Cisco APIC Container Plug-in Release Notes, Release 5.2(1)

Introduction

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, Rancher RKE, 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

For more information about this product, see " Related Content."

Note: The documentation set for this product strives to use bias-free language. For the purposes of this documentation set, bias-free is defined as language that does not imply discrimination based on age, disability, gender, racial identity, ethnic identity, sexual orientation, socioeconomic status, and intersectionality. Exceptions may be present in the documentation due to language that is hardcoded in the user interfaces of the product software, language used based on RFP documentation, or language that is used by a referenced third-party product.

Date	Description
February 28, 2022	Updated scale details for OpFlex hosts per leaf.
July 22, 2021	Added information for OpFlex hosts per port. See *Supported Scale* section.
June 7, 2021	Release 5.2(1) became available.

New Software Features

Feature	Description
Support for Kubernetes 1.21	Cisco ACI supports Kubernetes 1.21 using the Cisco ACI Container Network Interface (CNI) plug-in and installed with kubeadm on Ubuntu 20.04.
Support for OpenShift 4.6 in OpenStack Red Hat OpenStack Platform (OSP) 16.1	Cisco ACI supports Red Hat OpenShift 4.6 nested in Red Hat (OSP) 16.1. To enable this support, Cisco ACI provides customized Ansible modules to complement the upstream OpenShift installer. For more information, see *Installing OpenShift 4.6 on OpenStack 16.1* guide.
Support for OpenShift 4.6 on VMware vSphere	Cisco ACI supports Red Hat OpenShift 4.6 nested in VMware vSphere 7. To enable this support, Cisco ACI provides customized Ansible modules to complement the upstream OpenShift installer. For more information, see *Installing OpenShift 4.6 on VMware vSphere* guide.
Support for Session Affinity in Kubernetes Services	Cisco ACI supports passing connections from a particular client to the same Pod each time based on the client's IP address and a configurable session stickiness time.

Feature	Description
Topology-aware Routing of Service Endpoint Traffic	Cisco ACI supports a Kubernetes service to route traffic based upon the Node topology of the cluster using Topology Keys. Cisco APIC Release 5.2(1) provides a preview of this feature and it is expected to evolve to use Topology-aware Hints in the upcoming releases.

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

Supported Scale

For the verified scalability limits (except for CLI limits), see the Verified Scalability Guide for this release. For Kubernetes-based Integrations (including Docker, OpenShift, and Rancher), and OpenStack Platform Scale Limits, see the following table.

Note: The scalability information in the following table applies to Kubernetes or OpenStack resources integrated with OpFlex into the Cisco ACI fabric. It does not apply to Microsoft SCVMM hosts or Cisco ACI Virtual Edge instances.

Limit Type	Maximum Supported
Number of OpFlex hosts per leaf	120[1]
Number of OpFlex hosts per port	20
Number of vPC links per leaf	40
Number of endpoints per leaf	10,000
Number of endpoints per host	400
Number of virtual endpoints per 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.

Notes:

- For containers, an endpoint corresponds to a pod's network interface.
- For OpenStack, an endpoint corresponds to any of the following:
  - A virtual machine (VM) interface (also known as vnic)
  - A DHCP agent's port in OpenStack (if in DHCP namespace on the network controller)
  - A floating IP address
- 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 OpFlex hosts per port — a port is either a physical port or a vPC. One vPC equals one port.The number of member ports in a vPC is inconsequential.
- For the CLI verified scalability limits, see the *Cisco NX-OS Style Command-Line Interface Configuration Guide* for this release.

Known Limitations

- The NodePort service statistics exported to Prometheus get accounted under ClusterIp service statistics in on-premise deployments.
- A pod selector has to be always provided to map a port name to the port number, and an empty pod selector is not supported in the ingress direction.
- Istio 1.6.5 deployment and Topology-aware traffic routing are preview features, that are expected to rapidly evolve in the upcoming release(s) in response to upstream community changes. This may create issues with backward compatibility, and as such these features should only be used in experimental or pilot deployments.
- 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 policy configuration is not applicable to traffic within the same cluster.
- Due to Python 3 dependencies that are currently available only on RHEL8, acc-provision tool is supported on RHEL8 operating system, but not on RHEL7 operating system.

Usage Guidelines

- Note that upgrading a Cisco ACI CNI cluster requires running acc-provision with the " --upgrade" option.
- Metrics data for various features is pulled and exposed from opflex-agent at different timer intervals. In a steady state, statistics will become eventually consistent. However, if traffic is flowing continuously, there could be slight discrepancy when comparing statistics of various exported metrics. For example, the POD and Service counts are pulled at different intervals of time. While checking these at a particular instant of time can show variations, once in a steady state the statistics will be eventually consistent.
- It is recommended that the in-cluster Prometheus server have a low retention time since Prometheus doesn't support horizontal auto-scaling. In general, there should be a federated Prometheus server with long term storage solution to avoid hogging CPU and memory of a cluster.
- To enable drop logging, perform the following configuration in the acc-provision input file:

```
drop_log_config:
    enable: True
```

  For more information, see Enabling the OpFlex Drop Log Feature.

- 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".

- The aci-containers-controller pod 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, 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 or later.

- 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
```

- The Multus CNI deployment can be enabled in the OpenShift installation by performing the following configuration in the acc-provision input file:

```
multus:
  disable: False
```

- 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 5.2(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: 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.

- Starting with Cisco APIC Release 5.2(1), a fault (vmmClusterFaultInfo) is generated in ACI, if a Kubernetes namespace, deployment, or pod is annotated with an EPG name that does not resolve to an existing EPG. A log statement is added in the aci-containers-controller log to alert the user. The fault will be cleared upon the next correct annotation, or when the aci-containers-controller restarts, or when the annotated namespace, deployment, or pod is deleted.
- 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*
  - *Installing OpenShift 4.6 on OpenStack 16.1*
  - *Installing OpenShift 4.6 on VMware vSphere*
  - *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>
```

Open Issues

There are no open issues in this release.

Resolved Issues

Click the bug ID to access the Bug Search tool and see additional information about the bug.

Bug ID	Description
CSCvy21069	Fix possible snatglobalinfo cache inconsistencies.
CSCvx48097	acc-provision blindly setting gateway on apic.json to .1.
CSCvy36279	Add defensive code to recover from node information missing from globalsnatpolicy CR.

Known Issues

There are no known issues in this release.

Related Content

See the Cisco Application Policy Infrastructure Controller (APIC) page for the documentation.

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 APIC documentation website, you can narrow down the displayed documentation list to make it easier to find the desired document.

You can watch videos that demonstrate how to perform specific tasks in the Cisco APIC on the Cisco Data Center Networking YouTube channel.

Documentation Feedback

To provide technical feedback on this document, or to report an error or omission, send your comments to apic-docfeedback@cisco.com. We appreciate your feedback.

Legal Information