need to be updated accordingly.
- The use of acc-provision will, by default, install Istio control plane (version 1.5.2) on the cluster using a “demo” configuration profile. This installation is driven from aci-containers-controller pod. Istio control plane pods are brought up in “istio-system” namespace and are isolated in “aci-containers-istio” EPG. The constructs to achieve this isolation such as contracts, filters, and contract-relationships are automatically configured on the Cisco APIC.
- The scope of the SNAT service graph contract can be configured by the user. Please refer to the Usage Guidelines to perform this configuration.

- The "annotation" property for Cisco APIC objects created by acc-provision is set to `orchestrator:aci-containers-controller`. This is also reflected in a unique icon on the Cisco APIC objects in the Cisco APIC GUI.
- `acikubectl` has been enhanced to collect relevant config maps.

Known Limitations

- Istio 1.5.2 deployment is a preview feature and is expected to rapidly evolve in the next release in response to upstream community changes. This may create issues with backward compatibility and as such this feature should only be used in experimental or pilot deployments.
- OpenShift 4.3 on AWS does not currently support policy enforcement in a hybrid/multi-site deployment.
- The Cisco ACI CNI Plug-in is not integrated with the Multi-Site Orchestrator. When deploying to a Multi-Site deployment, the Cisco ACI configurations implemented by the plug-in must not be affected by the Multi-Site Orchestrator.
- SNAT is not supported for services inside the same fabric.
- The following issues are specific to the 5.0(1) release and are fixed in the 5.0(2) release:
  - The acikubectl executable is not packaged correctly in the dist-rpms-5.0.1.0.tar.gz file in the 5.0(1) release. Use the dist-rpms-5.0.2.0.tar.gz file from the 5.0(2) release to install the acikubectl executable.
  - The SNAT policy associated with a pod is lost if the aci-containers-host pod is restarted on that node.
  - When a packet is bounced in the SNAT reverse path, the router's MAC address is being preserved, This results in Cisco ACI learning that MAC address on different leaf interfaces if more than one node bounces the packets. Also, eventually a fault will be raised indicating the endpoint is in freeze state. Starting with 5.0(2) release, the MAC address of the interface on the node bouncing the packet will be set as the source MAC address of the bounced packet.
  - The SNAT IP allocated to a pod and other operational information relevant to the SNAT feature can be obtained by:

    ```
    kubectl describe snatlocalinfos.aci.snat -n aci-containers-system
    ```

  - Filter-chains are now used to configure the SNAT service-graph in the APIC. This prevents RST packets sent by the destination from being dropped (earlier only ACK packets were allowed on account of the filter configuration used). This change is available only when using ACI 4.2(4i) or later.

Usage Guidelines

- If you are upgrading a Kubernetes cluster that was provisioned with ACI CNI version 4.2(2), or you are deploying a new cluster with kubernetes-1.15 or kubernetes-1.16 flavors, you need to add the following configuration to the original acc-provision input file:

  ```
  aci_config:
    use_legacy_kube_naming_convention: True
  ```

  Also note that upgrading a Cisco ACI CNI cluster requires running acc-provision with the "-a" option.

- Istio installation can be disabled by setting the config parameter “install-istio” to False in the acc-provision-input file and generate/apply the deployment file.
- The scope of the SNAT service graph contract can be configured by the user in the acc-provision input file as follows:

  ```
  kube_config:
    snat_operator:
      contract_scope: <scope name>
  ```

  Valid values (as allowed by Cisco APIC) are "global", "tenant" and "context". The default is set to "global".
- ACC subscribes for notifications on certain objects to the Cisco APIC. There is a timeout associated with this subscription. A shorter timeout requires more frequent subscription renewals. The timeout is set to 900 seconds for Cisco APIC 4.x and can be changed by configuring the acc-provision input file:

  ```
  aci_config:
    apic_refreshtime: 1200
  ```

  Note: The subscription timeout is configurable only in Cisco APIC 4.x.
- The memory limit for the Open vSwitch container is set to 1GB. It can be changed by configuring the `acc-provision` input file as follows:

  ```
  kube_config:
    ovs_memory_limit: 5Gi
  ```
- Policy Based Routing (PBR) tracking can be enabled for the Cisco APIC service graph created for supporting the SNAT feature. More details on PBR tracking can be found in the chapter “Configuring Policy-Based Redirect” In the *Cisco APIC Layer 4 to Layer 7 Services Deployment Guide, Release 4.2(x)*.

