



SUSE Enterprise Storage 7

SUSE Enterprise Storage for Windows guide

SUSE Enterprise Storage for Windows guide


SUSE Enterprise Storage 7

by Mike Latimer and Alexandra Settle

Publication Date: 12 Dec 2024

<https://documentation.suse.com> 

Copyright © 2020–2024 SUSE LLC and contributors. All rights reserved.

Except where otherwise noted, this document is licensed under Creative Commons Attribution-ShareAlike 4.0 International (CC-BY-SA 4.0): <https://creativecommons.org/licenses/by-sa/4.0/legalcode> .

For SUSE trademarks, see <http://www.suse.com/company/legal/> . All third-party trademarks are the property of their respective owners. Trademark symbols (®, ™ etc.) denote trademarks of SUSE and its affiliates. Asterisks (*) denote third-party trademarks.

All information found in this book has been compiled with utmost attention to detail. However, this does not guarantee complete accuracy. Neither SUSE LLC, its affiliates, the authors nor the translators shall be held liable for possible errors or the consequences thereof.

Contents

About this guide vi

- 1 Available documentation vi
- 2 Giving feedback vii
- 3 Documentation conventions viii
- 4 Support x
 - Support statement for SUSE Enterprise Storage x • Technology previews xi
- 5 Ceph contributors xi
- 6 Commands and command prompts used in this guide xii
 - Salt-related commands xii • Ceph related commands xii • General Linux commands xiii • Additional information xiv
- 1 Ceph for Microsoft Windows 1**
 - 1.1 Introduction 1
 - 1.2 Technology preview 1
 - 1.3 Supported platforms 2
 - 1.4 Compatibility 2
 - 1.5 Installing and configuring 2
 - 1.6 RADOS Block Device (RBD) 3
 - Mapping images 3 • Hyper-V VM disks 4 • Configuring Microsoft Windows partitions 4
 - 1.7 RBD Microsoft Windows service 5
 - 1.8 Configuring CephFS 5

- A Sample configuration files 7
 - B Troubleshooting tips 8
 - C Upstream projects 9
 - D Ceph maintenance updates based on upstream
'Octopus' point releases 10
- Glossary 15

About this guide

This guide describes the integration, installation and configuration of Microsoft Windows environments and SUSE Enterprise Storage using the Windows Driver.

SUSE Enterprise Storage 7 is an extension to SUSE Linux Enterprise Server 15 SP2. It combines the capabilities of the Ceph (<http://ceph.com/>) storage project with the enterprise engineering and support of SUSE. SUSE Enterprise Storage 7 provides IT organizations with the ability to deploy a distributed storage architecture that can support a number of use cases using commodity hardware platforms.

1 Available documentation



Note: Online documentation and latest updates

Documentation for our products is available at <https://documentation.suse.com>, where you can also find the latest updates, and browse or download the documentation in various formats. The latest documentation updates can be found in the English language version.

In addition, the product documentation is available in your installed system under `/usr/share/doc/manual`. It is included in an RPM package named `ses-manual_LANG_CODE`. Install it if it is not already on your system, for example:

```
# zypper install ses-manual_en
```

The following documentation is available for this product:

Deployment Guide (<https://documentation.suse.com/ses/html/ses-all/book-storage-deployment.html>)

This guide focuses on deploying a basic Ceph cluster, and how to deploy additional services. It also cover the steps for upgrading to SUSE Enterprise Storage 7 from the previous product version.

Administration and Operations Guide (<https://documentation.suse.com/ses/html/ses-all/book-storage-admin.html>) ↗

This guide focuses on routine tasks that you as an administrator need to take care of after the basic Ceph cluster has been deployed (day 2 operations). It also describes all the supported ways to access data stored in a Ceph cluster.

Security Hardening Guide (<https://documentation.suse.com/ses/html/ses-all/book-storage-security.html>) ↗

This guide focuses on how to ensure your cluster is secure.

Troubleshooting Guide (<https://documentation.suse.com/ses/html/ses-all/book-storage-troubleshooting.html>) ↗

This guide takes you through various common problems when running SUSE Enterprise Storage 7 and other related issues to relevant components such as Ceph or Object Gateway.

SUSE Enterprise Storage for Windows Guide (<https://documentation.suse.com/ses/html/ses-all/book-storage-windows.html>) ↗

This guide describes the integration, installation, and configuration of Microsoft Windows environments and SUSE Enterprise Storage using the Windows Driver.

