



# Cisco Network Operator- VMM-Lite User Guide

## Cisco Network Operator - VMM-Lite for OpenShift virtualization

This guide provides instructions and information about how to use CNO VMM-Lite for OpenShift virtualization, a feature designed to provide a similar workflow to OpenShift environments for customers who migrate from VMware + ACI VMM integration, and to address Green Field and Brown Field customers who cannot use the current CNO model where CNO pushes configuration to ACI.

### Overview

CNO VMM-Lite dynamically creates Network Attachment Definitions (NADs) in an OpenShift cluster by monitoring configuration changes to Attachable Access Entity Profiles (AAEPs) and End-Point Groups (EPGs) attached to them in Cisco ACI. This allows for automated network provisioning for virtual machines (VMs) running on OpenShift Virtualization.

### Key concepts

1. **Network Attachment Definition (NAD):** A Kubernetes custom resource that defines how a pod or VM can connect to a network.
2. **Attachable Access Entity Profile (AAEP):** An ACI construct that defines the physical infrastructure (ports, VLANs) that EPGs can use.
3. **End-Point Group (EPG):** An ACI constructs that groups endpoints (e.g., VMs, bare-metal servers) that share common policy.
4. **CNO (Cisco Network Operator):** The operator responsible for managing network configurations in OpenShift, interacting with ACI.

## Installation of VMM Lite

### Prerequisites

1. **OpenShift configuration**
  - Worker node bonds: Create network bonds on all OpenShift worker nodes that will host VM traffic.
  - Linux bridge: Create a Linux bridge with the same name on all OpenShift worker nodes, using the bond created previously.
  - Trunk VLANs: Ensure VLANs are trunked on the bridge uplinks.
2. **ACI configuration**
  - Dedicated AAEP: Place the OpenShift Node uplink interface used for VM traffic in one dedicated AAEP.

### VMM Lite Installation

1. **Install acc-provision on a host that has access to APIC.**

```
pip install acc-provision
```

2. **Prepare YAML file named “acc-provision-input.yaml”.** Please refer example “acc-provision-input.yaml” file for reference.

```
aci_config:
  # List of all IP addresses of the APIC Controllers
  apic_hosts:
```

```
- 10.30.120.61
logging:
  controller_log_level: debug
vmm_lite_config:
  aaep_monitoring_enabled: true
  bridge_name: "bridge-net-1"
  cno_identifier: "cno-dev"
  apic_username: "ocpbm3"
  bridge_nad_config_file: "bridge_cni_config.yaml"
  kubeapi_vlan: 1133
```

**Note:** Sample “acc-provision-input.yaml” and “bridge\_cni\_config.yaml” contents can be collected using following command:

```
acc-provision --sample -f openshift-vmm-lite-baremetal
```

### 3. Specific parameters for VMM-lite feature:

- `vmm_lite_config.aaep_monitoring_enabled` : Mandatory field, enables VMM Lite feature for OpenShift Virtualization.
- `vmm_lite_config.bridge_name` : Mandatory field, name of Linux bridge created on OpenShift worker nodes as mentioned in pre-requisites.
- `vmm_lite_config.cno_identifier` : Optional field, used in EPG annotations and NAD annotations to identify resources of CNO specific cluster. By default, value is “cno”.
- `vmm_lite_config.apic_username` : Mandatory field, name of APIC user used to generate Key and cert and used to communicate with APIC.
- `vmm_lite_config.bridge_nad_config_file` : Optional field, used to get custom bridge CNI configuration from user and used to create NAD.
- `vmm_lite_config.kubeapi_vlan` : Optional field, VLAN ID used to communicate with kubeapi-server.

### 4. Execute the acc-provision command with the appropriate flavor and input file. Run command on the host that has access to APIC.

```
acc-provision -a -c acc-provision-input.yaml -f openshift-vmm-lite-baremetal -u <apic
username> -p <apic password> -o aci_deployment_vmm_lite.yaml
```

This command will:

- Use the openshift-vmm-lite-baremetal flavor and acc-provision-input.yaml configuration file.
- Generate the VMM Lite Custom Resource Definition (CRD) and other CNO manifests into aci\_deployment\_vmm\_lite.yaml.

