

Connect

What's new in OpenShift virtualization (4.17)

Dieter De Moitie Specialist Solution Architect



What is OpenShift Virtualization?

- Included feature of the OpenShift application platform
- Unified platform for VMs and Containers
 - VM technically runs inside a container
 - Reuse compute/networking/storage for both
- Performance, stability, scalability, and reliability of KVM, the Linux kernel-based hypervisor (15+ years of production use)
- Support for RHEL and Windows
- Manageability and ecosystem of OpenShift





User Interface





Multicluster Virtual Machine Observability with RHACM

Targeted for ACM 2.12 (next month)



Comprehensive set of dashboards addressing important use cases

- Flexible Single VM/Single-Cluster view versus Multi
 VM/MultiCluster view
- Based on a rich set of metrics to retrieve valuable status of individual VM's and inventories



Comprehensive VM alerting

Red Hat Advanced Cluster Management for Kubernetes

Targeted for ACM 2.12 (next month)

SEE: Five key dashboards - observability **FIND**: Virtual machines easily with enhanced search capabilities (Dev Preview).

DO: Stop, start, restart, and pause VMs directly from ACM (Tech Preview).





Networking

Red Hat

Virtual Machine Networking

Recap





OVN Kubernetes secondary network

Topology layer 2 or localnet



User Defined Networks

Technology Preview!

Support for multiple (overlapping) networks/VRFs to enable user-defined native network isolation

Brings flexibility of secondary networks to the primary network (for pods)

Support for:

- (Admin) Network Policy
- primary (default) and secondary UDNs
- overlapping pod IPs across UDNs
- clusterIP services and external services

Future: support for BGP+EVPN

- extend UDN into provider networks
- No NAT needed to reach the VM





Storage



Migrate storage classes for running and stopped VMs.

Tech Preview

apiVersion: kubevirt.io/v1 kind: VirtualMachine kubevirt.io/vm: vm-dv name: vm-dv spec: + updateVolumesStrategy: Migration dataVolumeTemplates: - metadata: name: src-pvc name: dst-dv spec: storage: storageClassName: ocs-storagecluster-ceph-rbd storageClassName: ocs-storagecluster-ceph-rbd-virtualization volumes: - dataVolume: name: src-pvc name: dst-dv name: datavolumedisk1

12

Enabling this feature (cluster becomes unsupported): https://access.redhat.com/solutions/7089972

Follow progress:

oc get virtualmachineinstancemigrations -l
kubevirt.io/volume-update-migration=<vm-name>

Upstream documentation:

https://kubevirt.io/user-guide/storage/volume_migration/



Multi-DC deployment options





Disaster Recovery for VMs with ACM/ODF





Technology Preview

Regional-DR with ODF and ACM



Generally Available

Metro-DR with ODF and ACM

Note: Synchronous (metro) DR and Asynchronous DR with Portworx is Generally Available

4 or 5 node control plane



4 or 5 Node HA Control-Plane (Bare Metal Only)



- Active-active deployments across two locations
- Designed for traditional applications like
 OpenShift Virtualization VMs
- Enhances resiliency with 2+2 or 3+2 configurations
- Supported on bare metal platform only



Trade-Offs of Control Plane (etcd) Sizes





4-node Control-Plane Quorum Loss Scenarios





Requirements for quorum recovery



User with cluster-admin role through a Certificate-Based Kubeconfig file



A healthy control-plane node to use as

recovery host



An existing backup* directory containing the etcd snapshot

and the **resources for the static pods** from the same backup

snapshot_<datetimestamp>.db

static_kuberesources_<datetimestamp>.tar.gz



*This requirement goes away in OpenShift 4.18 and a copy will always be available

4 or 5 node Control Plane (bare metal only)

Recovery procedure

- Assure fencing of the lost data center.
- SSH into one of the remaining nodes
- Restore database from etcd snapshot (requirements goes away in 4.18)
- Add more nodes to the etcd cluster

20

https://docs.openshift.com/container-platform/4.17/backup_and_restore/control_plane_backup_an d_restore/disaster_recovery/scenario-2-restoring-cluster-state.html



Automated Recovery with 4-5 Node Control Plane

Targeted for 4.18

- Reduced Complexity Eliminates the need to follow a manual recovery process
- Improved Efficiency Saves time and minimizes human error during disaster recovery situations
- Enhanced Support Makes it easier for the etcd team to support and test disaster recovery workflows

How it works

- 1. Admin initiates automation on a designated recovery host (e.g., via CLI)
- 2. Required inputs gathered (backup files, static pod configs, SSH access details)
- 3. Recovery steps executed
- 4. Control plane is restored and fully functional with standing nodes
- 5. Recovered nodes are added to the cluster



