Red Hat
Data Services

What's New in Red Hat Storage

ODF, RHCS, ODS

October 2021

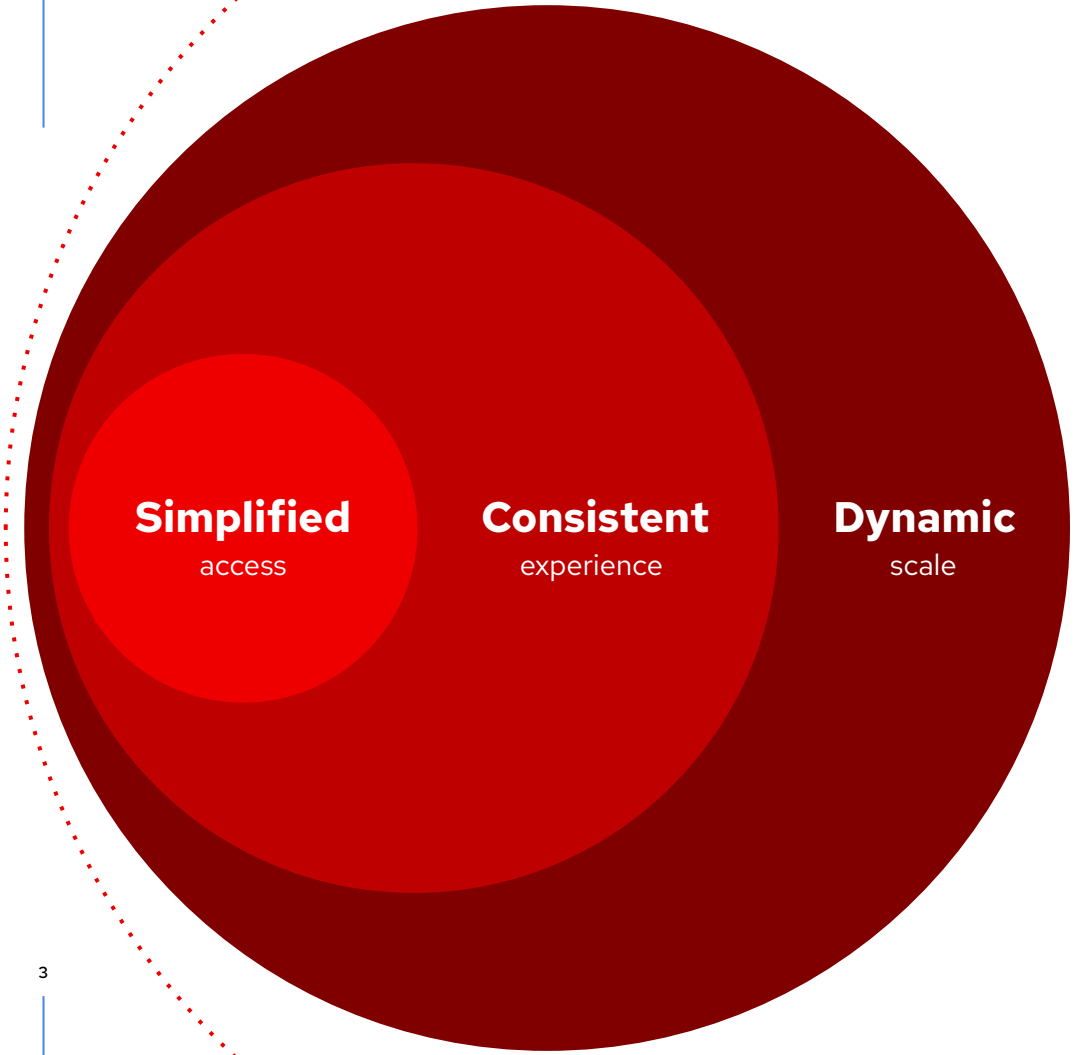
Shawn Houston
Principal Specialist Solutions Architect

Red Hat
Data Services

[What's New] Red Hat Storage [Oct-2021]

MN-WI RHUG

- Please feel free to ask questions at any time
- ODF: Openshift Data Foundation
  - Recently renamed, also known as Openshift Container Storage (OCS)
- RHCS: Red Hat Ceph Storage
- RHGS and RHHI will not be covered here
  - Both are facing upcoming end of support (EOS) for the products
- ODS: Open Data Science



Red Hat Data Services mission:

To make data accessible to applications across the hybrid cloud, unlocking its power in new and impactful ways

Delivering on the Red Hat OpenShift promise:

Innovation without limitation

Introducing Red Hat Data Services

About Tech preview

Tech Preview

Provides early access to upcoming product innovations, enabling customers to test functionality and provide feedback during the development process.

These features are not fully supported under Red Hat Subscription Level Agreements, may not be functionally complete, and are not intended for production use.

As Red Hat considers making future iterations of Technology Preview features generally available, we will attempt to resolve any issues that customers experience when using these features.