2 Giving feedback

We welcome feedback on, and contributions to, this documentation. There are several channels for this:

Service requests and support

For services and support options available for your product, see <http://www.suse.com/support/> ↗.

To open a service request, you need a SUSE subscription registered at SUSE Customer Center. Go to <https://scc.suse.com/support/requests> ↗, log in, and click *Create New*.

Bug reports

Report issues with the documentation at <https://bugzilla.suse.com/> ↗. Reporting issues requires a Bugzilla account.

To simplify this process, you can use the *Report Documentation Bug* links next to headlines in the HTML version of this document. These preselect the right product and category in Bugzilla and add a link to the current section. You can start typing your bug report right away.

Contributions

To contribute to this documentation, use the *Edit Source* links next to headlines in the HTML version of this document. They take you to the source code on GitHub, where you can open a pull request. Contributing requires a GitHub account.



For more information about the documentation environment used for this documentation, see the repository's README at <https://github.com/SUSE/doc-ses> .

Mail

You can also report errors and send feedback concerning the documentation to doc-team@suse.com. Include the document title, the product version, and the publication date of the document. Additionally, include the relevant section number and title (or provide the URL) and provide a concise description of the problem.

3 Documentation conventions

The following notices and typographic conventions are used in this document:

- /etc/passwd: Directory names and file names
- PLACEHOLDER: Replace PLACEHOLDER with the actual value
- PATH: An environment variable
- ls, --help: Commands, options, and parameters
- user: The name of user or group
- package_name: The name of a software package
- **Alt**, **Alt - F1**: A key to press or a key combination. Keys are shown in uppercase as on a keyboard.
- *File*, *File > Save As*: menu items, buttons
- **AMD/Intel** This paragraph is only relevant for the AMD64/Intel 64 architectures. The arrows mark the beginning and the end of the text block. 
- **IBM Z, POWER** This paragraph is only relevant for the architectures IBM Z and POWER. The arrows mark the beginning and the end of the text block. 
- *Chapter 1, “Example chapter”*: A cross-reference to another chapter in this guide.

- Commands that must be run with root privileges. Often you can also prefix these commands with the sudo command to run them as non-privileged user.

```
# command  
> sudo command
```

- Commands that can be run by non-privileged users.

```
> command
```

- Notices



Warning: Warning notice

Vital information you must be aware of before proceeding. Warns you about security issues, potential loss of data, damage to hardware, or physical hazards.



Important: Important notice

Important information you should be aware of before proceeding.



Note: Note notice

Additional information, for example about differences in software versions.



Tip: Tip notice

Helpful information, like a guideline or a piece of practical advice.

- Compact Notices



Additional information, for example about differences in software versions.



Helpful information, like a guideline or a piece of practical advice.

4 Support

Find the support statement for SUSE Enterprise Storage and general information about technology previews below. For details about the product lifecycle, see <https://www.suse.com/lifecycle>. If you are entitled to support, find details on how to collect information for a support ticket at <https://documentation.suse.com/sles-15/html/SLES-all/cha-adm-support.html>.

4.1 Support statement for SUSE Enterprise Storage

To receive support, you need an appropriate subscription with SUSE. To view the specific support offerings available to you, go to <https://www.suse.com/support/> and select your product.

The support levels are defined as follows:

L1

Problem determination, which means technical support designed to provide compatibility information, usage support, ongoing maintenance, information gathering and basic troubleshooting using available documentation.

L2

Problem isolation, which means technical support designed to analyze data, reproduce customer problems, isolate problem area and provide a resolution for problems not resolved by Level 1 or prepare for Level 3.

L3

Problem resolution, which means technical support designed to resolve problems by engaging engineering to resolve product defects which have been identified by Level 2 Support.

For contracted customers and partners, SUSE Enterprise Storage is delivered with L3 support for all packages, except for the following:

- Technology previews.
- Sound, graphics, fonts, and artwork.
- Packages that require an additional customer contract.
- Some packages shipped as part of the module *Workstation Extension* are L2-supported only.
- Packages with names ending in `-devel` (containing header files and similar developer resources) will only be supported together with their main packages.

SUSE will only support the usage of original packages. That is, packages that are unchanged and not recompiled.

4.2 Technology previews

Technology previews are packages, stacks, or features delivered by SUSE to provide glimpses into upcoming innovations. Technology previews are included for your convenience to give you a chance to test new technologies within your environment. We would appreciate your feedback! If you test a technology preview, please contact your SUSE representative and let them know about your experience and use cases. Your input is helpful for future development.