Overcommitment



Resources in Kubernetes

Recap

In Kubernetes

23

- CPU = compressible resource
 - CPU gets throttled
 - · CPU cycles divided based Quality of Service classes
- Memory = non-compressible resource
 - Memory reclaimed by killing containers (soft or hard eviction)
 - Eviction based on Quality of Service Classes
 - eviction might not be an issue for cloud-native workloads
 - Hard: no grace period (default)
 - · Soft: eviction-max-pod-grace-period

In OpenShift virtualization, by default:

- 1/10th of CPU is reserved (CPU overcommitting by default)
- All memory of a VM is reserved
- Live migration for soft eviction



Memory Overcommitting options

- ► Free Page Reporting
 - Enabled by default
 - FPR allows guests to report (release) unused memory back to the hypervisor
 - https://developers.redhat.com/articles/2024/03/13/save-memory-openshift-virtualization-usin g-free-page-reporting?source=sso#
- Enable Kernel Same Page Merging (KSM)
 - allows KVM guests to share identical memory pages
 - only use with trusted workloads.
- SWAP using wasp-agent
 - Out of tree agent: <u>https://github.com/openshift-virtualization/wasp-agent</u>
 - facilitates memory overcommitment by assigning swap resources
 - manages pod evictions when nodes are at risk due to high swap I/O traffic or high utilization.
 - Only for Burstable pods
 - Can decrease performance



VM distribution

Soft eviction

Migrates VM when nodes are at risk

Descheduler

- Automatic workload balancing using Descheduler
- LowNodeUtilization: evicts pods from overutilized nodes when there are any underutilized nodes. The destination node for the evicted pod will be determined by the scheduler.
- A node is considered underutilized if its usage is below 20% for all thresholds (CPU, memory, and number of pods).
- A node is considered overutilized if its usage is above 50% for any of the thresholds (CPU, memory, and number of pods).



Lightspeed

Red Hat

OpenShift Lightspeed

(Technology Preview) Generative AI based chat assistant

0

0

- 0

1

Help where you need it

Integrated directly into the Red Hat OpenShift web console

Explain, investigate and learn more

Provides assistance with explaining and investigating cluster resources

Generative Al

Powerful, pluggable LLMs combined with the latest OpenShift documentation

- RHEL AI
- Red Hat OpenShift Al
- OpenAl
- Azure Al
- Watson X

Red Hat OpenShift Lightspeed

Explore deeper insights, engage in meaningful discussions, and unlock new possibilities with Red Hat OpenShift Lightspeed. Answers are provided by generative AI technology, please use appropriate caution when following recommendations.

1 Important

OpenShift Lightspeed can answer questions related to OpenShift. Do not include personal or business sensitive information in your input. Interactions with OpenShift Lightspeed may be reviewed and used to improve our products and services.





AL VMCannotBeEvicted A Warning

Eviction policy for VirtualMachine centos-stream9-gray-otter-66 in namespace NS openshift-cnv (on node) is set to Live Migration but the VM

Alert details



 Name
 Source

 VMCannotBeEvicted
 Platform

 Severity
 State

 ▲ Warning
 ▲ Firing

 Description
 Mathematical State

Eviction policy for VirtualMachine centos-stream9-gray-otter-66 in namespace **NS** openshift-cnv (on node) is set to Live Migration but the VM is not migratable

Summary

The VM's eviction strategy is set to Live Migration but the VM is not migratable

 $\xrightarrow{|X|}{|x^{*}}$

Red Hat OpenShift Lightspeed

Explore deeper insights, engage in meaningful discussions, and unlock new possibilities with Red Hat OpenShift Lightspeed. Answers are provided by generative AI technology, please use appropriate caution when following recommendations.

Important

OpenShift Lightspeed can answer questions related to OpenShift. Do not include personal or business sensitive information in your input. Interactions with OpenShift Lightspeed may be reviewed and used to improve our products and services.

explain the alert and how to fix it

<u>o</u> C

욹

AL VMCannotBeEvicted 🗙

Always check AI/LLM generated responses for accuracy prior to use.



Red Hat OpenShift Lightspeed



Val	
rou	

explain the alert and how to fix it





OpenShift Lightspeed

The alert VMCannotBeEvicted is triggered when the eviction strategy of a virtual machine (VM) is set to LiveMigration, but the VM is not migratable. This condition can prevent node eviction, affecting operations such as node drain and updates.

Impact

Non-migratable VMs prevent node eviction, which can disrupt maintenance activities like draining nodes for updates or other administrative tasks.

Diagnosis

- To diagnose this issue, follow these steps:
- Check the VMI Configuration: Verify if the evictionStrategy of the Virtual Machine Instance (VMI) is set to LiveMigrate.