  One HealthGroup for each node is created, and it is associated with the redirect policy of the SNAT service graph with the internet protocol service level agreement (IP SLA) interval set to 5 seconds. This interval is configurable through the acc- provision input file:

  ```
  net_config:
    service_monitor_interval: 10
  ```

  If the service_monitor_interval is set to zero, PBR tracking is disabled.

  PBR tracking can be also be enabled for other Cisco APIC service graphs created for each Kubernetes external service, setting the following configuration in the `acc-provision` input file:

  ```
  net_config:
    pbr_tracking_non_snat: true
  ```

  If enabled, the service_monitoring_interval described earlier applies here as well.

  Note that in a Cisco ACI CNI-based cluster, the same worker node is used to provide both the external Layer 4 load balancer and SNAT services. So if PBR tracking is enabled, and if the worker node reports unhealthy status for SNAT, a fault appears in the redirect policies associated with all other (non-SNAT) service graphs that have this node. However, this fault does not actually affect those other services and traffic from those services is still distributed to that node. The fault manifests for those other services only in the Cisco APIC GUI.

Note: The following are general usage guidelines that also apply to this release:

- The Cisco ACI CNI Plug-in is supported with the following container solutions:
  - Canonical Kubernetes on Ubuntu 18.04

  - Red Hat OpenShift on Red Hat Enterprise Linux 7

- You should be familiar with installing and using Kubernetes or OpenShift. Cisco ACI does not provide the Kubernetes or OpenShift installer. Refer to the following documents on Cisco.com for details:
  - *Cisco ACI and Kubernetes Integration*
  - *Cisco ACI and OpenShift Integration*
  - *Cisco ACI CNI Plugin for Red Hat OpenShift Container Platform Architecture and Design Guide*
  - *Upgrading the Cisco ACI CNI Plug-in*
- The Cisco ACI CNI plug-in implements various functions running as containers inside pods. The released images for those containers for a given version are available on the Docker Hub website under user noiro. A copy of those container images and the RPM/DEB packages for support tools (`acc-provision` and `acikubectl`) are also published on the Software Download page on Cisco.com.
- OpenShift has a tighter security model by default, and many off-the-shelf Kubernetes applications, such as guestbook, may not run on OpenShift (if, for example, they run as root or open privileged ports like 80).
- Refer to the article “Getting any Docker image running in your own OpenShift cluster” on the Red Hat OpenShift website for details. The Cisco ACI CNI Plug-in is not aware of any configuration on OpenShift cluster or pods when it comes to working behind a proxy. Running OpenShift "`oc new-app`," for instance, may require access to Git Hub, and if the proxy settings on the OpenShift cluster are not correctly set, this access may fail. Ensure your proxy settings are correctly set.
- In this release, the maximum supported number of PBR based external services is 250 virtual IP addresses (VIPs). Scalability is expected to increase in upcoming releases.

