Cisco Application Policy Infrastructure Controller Container Plug-in Release 4.2(3), 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

Note: There is no Cisco ACI CNI version 4.2(3). However, you need Cisco APIC 4.2(3) to take advantage of some extra capabilities. If you use Cisco APIC 4.2(3), all the switches in the fabric need to run version 14.2(3).

Table 1 shows the online change history for this document.

Table 1: Online History Change

Date	Description
2021-11-24	Updated the Known Behaviors list.
2020-03-30	Added information in New Software Features section about bug and feature relevant only when using Cisco APIC 4.2(3).
2020-01-14	Added issue affecting Kubernetes Network Policy to Known Limitations section.
2020-01-13	Moved CSCvn13789 from Known Behaviors table to Resolved Bugs table.
2019-12-13	Release 4.2(3) 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 and includes the following topics:

New Software Features

The following are the new software features for this release:

- Distributed Source Network Address Translation (SNAT) for Egress Traffic enables you to configure and manage NAT with the Cisco ACI CNI Plug-in from inside the Cisco ACI fabric.
- Docker EE 3.0 with UCP 3.2 on Red Hat 7.6 and 7.7. See *Cisco ACI and Docker EE Integration* on Cisco.com.
- Support for Kubernetes 1.14, 1.15, and 1.16.
- Container base image upgrade to Alpine 3.10.2.

Note: The SNAT feature requires OVS NAT support, as documented in the Open vSwitch Release FAQs on the Open vSwitch website. It has been tested to work with RHEL 7.6 and Ubuntu 18.04.

The following are relevant only if you use Cisco ACI CNI with Cisco APIC 4.2(3):

- Resolved bug: CSCvn13789–The Cisco ACI CNI plug-in does not support N/S load-balancer for pods hosted on UCS-B with FI connectivity or for VMs in nested mode that can vMotion.
- eLag support: For the VMware VDS integration, refer to the Enhanced Link Aggregation Group (eLAG) configured through Cisco APIC by using the following configuration in the acc-provision input file nested inside:

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

Changes In Behavior

This section lists changes in behavior in this release.

- The Cisco APIC Release 3.2 introduced using annotations for Cisco APIC objects instead of using tags (wherever applicable) for scalability reasons. Until Cisco APIC Release 4.1(x), The Cisco ACI Containers Controller (ACC) tagged the objects that it was creating in Cisco APIC.

  In this release, we have moved to applying annotations instead of tags. When ACC starts, Cisco APIC is queried to get the current version. Based on the version, it determines whether to use the tagInst (for tags) or tagAnnotation (for annotations) objects. When you upgrade from Cisco APIC Release 3.2(x) to 4.2(x), the migration to tagAnnotation occurs automatically. Also, existing tagInst objects are removed. If the Cisco APIC version is earlier than Release 3.2(x), no changes occur to the existing tags. If Cisco APIC is downgraded from one version of Release 3.2(x) to an earlier version of 3.2(x), ACC automatically restarts and reverts to using tags. Cisco APIC versions earlier than Release 3.2(x) are not supported for Cisco ACI CNI release 4.2.

Note: This change is automatic and does not require any action on part of the user.

- Memory leaks have been observed in OVS 2.10. There are several discussions in the upstream community around those. See the article “ovs-vswitchd: possible memory leak in 2.10.1 (Alpine 3.9)” on the bugs.alpinelinux.org Website and the article “ovs-vswitchd: possible memory leak in 2.10.1 #168” on the GitHub Open vSwitch website.

  To work around this issue, the memory resource limit on the OVS container is restricted to 1GB. If the OVS container memory consumption hits this 1GB limit, Kubernetes automatically restarts it. (No action is required on the part of the user.) The 1GB resource limit is configurable; its configuration is explained in the Usage Guidelines section of this document. Note that the OpFlex Agent container is automatically restarted when the OVS container is restarted.

- When you upgrade from an older Cisco ACI CNI release, you must generate a new Cisco ACI CNI Kubernetes deployment file using `acc-provision`. See *Upgrading the Cisco ACI CNI Plug-in* on Cisco.com

- Cloud Foundry and Pivotal Cloud Foundry are no longer supported.

- Ubuntu 16.04 is not supported.

Known Limitations

This section lists the known limitations:

- When downgrading from Cisco ACI CNI-plugin 4.2(1) to an older release, the vxlan_sys_8472 interface on the host should be first deleted. This interface is recreated when the host-agent pod restarts.