\$ oc get vmis -o yaml

 Identify Non-Migratable VMIs: Look for a False status in the LIVE-MIGRATABLE column to identify VMIs that are not migratable.

\$ oc get vmis -o wide



0



.



Project: openshift-cnv 🔹

VirtualMachines > VirtualMachine details

VM centos-stream9-gray-otter-66 GErrorUnschedulable Configuration Events Snapshots Diagnostics 1 Overview Metrics YAML Console Status conditions ③ Status & Conditions Red Hat OpenShift Lightspeed ▼ Filter ▼ Search by reason... 1 Explore deeper insights, engage in meaningful discussions, and unlock new Guest system log possibilities with Red Hat OpenShift Lightspeed. Answers are provided by generative AI technology, please use appropriate caution when following Sta... Reason > Type Message recommendations. Ready False GuestNotRunning Guest VM is not reported as running Important OpenShift Lightspeed can answer questions related to OpenShift. Do not include personal or business sensitive information in your input. Interactions with OpenShift Lightspeed may be reviewed and used to improve our products and services. Guest VM is not reported as running **NotAllDVsReady** Not all of the VMI's DVs are ready ~ DataVolumesReady False Attach VM logs/ O Not all of the VMI's DVs are ready events/yaml/alerts PodScheduled False Unschedulable 0/2 nodes are available: 2 Insufficient deviv 0 What is wrong with my VM Always check AI/LLM generated responses for accuracy prior to use. 0/2 nodes are available: 2 Insufficient devices.kubevirt.io/kvm. preemption: 0/2 nodes are available: 2 No

00

1 of 1

<< <

1

Red Hat

53 _





Red Hat OpenShift Lightspeed

You

what are errors and how solve the errors



.....

OpenShift Lightspeed

The provided YAML configuration for the Virtual Machine resource from KubeVirt indicates several errors in the status section. Here are the identified errors and potential solutions:

Errors

1. Guest VM is not reported as running:

- Type: Ready
- Status: False
- Reason: GuestNotRunning
- Message: Guest VM is not reported as running

2. Not all of the VMI's DVs are ready:

- Type: DataVolumesReady
- Status: False
- Reason: NotAllDVsReady
- Message: Not all of the VMI's DVs are ready
- 3. Pod scheduling issues due to node affinity/selector and preemption constraints:
- 0/2 nodes are available.
- 2 nodes didn't match Pod's node affinity/selector.
- Preemption is not helpful for scheduling.

Solutions

1. Guest VM is not reported as running:

This error suggests that the virtual machine instance (VMI) has started but isn't fully operational.

Solution:

1. Check logs for more detailed error messages using kubectl logs <pod-name>.

2. Ensure that all required resources (CPU, memory) are available on your cluster.

3. Verify that any dependencies or services required by this VM are up and running.

2. Not all of the VMI's DVs are ready:

This indicates an issue with one or more data volumes associated with this virtual machine instance. Solution:

Send a message...



#

-

Clear chat



Red Hat OpenShift Lightspeed



You

How can i delete my virtual machine



OpenShift Lightspeed

To delete a virtual machine in OpenShift, you have two options: using the web console or the command line interface (CLI).

Deleting a Virtual Machine Using the Web Console

- In the Red Hat OpenShift Container Platform console, click Virtualization -> VirtualMachines from the side menu.
- Click the Options menu (:) beside a virtual machine and select **Delete**.
 Alternatively, click on the virtual machine name to open its details page and then click **Actions** → **Delete**.
- Optional: Select With grace period or clear Delete disks, depending on your requirements.
- 4. Click Delete to permanently remove the virtual machine.

Deleting a Virtual Machine Using the CLI

1. Identify the name of the virtual machine that you want to delete.

2. Run the following command:

\$ oc delete vm <vm_name>

Note:

О

 This command only deletes a VM in your current project. If your VM is in another project or namespace, specify it with -n <project_name>.
 Would you like more detailed steps for either method?

Send a message...