  Note: With OpenShift, master nodes and router nodes are tainted by default, and you might see lower scale than an upstream Kubernetes installation on the same hardware.
- Some deployments require installation of an “allow” entry in IP Tables for IGMP. This must be added to all hosts running an OpFlex agent and using VXLAN encapsulation to the leaf. The rule must be added using the following command:

  ```
  $ iptables -A INPUT -p igmp -j ACCEPT
  ```

  In order to make this change persistent across reboots, add the command either to `/etc/rc.d/rc.local` or to a cron job that runs after reboot.
- Both RHEL and Ubuntu distributions set net.ipv4.igmp_max_memberships set to 20 by default. This limits the number of end point groups (EPGs) that can be used in addition to the kube-default EPG for pod networking. If you anticipate using more than 20 EPGs, set the value to the desired number of EPGs on each node as follows:

  ```
  $ sysctl net.ipv4.igmp_max_memberships=desired_number_of_epgs
  ```
- For the VMware VDS integration, you can refer to the Enhanced Link Aggregation Group (eLAG) configured through the Cisco APIC by using the following configuration in the acc-provision input file:

  ```
  nested_inside:
   type: vmware
  …
  elag_name: <eLAG-name-used>
  ```

Supported Scale

The Kubernetes, OpenShift, Cloud Foundry, and Pivotal Cloud Foundry Platform scale limits are shown in the following table:

Table 3: Supported Scale Limits

Limit Type	Maximum Supported
Nodes/Leaf (or OpFlex hosts per leaf)	120[1]
Nodes/interface on Leaf (or OpFlex hosts per port)	20
VPC links/Leaf	40
Endpoints[2]/Leaf	10000
Endpoints/Host	400
Virtual endpoints[3]/Leaf	40,000

1- The indicated scale value is for Cisco ACI version 5.0(1) and later. If the ACI version is less than 5.0(1), the number of supported OpFlex hosts are 40.

2- An endpoint corresponds to a Pod's network interface.

3- Total virtual endpoints on a leaf can be calculated as: Virtual endpoints / leaf = VPCs x EPGs where:

- VPCs is the number of VPC links on the switch in the attachment profile used by the OpenStack Virtual Machine Manager (VMM).
- EPGs is the number of EPGs provisioned for the OpenStack VMM.

For the CLI verified scalability limits, see the *Cisco NX-OS Style Command-Line Interface Configuration Guide* for this release.

Bugs

This section contains lists of bugs and known behaviors.

Open Bugs

Table 4: Open bugs in the 5.0(2) release

Bug ID	Description
CSCvv47774	Stale RD policy in opflexp policy cache with multiple L3outs attached to VRF
CSCvu58070	Incorrect PBR connector class ID for inter-VRF contract

Resolved Bugs

Table 5: Resolved bugs in the 5.0(2) release

Bug ID	Description
CSCvv04675	ACI CNI 4.2.2.2 IPAM exhaustion
CSCvv23505	Change Adjacency type to L3 for ACI CNI-created SG template
CSCvv87025	Constant OVS reprograming after Docker EE CNI upgrade from 4.1.1 to 5.0.2
CSCvv90680	ACI/OpenShift: CNI plugin - aci-containers-host pod is restarted
CSCvq96281	aci-container-controllers: Add error message if VRF config is inconsistent
CSCvt02599	Add support for additional Local Subnets for SNAT
CSCvs41278	SNAT is applied to all Namespaces
CSCvs41252	SNAT for Services with Multiple ports does not work
CSCvs61926	acc-provision: remove extern_static for OpenShift Flavor
CSCvu92591	K8s ACC fails to push changes to APIC cluster after observed panic

Known Behaviors

This section lists known behaviors. Click the Bug ID to access the Bug Search Tool and see additional information about the bug.

Table 5: Known Behaviors in the 5.0(2) release

Bug ID	Description
CSCvm66785	Containers IP is shown as learned on all cluster interfaces.

Related Documentation

The Cisco Application Policy Infrastructure Controller (APIC) documentation can be accessed from the following website:

https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products- support-series-home.html

The documentation includes installation, upgrade, configuration, programming, and troubleshooting guides, technical references, release notes, and knowledge base (KB) articles, as well as other documentation. KB articles provide information about a specific use case or a specific topic.

By using the "Choose a topic" and "Choose a document type" fields of the Cisco APIC documentation website, you can narrow down the displayed documentation list to make it easier to find the desired document.