- OpenShift has the following issues:

  — The 3.10 installation has open issues regarding proxy configuration. See the article “Failed to deploy 3.10 when using proxy” on the GitHub website.

  `Pod /etc/resolv.conf` is populated with an incorrect nameserver, possibly with an IP address read from the wrong interface. See the article “Bug 1680059 - [3.11] Pod /etc/resolv.conf populated with incorrect nameserver possibly with an IP read from the wrong interface” on the Bugzilla website and the article “[release-3.11] Added more conditions to set a reasonable DNSIP #21866” on the GitHub website.

Use the suggested workaround in the issue description.

- The Cisco ACI CNI pods are marked with “priorityClassName: system-node-critical” to prevent them from being first candidates for eviction. However, in Kubernetes 1.11 and earlier, only pods belonging to the kube-system namespace can be marked in this way. Consequently, in older Kubernetes versions, or any distributions that rely on those versions (including OpenShift 3.11 and earlier), Cisco ACI CNI pods might get evicted before other application pods.
- 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 combination `podSelector` and `namespaceSelector` with an AND (intersection) operation is not supported for Kubernetes Network Policy ingress or egress sections.

Usage Guidelines

- 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.

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. The Cisco CNI plug-in and the corresponding deployment file are provided to enable networking for an existing installer such as such as kubeadm, Kubespray, or openshift-ansible. 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.
- For OpenShift, the external IP address used for the LoadBalancer service type is automatically chosen from the subnet pool specified in the ingressIPNetworkCIDR configuration in the `/etc/origin/master/master-config.yaml` file. This subnet should match the extern_dynamic property configured in the input file provided to the `acc_provision` script. If a specific IP address is desired from this subnet pool, it can be assigned to the "loadBalancerIP" property in the LoadBalancer service specification. For more details, see the article “Configuring the Cluster to Use Unique External IPs” on the Red Hat OpenShift website.

  Note: The extern_static subnet configuration in the `acc_provision` input is not used for OpenShift.

- 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 Aggregration Group (eLAG) configured via the 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 Table 2:

Table 2 Supported Scale Limits

Limit Type	Maximum Supported
Nodes/Leaf	40
VPC links/Leaf	40
Endpoints[1]/Leaf	4000
Endpoints/Host	400
Virtual endpoints[2]/Leaf	40,000

[1] An endpoint corresponds to a Pod’s network interface.

[2] 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.

- Resolved Bugs
- Known Behaviors

Resolved Bugs

This section lists the resolved bugs for the release. Click the bug ID to access the Bug Search Tool and see additional information about the bug.

Table 3 Resolved bugs in the 4.2(3) Release

Bug ID	Description
CSCvn13789	Cisco ACI CNI plug-in does not support N/S load-balancer for pods hosted on UCS-B with FI connectivity or for VMs in nested mode that can vMotion.
CSCvr99112	Evicted aci-container-controller does not release lock
CSCvr91947	dockerEE: Missing loopback in /opt/cni/bin
CSCvr88382	acikubectl: cluster report does not work on Kubernetes
CSCvr67770	Cisco ACI-CNI: Shared services do not work
CSCvq63491	Cisco ACI CNI plug-in on Kubernetes/OpenShift Open vSwitch memory leak
CSCvq19984	The acc deletes L3Out contract relationships if it loses connectivity to Cisco APIC during refresh API Call
CSCvp86156	Cisco ACI CNI: aci-containers-controller: ApicRefreshTimer parameter format is invalid
CSCvp26247	Cisco ACI CNI: opflex-agent should ignore Headless Services
CSCvo67047	Docker-ucp flavor should open additional ports
CSCvo58994	ACI CNI not cleaning up L4-7 devices in APIC after worker nodes are removed

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 4 Known Behaviors in the 4.2(3) Release

Bug ID	Description
CSCvm66785	Containers IP is shown as learned on all cluster interfaces.
CSCvy17504	OpflexP does not trigger opflex disconnect until ARP entry expired.
CSCvz17367	opflexODev object points to old compute node after vmotion.
CSCvz98577	Remote EP relationships are not updated after VM migration.

Bug ID	Description
CSCwa22996	intra cluster communication broken in OpenShift after upgrade.
CSCwa03344	Immediately update Remote EP relationships after VM migration.

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.