### 5. Apply the generated manifests to the OpenShift cluster:

```
oc apply -f aci_deployment_vmm_lite.yaml
```

This will:

- Bring up the CNO.

- Install the VMM Lite CRD (AaepMonitor).

## 6. After manifests are applied to the OpenShift cluster, you should see following resource:

```
noiro@ocp-3-ext-rtr:~$ oc get po -n aci-containers-system -o wide
```

| NAME                                      |                | READY     | STATUS  | RESTARTS | AGE | IP           |
|-------------------------------------------|----------------|-----------|---------|----------|-----|--------------|
| NODE                                      | NOMINATED NODE | READINESS | GATES   |          |     |              |
| aci-containers-controller-7f5d4bf68-lwvjv |                | 1/1       | Running | 0        | 38m | 192.168.23.6 |
| worker1.ocpbm3.noiro.local                | <none>         |           | <none>  |          |     |              |
| aci-containers-webhook-5475f7c4c-jm2n7    |                | 1/1       | Running | 0        | 38m | 10.2.7.116   |
| worker1.ocpbm3.noiro.local                | <none>         |           | <none>  |          |     |              |

### Installed CRD:

```
[noiro@ocpbm-fablocpbm2-ext-rtr ~] $ oc get crd aaepmonitors.aci.attachmentmonitor -o yaml
apiVersion: apiextensions.k8s.io/v1
kind: CustomResourceDefinition
metadata:
  name: aaepmonitors.aci.attachmentmonitor
spec:
  conversion:
    strategy: None
  group: aci.attachmentmonitor
  names:
    kind: AaepMonitor
    listKind: AaepMonitorList
    plural: aaepmonitors
    singular: aaepmonitor
  scope: Cluster
  versions:
  - name: v1
    schema:
      openAPIV3Schema:
        description: AaepMonitor is the Schema for AttachmentMonitors to monitor AAEPs
        properties:
          apiVersion:
            type: string
          kind:
            type: string
          metadata:
            type: object
          spec:
            properties:
              aaeps:
```

```

        description: List of AAEPs to monitor.
        items:
          type: string
        type: array
      required:
      - aaeps
      type: object
    status:
      properties:
        status:
          type: string
        type: object
      required:
      - spec
      type: object
    x-kubernetes-validations:
      - message: Only one instance with name aaepmonitor allowed
        rule: self.metadata.name == 'aaepmonitor'
  served: true
  storage: true
  subresources:
    status: {}

```

Ensure to configure AAEP monitoring, after VMM-lite installation.

## Configure AAEP monitoring

Apply the manifests, and then you need to create an AaepMonitor Custom Resource (CR) to specify which AAEPs CNO VMM-Lite should monitor.

- Only one AaepMonitor CR is allowed, and its name must be aaepmonitor.
- It should be created in the aci-containers-system namespace.

Example AaepMonitor CR:

```

apiVersion: aci.attachmentmonitor/v1
kind: AaepMonitor
metadata:
  name: aaepmonitor
  namespace: aci-containers-system
spec:
  aaeps: # List of AAEP names to monitor
    - AAEP_1
    - AAEP_2

```

**Note:** Currently the only use case for having multiple AAEPs in this CR is to support OpenShift cluster spread over multiple AAEPs.

---

Apply this CR to your cluster:

```
oc apply -f your-aaepmonitor-cr.yaml
```

After AAEPs are configured, you need NADs to launch the VMs.

## NAD creation logic

CNO VMM-Lite dynamically creates NADs based on monitored AAEPs and EPGs.

### 1. Conditions for NAD creation

A NAD will be created for an EPG if all the following conditions are met:

- The EPG is attached to any of the AAEPs specified in the AaepMonitor CR.
- The EPG is annotated with a namespace name that exists in the OpenShift setup.
  - Annotation Key: < cno-identifier >-namespace (e.g., cno-namespace if cno\_identifier is "cno" )
  - Annotation Value: < namespace-name >

Overlapping VLAN is not used when attaching the EPG with the AAEP. In this context, an overlapping VLAN is a VLAN that is being used while a previous EPG attachment and Network Attachment Definition (NAD) already exist for that VLAN.