Technology previews have the following limitations:

- Technology previews are still in development. Therefore, they may be functionally incomplete, unstable, or in other ways *not* suitable for production use.
- Technology previews are *not* supported.
- Technology previews may only be available for specific hardware architectures.
- Details and functionality of technology previews are subject to change. As a result, upgrading to subsequent releases of a technology preview may be impossible and require a fresh installation.
- SUSE may discover that a preview does not meet customer or market needs, or does not comply with enterprise standards. Technology previews can be removed from a product at any time. SUSE does not commit to providing a supported version of such technologies in the future.

For an overview of technology previews shipped with your product, see the release notes at https://www.suse.com/releasenotes/x86_64/SUSE-Enterprise-Storage/7.

5 Ceph contributors

The Ceph project and its documentation is a result of the work of hundreds of contributors and organizations. See <https://ceph.com/contributors/> for more details.

6 Commands and command prompts used in this guide

As a Ceph cluster administrator, you will be configuring and adjusting the cluster behavior by running specific commands. There are several types of commands you will need:

6.1 Salt-related commands

These commands help you to deploy Ceph cluster nodes, run commands on several (or all) cluster nodes at the same time, or assist you when adding or removing cluster nodes. The most frequently used commands are `ceph-salt` and `ceph-salt config`. You need to run Salt commands on the Salt Master node as `root`. These commands are introduced with the following prompt:

```
root@master #
```

For example:

```
root@master # ceph-salt config ls
```

6.2 Ceph related commands

These are lower-level commands to configure and fine tune all aspects of the cluster and its gateways on the command line, for example `ceph`, `cephadm`, `rbd`, or `radosgw-admin`.

To run Ceph related commands, you need to have read access to a Ceph key. The key's capabilities then define your privileges within the Ceph environment. One option is to run Ceph commands as `root` (or via `sudo`) and use the unrestricted default keyring 'ceph.client.admin.key'. The safer and recommended option is to create a more restrictive individual key for each administrator user and put it in a directory where the users can read it, for example:

```
~/ceph/ceph.client.USERNAME.keyring
```



Tip: Path to Ceph keys

To use a custom admin user and keyring, you need to specify the user name and path to the key each time you run the `ceph` command using the `-n client.USER_NAME` and `--keyring PATH/TO/KEYRING` options.

To avoid this, include these options in the `CEPH_ARGS` variable in the individual users' `~/.bashrc` files.

Although you can run Ceph-related commands on any cluster node, we recommend running them on the Admin Node. This documentation uses the `cephuser` user to run the commands, therefore they are introduced with the following prompt:

```
cephuser@adm >
```

For example:

```
cephuser@adm > ceph auth list
```



Tip: Commands for specific nodes

If the documentation instructs you to run a command on a cluster node with a specific role, it will be addressed by the prompt. For example:

```
cephuser@mon >
```

6.2.1 Running `ceph-volume`

Starting with SUSE Enterprise Storage 7, Ceph services are running containerized. If you need to run `ceph-volume` on an OSD node, you need to prepend it with the `cephadm` command, for example:

```
cephuser@adm > cephadm ceph-volume simple scan
```

6.3 General Linux commands

Linux commands not related to Ceph, such as `mount`, `cat`, or `openssl`, are introduced either with the `cephuser@adm >` or `#` prompts, depending on which privileges the related command requires.

6.4 Additional information

For more information on Ceph key management, refer to *Book "Administration and Operations Guide", Chapter 30 "Authentication with cephx", Section 30.2 "Key management"*.

1 Ceph for Microsoft Windows

1.1 Introduction

Ceph is a highly-resilient software-defined-storage offering, which has only been available to Microsoft Windows environments through the use of iSCSI or CIFS gateways. This gateway architecture introduces a single point of contact and limits fault-tolerance and bandwidth, in comparison to the native I/O paths of Ceph with RADOS.

In order to bring the benefits of native Ceph to Microsoft Windows environments, SUSE partnered with Cloudbase Solutions to port Ceph to the Microsoft Windows platform. This work is nearing completion, and provides the following functionality:


- RADOS Block Device (RBD)
- CephFS

You can find additional information on the background of this effort through the following SUSECON Digital session:

- Ceph in a Windows World (TUT-1121 (<https://www.youtube.com/watch?v=BWZlwXLcNts>) )
Presenters: Mike Latimer (SUSE) Alessandro Pilotti (Cloudbase Solutions)

1.2 Technology preview

SUSE Enterprise Storage Driver for Windows is currently being offered as a technology preview. This is a necessary step toward full support as we continue work to ensure this driver performs well in all environments and workloads. You can contribute to this effort by reporting any issues you may encounter to SUSE Support.