Source

About Dev preview

Dev Preview

Development Preview releases are meant for customers who are willing to evaluate new products or releases of products in an early stage of product development.

It’s a vehicle for developers that provides early access to new unreleased features.

These features are not supported under Red Hat Subscription Level Agreements, may not be functionally complete, and are not intended for production use.

Dev preview features are also not documented in the official release documentation.

Red Hat
Data Services

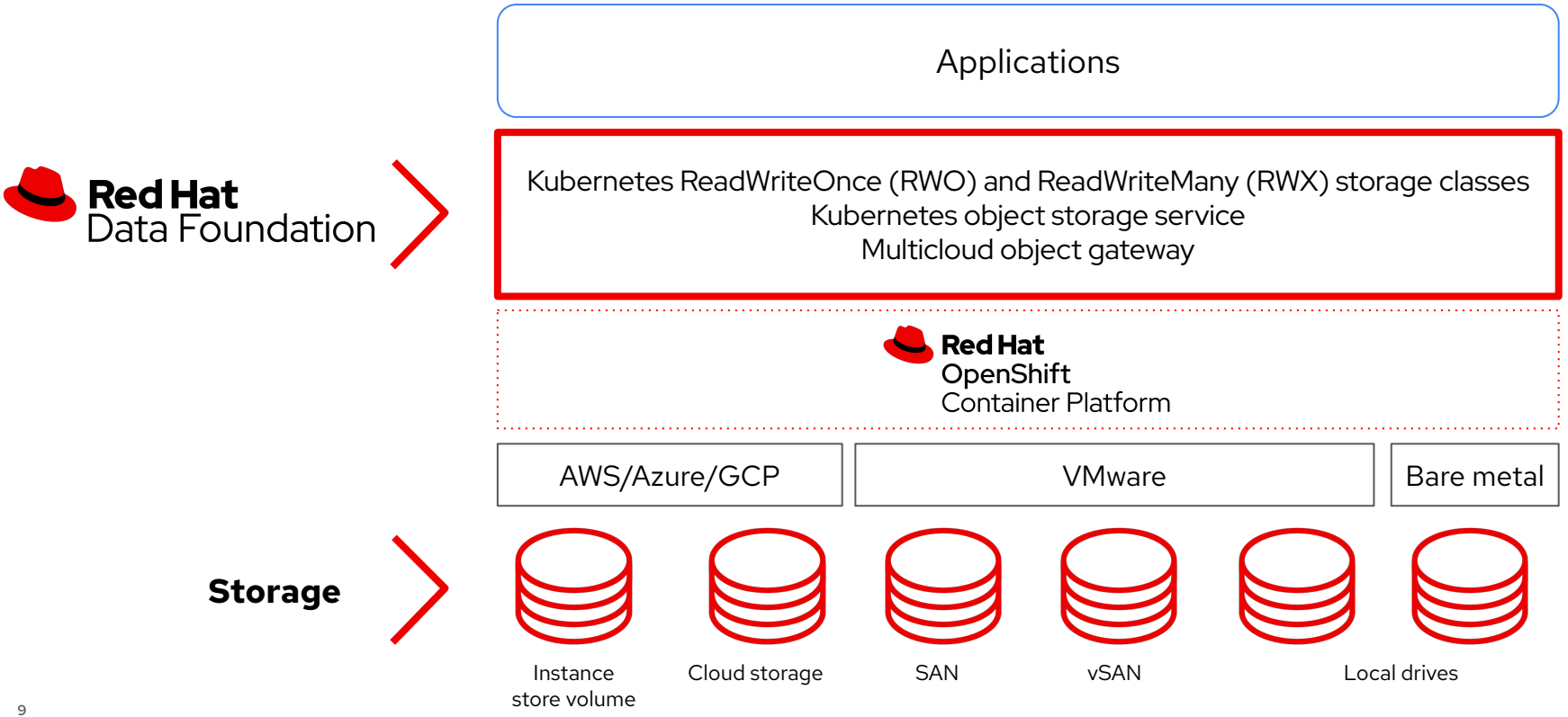
What's New in Red Hat OpenShift Data Foundation 4.8

Red Hat
Data Services

Data is the most significant asset in today's businesses—give it data foundation

- Data foundation focuses on infrastructure and application needs so they can run and interact with ease and efficiency
- Provides a foundational data layer for applications to function and interact with data in a simplified, consistent and scalable manner
- Red Hat Ceph Storage is a foundational component to drive data services

The Red Hat Data Foundation stack



What's new?

Data resilience with Red Hat OpenShift Data Foundation 4.8

FUNCTIONALITY

Greater control and manageability with about 10 new functional features

SECURITY

Enhanced protection with data encryption for RBD and additional protection with snapshotting and cloning

PERFORMANCE

Improved segregation of storage and network resources. Faster upgrade by component rescheduling improvement

EFFICIENCY

Extended flexibility by component selectability and new caching capabilities

What's new?