Delete d de come entreten





Reference Materials



Reference material

OpenShift Virtualization

Reference Implementation Guide

Hybrid Platforms Business Unit

version: 1.0.2

September 2024



Ecosystem Catalog Sol	utions Certified products Artifacts F	Partners All - Q. Search Eco	System Catalog - Contact us Resou
dilla Francisco Babilio	Overview Workloads	Infrastructure Services Conta	ct
Coffware Workloade			
Software workloads			
Storage & networking Ba	ckup & disaster recovery		
hese certified products integrate Container Storage Interface (CSI)	with OpenShift Virtualization using na or Container Networking Interface (CN	tive Kubernetes interfaces: II).	1 of 3
hese certified products integrate container Storage Interface (CSI)	e with OpenShift Virtualization using na or Container Networking Interface (Ch	tive Kubernetes interfaces: al).	
hese certified products integrate container Storage Interface (CSI)	e with OpenShift Virtualization using na or Container Networking Interface (CN TONETAPP Astra Trident	tive Kubernetes interfaces: II). IBM Fusion	Infinidat InfiniBox®
hese certified products integrate container Storage Interface (CSI)	e with OpenShift Virtualization using nation Container Networking Interface (CN Interface (CN Astra Trident by NetApp Data connectivity to persistent	IBM Fusion by IBM Japan, Ltd.	Infinidat InfiniBox® Container Storage Interface (CSI) Driver
hese certified products integrate container Storage Interface (CSI) Dell Container Storage Modules by DELL Easily install and manage Dell's	e with OpenShift Virtualization using nation or Container Networking Interface (CN Interface (CN Astra Trident by NetApp Data connectivity to persistent data stores for Kubernetes	IBM Fusion by IBM Japan, Ltd. The easiest way to deploy OpenShift applications and	Infinidat InfiniBox® Container Storage Interface (CSI) Driver by Infinidat Ltd.
hese certified products integrate container Storage Interface (CSI) Dell Container Storage Modules by DELL Easily install and manage Dell's CSI Drivers and CSM	e with OpenShift Virtualization using na or Container Networking Interface (CN INTINETAPP Astra Trident by NetApp Data connectivity to persistent data stores for Kubernetes applications	tive Kubernetes interfaces: I). IBM Fusion by IBM Japan, Ltd. The easiest way to deploy OpenShift applications and watsonx.	Infinidat InfiniBox* Container Storage Interface (CSI) Driver by Infinidat Ltd. CNCF-compliant Kubernetes integration for InfiniBox storage systems, offering

Hardware infrastructure

Select from a broad range of trusted systems for your OpenShift clusters, and take advantage of Red Hat's close collaboration with all major hardware providers.



red.ht/workswithvirt



1 of 2



Reference material

Red Hat				ads Red Hat Console Get Support	
ome > Products > Workload Availa	ability for Red Hat OpenShift > 24	4.3 ➤ Remediation, fencing, and maintenance		er Portal Products Knowledge Security Support	Search
			Products &	Services > Articles > Red Hat OpenShift Virtualization disaster recovery guid	le
Focus mode		Remediation, fencing, and maintenance		Red Hat OpenShift Virtualization disaster recovery guide	
diation, fencing, and enance		± PDF	14	Updated November 13 2023 at 4:08 PM - English -	
ace		WORKI OAD AVAIL ABILITY FOR RED HAT OPENSHIET 24.3	Ģ	1. Disaster recovery solutions	5.1 Overview
iding feedback on Workload lability for Red Hat OpenShift umentation		Workload Availability remediation, fencing, and maintenance		1.1 Application failover 2. Metro-DR	5.2 Elasticsearch cluster virtual machines 5.3 Kibana container
but node remediation, fencing,	>	Red Hat Customer Content Services		2.1 Overview 2.2 Network considerations	5.4 GitOps considerations 5.5 Network considerations
aintenance ng Self Node Remediation	>	Abstract		2.3 Recovery objectives	5.6. Storage considerations
ng Fence Agents Remediation	>	Information about workload availability operators and their usage		3.1 Overview	6. Using OpenShift GitOps with OpenShift Virtualization scenarios
ng Machine Deletion diation	>			3.2 Network considerations 3.3 Recovery Objectives	6.1 GitOps overview 6.2 About OpenShift Virtualization resources
rediating nodes with Machine Checks	>			4. Disaster recovery procedure	6.3 How can I protect volumes from being deleted in a DI6.4 OpenShift GitOps behavior in different versions of O
nediating nodes with Node Checks	>			4.2. Failover to passive cluster	Virtualization 6.4.1 OpenShift 4.13 and 4.12
ing nodes in maintenance	>			4.3 Failback to primary cluster 4.4 Red Hat Advanced Cluster Management configuration	6.4.2 OpenShift 4.14 Conclusion
de with Node Maintenance erator				4.5 Failover with Kustomize 5. Example composite application with VMs and containers	Glossary
	Workload	Availability Guide			Appendix A. Volsync

https://docs.redhat.com/en/documentation/workl oad availability for red hat openshift/24.3 Fencing and VM High Availability Guide https://access.redhat.com/articles/7057929

Disaster recovery guide:

https://access.redhat.com/articles/7041594





Connect

Thank you



linkedin.com/company/red-hat



youtube.com/user/RedHatVideos



facebook.com/redhatinc



twitter.com/RedHat