CephFS functionality requires a third party FUSE wrapper provided through the [Dokany project](https://github.com/dokan-dev/dokany) (<https://github.com/dokan-dev/dokany>) . This functionality should be considered experimental, and is not recommended for production use.

1.3 Supported platforms

Microsoft Windows Server 2016 and 2019 are supported. Previous Microsoft Windows Server versions, including Microsoft Windows client versions such as Microsoft Windows 10, may work, but for the purpose of this document have not been thoroughly tested.



Note

Early builds of Microsoft Windows Server 2016 do not provide UNIX sockets, in which case the Ceph admin socket feature is unavailable.

1.4 Compatibility

RADOS Block Device images can be exposed to the OS and host Microsoft Windows partitions or they can be attached to Hyper-V VMs in the same way as iSCSI disks.



Note

At the moment, the Microsoft Failover Cluster refuses to use Windows Block Device (WNBD) driver disks as Cluster Shared Volumes (CSVs) underlying storage.

OpenStack integration has been proposed and may be included in the next OpenStack release. This will allow RBD images managed by OpenStack Cinder to be attached to Hyper-V VMs managed by OpenStack Nova.

1.5 Installing and configuring

Ceph for Microsoft Windows can be easily installed through the SES4Win.msi setup wizard. You can download this from [SES4Win](https://beta.suse.com/private/SLE15/SP2/download/SES7/SES4Win/) (<https://beta.suse.com/private/SLE15/SP2/download/SES7/SES4Win/>). This wizard performs the following functions:

- Installs Ceph-related code to the C:\Program Files\Ceph directory.
- Adds C:\Program Files\Ceph\bin to the %PATH% environment variable.
- Creates a Ceph RBD Mapping Service to automatically map RBD devices upon machine restart (using rbd-wnbd.exe).

After installing Ceph for Microsoft Windows, manual modifications are required to provide access to a Ceph cluster. The files which must be created or modified are as follows:

```
C:\ProgramData\ceph\ceph.conf  
C:\ProgramData\ceph\keyring
```

These files can be copied directly from an existing OSD node in the cluster. Sample configuration files are provided in [Appendix A, Sample configuration files](#).

1.6 RADOS Block Device (RBD)

Support for RBD devices is possible through a combination of Ceph tools and Microsoft Windows WNBBD. This driver is in the process of being certified by the Windows Hardware Quality Labs (WHQL).

Once installed, the WNBBD SCSI Virtual Adapter driver can be seen in the Device Manager as a storage controller. Multiple adapters may be seen, in order to handle multiple RBD connections. The **rbd** command is used to create, remove, import, export, map, or unmap images, exactly like it is used on Linux.

1.6.1 Mapping images

The behavior of the **rbd** command is similar to its Linux counterpart, with a few notable differences:

- Device paths cannot be requested. The disk number and path is picked by Microsoft Windows. If a device path is provided by the user when mapping an image, it is used as an identifier. This can also be used when unmapping the image.
- The **show** command was added, which describes a specific mapping. This can be used for retrieving the disk path.
- The **service** command was added, allowing **rbd-wnbd** to run as a Microsoft Windows service. All mappings are currently persistent and will be recreated when the service stops, unless they are explicitly unmapped. The service disconnects the mappings when being stopped.
- The **list** command also includes a **status** column.

The mapped images can either be consumed by the host directly or exposed to Hyper-V VMs.