### 2. NAD naming and annotations

- Default NAD Name: < tenant\_name >-< app\_profile\_name >-< epɡ\_name >-< hash >
- NAD Annotation for cno-name: If the EPG is annotated with a NAD name, the NAD will also be annotated with cno-name: < NAD-name-from-EPG-annotation >.
  - EPG Annotation Key: < cno-identifier >-nad
  - EPG Annotation Value: < Custom-NAD-name >

**Note:** This annotation only adds a NAD level annotation but will not change the name of the NAD itself. The name is always < tenant\_name >-< app\_profile\_name >-< epɡ\_name >-< hash >

- Managed by annotation: All NADs created by VMM-Lite will be annotated with managed-by: cisco-network-operator.
- Sync status annotation: NADs will initially be annotated with aci-sync-status: in-sync. Refer to the next section for more details.

### 3. Deferred creation: If an EPG is annotated with a namespace that does not yet exist, NAD creation will be deferred until the namespace is created.

### 4. Bridge CNI configuration for NADs

- Default values: If bridge\_nad\_config\_file is not provided in acc-provision-input.yaml, NADs will be created with default bridge CNI configuration values.
- Custom configuration: You can provide a custom YAML file via bridge\_nad\_config\_file to override default bridge CNI settings.

## NAD deletion logic

NADs are automatically deleted under several conditions:

- EPG deletion: If the associated EPG is deleted.
- EPG detachment: If the EPG is detached from the monitored AAEP.
- CR deletion: If the aaepmonitor CR is deleted.
- AAEP removal from CR: If an AAEP is removed from the aaepmonitor CR.
- Namespace annotation deletion: If the namespace annotation is removed from the EPG.
- VLAN overlap: If the EPG's VLAN changes, and the new VLAN overlaps with another EPG's VLAN (for which a NAD already exists).

**Note:** A NAD will only be deleted if no VMs/PODs are using it.

#### 1. **Handle pods associated with NADs**

- If a NAD is targeted for deletion but has one or more pods associated with it, the NAD will not be deleted.
- Instead, the NAD will be annotated with `aci-sync-status: out-of-sync`, and an event will be injected explaining the reason.
- If the EPG and AAEP association are later restored, the `aci-sync-status` annotation will revert to `in-sync`.

#### 2. **Manual NAD deletion**

- Manual deletion of a NAD is allowed only if it is not annotated with `managed-by: cisco-network-operator`.
- To manually clean up a NAD created by VMM-Lite, you must first remove the `managed-by: cisco-network-operator` annotation.

## NAD update logic

CNO VMM-Lite automatically updates NADs in response to changes in EPG configurations.

1. **EPG VLAN ID update:** If an EPG's VLAN ID changes, the corresponding NAD will be updated with the new VLAN ID, provided the NAD is managed by CNO.
  - If the new VLAN overlaps with another EPG's VLAN, the existing NAD will be deleted if not VMs/Pods using it, and a new NAD will not be created for that EPG.
2. **Namespace name change:** If an EPG's namespace annotation changes:
  - The NAD in the old namespace will be deleted (if no pods are using it). If pods are using it, the NAD will be marked `out-of-sync`.
  - A new NAD will be created in the new namespace.
3. **NAD name change:** If an EPG's NAD name annotation changes, the `cno-name` annotation on the NAD will be updated accordingly.

## Important considerations

1. **Single AaepMonitor CR:** Only one AaepMonitor CR named `aaepmonitor` is allowed per cluster.

- 
2. **VLAN updates with VMs:** VLANs for EPGs should not be updated if there are VMs currently using the associated NAD. Doing so will result in an inconsistent state where ACI is using the new VLAN and OpenShift is using the old one. This is a limitation of KubeVirt a NAD VLAN ID, which can't be updated if VMs are using it.
  3. **Cross-cluster coordination:** The current solution does not coordinate across clusters. If the same AAEP is monitored in different clusters, NADs will be created independently in each cluster.
  4. **First EPG rule:** When the first EPG is associated with an AAEP and a specific VLAN is used, a NAD is generated. Subsequent EPGs using the same AAEP and VLAN do not trigger additional NAD creation. If the initial EPG is detached/deleted, a new NAD will be created for one of the remaining EPGs sharing that VLAN.