1.6.2 Hyper-V VM disks

The following sample imports an RBD image and boots a Hyper-V VM using it.

```
# Feel free to use any other image. This one is convenient to use for
# testing purposes because it's very small (~15MB) and the login prompt
# prints the pre-configured password.
wget http://download.cirros-cloud.net/0.5.1/cirros-0.5.1-x86_64-disk.img `
    -O cirros-0.5.1-x86_64-disk.img

# We'll need to make sure that the imported images are raw (so no qcow2 or vhdx).
# You may get qemu-img from https://cloudbase.it/qemu-img-windows/
# You can add the extracted location to $env:Path or update the path accordingly.
qemu-img convert -O raw cirros-0.5.1-x86_64-disk.img cirros-0.5.1-x86_64-disk.raw

rbd import cirros-0.5.1-x86_64-disk.raw
# Let's give it a hefty 100MB size.
rbd resize cirros-0.5.1-x86_64-disk.raw --size=100MB

rbd-wnbd map cirros-0.5.1-x86_64-disk.raw

# Let's have a look at the mappings.
rbd-wnbd list
Get-Disk

$mappingJson = rbd-wnbd show cirros-0.5.1-x86_64-disk.raw --format=json
$mappingJson = $mappingJson | ConvertFrom-Json

$diskNumber = $mappingJson.disk_number

New-VM -VMName BootFromRBD -MemoryStartupBytes 512MB
# The disk must be turned offline before it can be passed to Hyper-V VMs
Set-Disk -Number $diskNumber -IsOffline $true
Add-VMHardDiskDrive -VMName BootFromRBD -DiskNumber $diskNumber
Start-VM -VMName BootFromRBD
```

1.6.3 Configuring Microsoft Windows partitions

The following sample creates an empty RBD image, attaches it to the host and initializes a partition:

```
rbd create blank_image --size=1G
rbd-wnbd map blank_image

$mappingJson = rbd-wnbd show blank_image --format=json
```

```
$mappingJson = $mappingJson | ConvertFrom-Json

$diskNumber = $mappingJson.disk_number

# The disk must be online before creating or accessing partitions.
Set-Disk -Number $diskNumber -IsOffline $false

# Initialize the disk, partition it and create a filesystem.
Get-Disk -Number $diskNumber | `
    Initialize-Disk -PassThru | `
    New-Partition -AssignDriveLetter -UseMaximumSize | `
    Format-Volume -Force -Confirm:$false
```

1.7 RBD Microsoft Windows service

In order to ensure that **rbd-wnbd** mappings survive host reboots, a new Microsoft Windows service, called the Ceph RBD Mapping Service has been created. This service automatically maintains mappings as they are added using the Ceph tools. All mappings are currently persistent and are recreated when the service starts, unless they are explicitly unmapped. The service disconnects all mappings when stopped.

This service also adjusts the Microsoft Windows service start order so that RBD images can be mapped before starting any services that may depend on them. For example, VMs.

RBD maps are stored in the Microsoft Windows registry at the following location:

```
SYSTEM\CurrentControlSet\Services\rbd-wnbd
```

1.8 Configuring CephFS



Note

The following feature is *experimental*, and is not intended for use in production environments.

Ceph for Microsoft Windows provides CephFS support through the Dokany FUSE wrapper. In order to use CephFS, install Dokany v1.4.1 or newer using the installers available here: <https://github.com/dokan-dev/dokany/releases>

With Dokany installed, and `ceph.conf` and `ceph.client.admin.keyring` configuration files in place, CephFS can be mounted using the `ceph-dokan.exe` command. For example:

```
ceph-dokan.exe -l x
```

This command mounts the default Ceph file system using the drive letter `X`. If `ceph.conf` is not placed at the default location (`C:\ProgramData\ceph\ceph.conf`), a `-c` parameter can be used to specify the location of `ceph.conf`.

The `-l` argument also allows using an empty folder as a mountpoint instead of a drive letter.

The UID and GID used for mounting the file system defaults to `0` and may be changed using the following `ceph.conf` options:

```
[client]
# client_permissions = true
client_mount_uid = 1000
client_mount_gid = 1000
```



Important


Microsoft Windows Access Control Lists (ACLs) are ignored. Portable Operating System Interface (POSIX) ACLs are supported but cannot be modified using the current CLI.



Important

CephFS does not support mandatory file locks, which Microsoft Windows heavily relies upon. At the moment, we are letting Dokan handle file locks, which are only enforced locally.

For debugging purposes, `-d` and `-s` may be used. The former enables debug output and the latter enables `stderr` logging. By default, debug messages are sent to a connected debugger.

You may use `--help` to get the full list of available options. Additional information on this *experimental* feature may be found in the upstream Ceph documentation: <https://docs.ceph.com/en/latest/cephfs/ceph-dokan> 

A Sample configuration files

C:\ProgramData\ceph\ceph.conf

```
[global]
    log to stderr = true
    ; Uncomment the following in order to use the Windows Event Log
    ; log to syslog = true

    run dir = C:/ProgramData/ceph
    crash dir = C:/ProgramData/ceph

    ; Use the following to change the cephfs client log level
    ; debug client = 2
[client]
    keyring = C:/ProgramData/ceph/keyring
    ; log file = C:/ProgramData/ceph/$name.$pid.log
    admin socket = C:/ProgramData/ceph/$name.$pid.asok

    ; client_permissions = true
    ; client_mount_uid = 1000
    ; client_mount_gid = 1000
[global]
    ; Specify IP addresses for monitor nodes as in the following example: ;
    mon host = [v2:10.1.1.1:3300,v1:10.1.1.1:6789] [v2:10.1.1.2:3300,v1:10.1.1.2:6789]
    [v2:10.1.1.3:3300,v1:1.1.1.3:6789]
```



Note

Directory paths in the ceph.conf must be delimited using forward-slashes.

C:\ProgramData\ceph\keyring

```
; This file should be copied directly from /etc/ceph/ceph.client.admin.keyring
; The contents should be similar to the following example:
[client.admin]
    key = ADCyl77eBBAAABDDjX72tAlj0wv04m121v/7yA==
    caps mds = "allow *"
    caps mon = "allow *"
    caps osd = "allow *"
    caps mgr = "allow *"
```

B Troubleshooting tips

If you encounter installation or driver problems, the following tips may be helpful.

Generating an installation log:

```
msiexec /i C:\path\to\ses4win.msi /l*v log.txt
```

You can identify driver loading issues in the Windows driver log:

```
C:\Windows\inf\Setupapi.dev.log
```

To manually uninstall, execute the following:

```
msiexec /x [C:\path\to\ses4win.msi|{GUID}]
```



Note

The GUID can be found under HKLM\Software\Microsoft\Windows\CurrentVersion\Uninstall.

Increase WNDB logging levels through:

```
wnbd-client set-debug 1
```

Basic I/O counters can be monitored through:

```
wnbd-client list  
wnbd-client stats [image_name]
```

C Upstream projects

The Ceph for Microsoft Windows effort is being done entirely in Open Source, and in conjunction with the upstream project(s). For more development-level details, feel free to join the discussion in the following projects:

- Ceph: <https://github.com/ceph/ceph/pull/34859> ↗
- WNBD: <https://github.com/cloudbase/wnbd> ↗
- Ceph Windows Installer: <https://github.com/cloudbase/ceph-windows-installer> ↗

D Ceph maintenance updates based on upstream 'Octopus' point releases

Several key packages in SUSE Enterprise Storage 7 are based on the Octopus release series of Ceph. When the Ceph project (<https://github.com/ceph/ceph>) publishes new point releases in the Octopus series, SUSE Enterprise Storage 7 is updated to ensure that the product benefits from the latest upstream bug fixes and feature backports.

This chapter contains summaries of notable changes contained in each upstream point release that has been—or is planned to be—included in the product.

Octopus 15.2.11 Point Release

This release includes a security fix that ensures the `global_id` value (a numeric value that should be unique for every authenticated client or daemon in the cluster) is reclaimed after a network disconnect or ticket renewal in a secure fashion. Two new health alerts may appear during the upgrade indicating that there are clients or daemons that are not yet patched with the appropriate fix.

To temporarily mute the health alerts around insecure clients for the duration of the upgrade, you may want to run:

```
cephuser@adm > ceph health mute AUTH_INSECURE_GLOBAL_ID_RECLAIM 1h
cephuser@adm > ceph health mute AUTH_INSECURE_GLOBAL_ID_RECLAIM_ALLOWED 1h
```

When all clients are updated, enable the new secure behavior, not allowing old insecure clients to join the cluster:

```
cephuser@adm > ceph config set mon auth_allow_insecure_global_id_reclaim false
```

For more details, refer to <https://docs.ceph.com/en/latest/security/CVE-2021-20288/>.

Octopus 15.2.10 Point Release

This backport release includes the following fixes:

- The containers include an updated `tcmmalloc` that avoids crashes seen on 15.2.9.
- RADOS: BlueStore handling of huge (> 4GB) writes from RocksDB to BlueFS has been fixed.
- When upgrading from a previous cephadm release, `systemctl` may hang when trying to start or restart the monitoring containers. This is caused by a change in the `systemd` unit to use `type=forking`.) After the upgrade, please run:

```
cephuser@adm > ceph orch redeploy nfs
cephuser@adm > ceph orch redeploy iscsi
cephuser@adm > ceph orch redeploy node-exporter
cephuser@adm > ceph orch redeploy prometheus
cephuser@adm > ceph orch redeploy grafana
cephuser@adm > ceph orch redeploy alertmanager
```

Octopus 15.2.9 Point Release

This backport release includes the following fixes:

- MGR: progress module can now be turned on/off, using the commands: `ceph progress on` and `ceph progress off`.
- OSD: PG removal has been optimized in this release.

Octopus 15.2.8 Point Release

This release fixes a security flaw in CephFS and includes a number of bug fixes:

- OpenStack Manila use of `ceph_volume_client.py` library allowed tenant access to any Ceph credential's secret.
- **ceph-volume**: The `lvm batch` subcommand received a major rewrite. This closed a number of bugs and improves usability in terms of size specification and calculation, as well as idempotency behaviour and disk replacement process. Please refer to <https://docs.ceph.com/en/latest/ceph-volume/lvm/batch/> for more detailed information.

- MON: The cluster log now logs health detail every `mon_health_to_clog_interval`, which has been changed from 1hr to 10min. Logging of health detail will be skipped if there is no change in health summary since last known.
- The `ceph df` command now lists the number of PGs in each pool.
- The `bluefs_preextend_wal_files` option has been removed.
- It is now possible to specify the initial monitor to contact for Ceph tools and daemons using the `mon_host_override` config option or `--mon-host-override` command line switch. This generally should only be used for debugging and only affects initial communication with Ceph's monitor cluster.

Octopus 15.2.7 Point Release


This release fixes a serious bug in RGW that has been shown to cause data loss when a read of a large RGW object (for example, one with at least one tail segment) takes longer than one half the time specified in the configuration option `rgw_gc_obj_min_wait`. The bug causes the tail segments of that read object to be added to the RGW garbage collection queue, which will in turn cause them to be deleted after a period of time.

Octopus 15.2.6 Point Release

This releases fixes a security flaw affecting Messenger V2 for Octopus and Nautilus.

Octopus 15.2.5 Point Release

The Octopus point release 15.2.5 brought the following fixes and other changes:

- CephFS: Automatic static sub-tree partitioning policies may now be configured using the new distributed and random ephemeral pinning extended attributes on directories. See the following documentation for more information: <https://docs.ceph.com/docs/master/cephfs/multimds/> 
- Monitors now have a configuration option `mon_osd_warn_num_repaired`, which is set to 10 by default. If any OSD has repaired more than this many I/O errors in stored data a `OSD_T00_MANY_REPAIRS` health warning is generated.

- Now, when `no scrub` and/or `no deep-scrub` flags are set globally or per pool, scheduled scrubs of the type disabled will be aborted. All user initiated scrubs are NOT interrupted.
- Fixed an issue with osdm maps not being trimmed in a healthy cluster.

Octopus 15.2.4 Point Release

The Octopus point release 15.2.4 brought the following fixes and other changes:

- CVE-2020-10753: rgw: sanitize newlines in s3 CORSConfiguration's ExposeHeader
- Object Gateway: The `radosgw-admin` sub-commands dealing with orphans—`radosgw-admin orphans find`, `radosgw-admin orphans finish`, and `radosgw-admin orphans list-jobs`—have been deprecated. They had not been actively maintained, and since they store intermediate results on the cluster, they could potentially fill a nearly-full cluster. They have been replaced by a tool, `rgw-orphan-list`, which is currently considered experimental.
- RBD: The name of the RBD pool object that is used to store RBD trash purge schedule is changed from `rbd_trash_trash_purge_schedule` to `rbd_trash_purge_schedule`. Users that have already started using RBD trash purge schedule functionality and have per pool or name space schedules configured should copy the `rbd_trash_trash_purge_schedule` object to `rbd_trash_purge_schedule` before the upgrade and remove `rbd_trash_purge_schedule` using the following commands in every RBD pool and name space where a trash purge schedule was previously configured:

```
rados -p pool-name [-N namespace] cp rbd_trash_trash_purge_schedule
    rbd_trash_purge_schedule
rados -p pool-name [-N namespace] rm rbd_trash_trash_purge_schedule
```

Alternatively, use any other convenient way to restore the schedule after the upgrade.

Octopus 15.2.3 Point Release

- The Octopus point release 15.2.3 was a hot-fix release to address an issue where WAL corruption was seen when `bluefs_preextend_wal_files` and `bluefs_buffered_io` were enabled at the same time. The fix in 15.2.3 is only a temporary measure (changing the

default value of `bluefs_preextend_wal_files` to `false`). The permanent fix will be to remove the `bluefs_preextend_wal_files` option completely: this fix will most likely arrive in the 15.2.6 point release.

Octopus 15.2.2 Point Release

The Octopus point release 15.2.2 patched one security vulnerability:

- CVE-2020-10736: Fixed an authorization bypass in MONs and MGRs

Octopus 15.2.1 Point Release

The Octopus point release 15.2.1 fixed an issue where upgrading quickly from Luminous (SES5.5) to Nautilus (SES6) to Octopus (SES7) caused OSDs to crash. In addition, it patched two security vulnerabilities that were present in the initial Octopus (15.2.0) release:

- CVE-2020-1759: Fixed nonce reuse in msgr V2 secure mode
- CVE-2020-1760: Fixed XSS because of RGW GetObject header-splitting

Glossary

General

Admin node

The host from which you run the Ceph-related commands to administer cluster hosts.

Alertmanager

A single binary which handles alerts sent by the Prometheus server and notifies the end user.

archive sync module

Module that enables creating an Object Gateway zone for keeping the history of S3 object versions.

Bucket

A point that aggregates other nodes into a hierarchy of physical locations.

Ceph Client

The collection of Ceph components which can access a Ceph Storage Cluster. These include the Object Gateway, the Ceph Block Device, the CephFS, and their corresponding libraries, kernel modules, and FUSE clients.

Ceph Dashboard

A built-in Web-based Ceph management and monitoring application to administer various aspects and objects of the cluster. The dashboard is implemented as a Ceph Manager module.

Ceph Manager

Ceph Manager or MGR is the Ceph manager software, which collects all the state from the whole cluster in one place.

Ceph Monitor

Ceph Monitor or MON is the Ceph monitor software.

Ceph Object Storage

The object storage "product", service or capabilities, which consists of a Ceph Storage Cluster and a Ceph Object Gateway.

Ceph OSD Daemon

The **ceph-osd** daemon is the component of Ceph that is responsible for storing objects on a local file system and providing access to them over the network.

Ceph Storage Cluster

The core set of storage software which stores the user's data. Such a set consists of Ceph monitors and OSDs.

ceph-salt

Provides tooling for deploying Ceph clusters managed by cephadm using Salt.

cephadm

cephadm deploys and manages a Ceph cluster by connecting to hosts from the manager daemon via SSH to add, remove, or update Ceph daemon containers.

CephFS

The Ceph file system.

CephX

The Ceph authentication protocol. Cephx operates like Kerberos, but it has no single point of failure.

CRUSH rule

The CRUSH data placement rule that applies to a particular pool or pools.

CRUSH, CRUSH Map

Controlled Replication Under Scalable Hashing: An algorithm that determines how to store and retrieve data by computing data storage locations. CRUSH requires a map of the cluster to pseudo-randomly store and retrieve data in OSDs with a uniform distribution of data across the cluster.

DriveGroups

DriveGroups are a declaration of one or more OSD layouts that can be mapped to physical drives. An OSD layout defines how Ceph physically allocates OSD storage on the media matching the specified criteria.

Grafana

Database analytics and monitoring solution.

Metadata Server

Metadata Server or MDS is the Ceph metadata software.

Multi-zone

Node

Any single machine or server in a Ceph cluster.

Object Gateway

The S3/Swift gateway component for Ceph Object Store. Also known as the RADOS Gateway (RGW).

OSD

Object Storage Device: A physical or logical storage unit.

OSD node

A cluster node that stores data, handles data replication, recovery, backfilling, rebalancing, and provides some monitoring information to Ceph monitors by checking other Ceph OSD daemons.

PG

Placement Group: a sub-division of a *pool*, used for performance tuning.

Point Release

Any ad-hoc release that includes only bug or security fixes.

Pool

Logical partitions for storing objects such as disk images.

Prometheus

Systems monitoring and alerting toolkit.

RADOS Block Device (RBD)

The block storage component of Ceph. Also known as the Ceph block device.

Reliable Autonomic Distributed Object Store (RADOS)

The core set of storage software which stores the user's data (MON + OSD).

Routing tree

A term given to any diagram that shows the various routes a receiver can run.

Rule Set

Rules to determine data placement for a pool.

Samba

Windows integration software.

Samba Gateway

The Samba Gateway joins the Active Directory in the Windows domain to authenticate and authorize users.

zonegroup