Red Hat OpenShift Data Foundation 4.8

FUNCTIONALITY

SECURITY

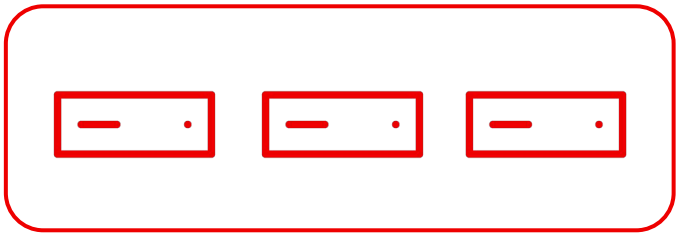
PERFORMANCE

EFFICIENCY

FUNCTIONALITY

Compact mode
with Red Hat OpenShift Data Foundation

Run Red Hat OpenShift including OpenShift Data Foundation deployed on three nodes in production, without distinct compute or worker nodes and inclusive storage



What's new?

Red Hat OpenShift Data Foundation 4.8

Tech Preview

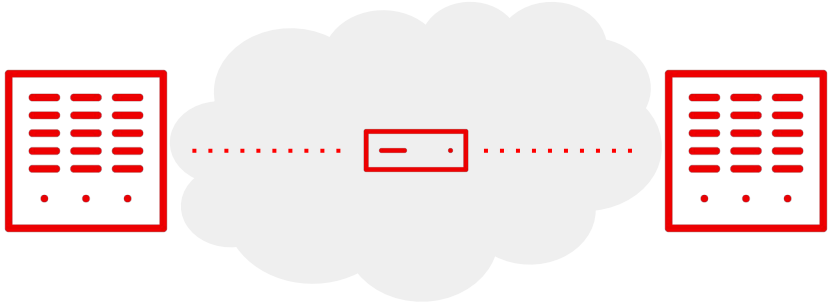
FUNCTIONALITY

Tech Preview

Metro DR-stretch cluster

Stretched cluster with arbiter

No data-loss recovery when only two data centers can be used. An arbiter will be used to get a valid quorum between the two data centers.
This concept enables for near-zero recovery point objective (RPO).
Recovery times vary, based on the volume type.



What's new?

Red Hat OpenShift Data Foundation 4.8

Dev Preview

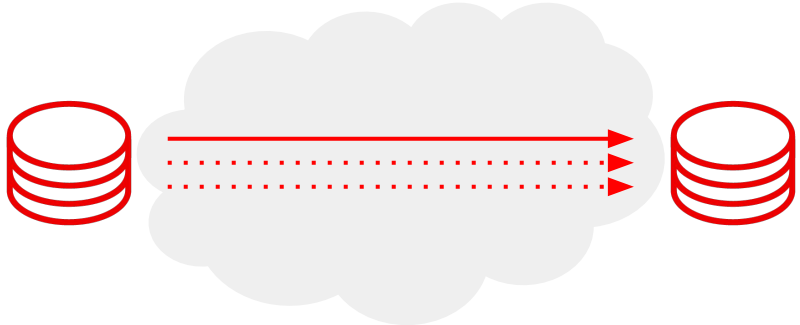
FUNCTIONALITY

Regional DR

Multi cluster persistent block volume async replication

Disaster recovery for persistent **block** volumes, using differential data for data transfer and time efficiency. Recovery point objective (RPO) and recovery time objective (RTO) times are in mins.

Capability for use with higher latency connections like WAN



What's new?

Red Hat OpenShift Data Foundation 4.8

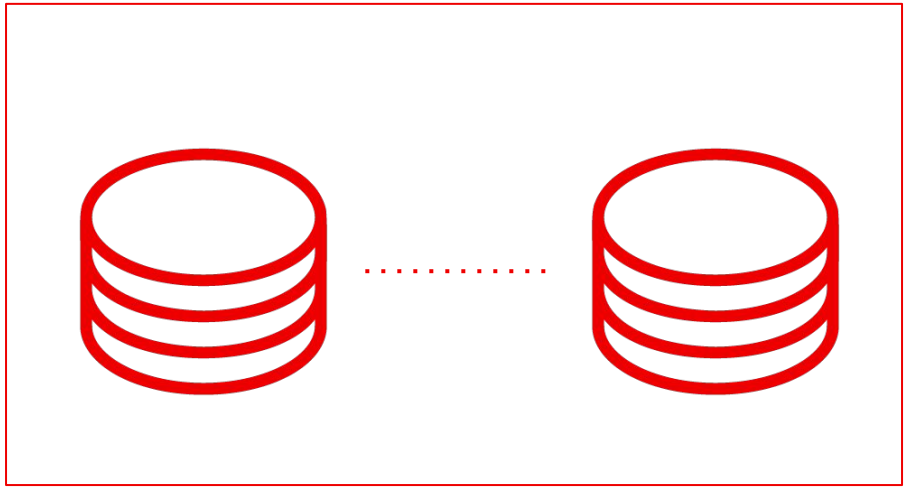
Dev Preview

FUNCTIONALITY

- CephFS
- RADOS block device

Replica-2

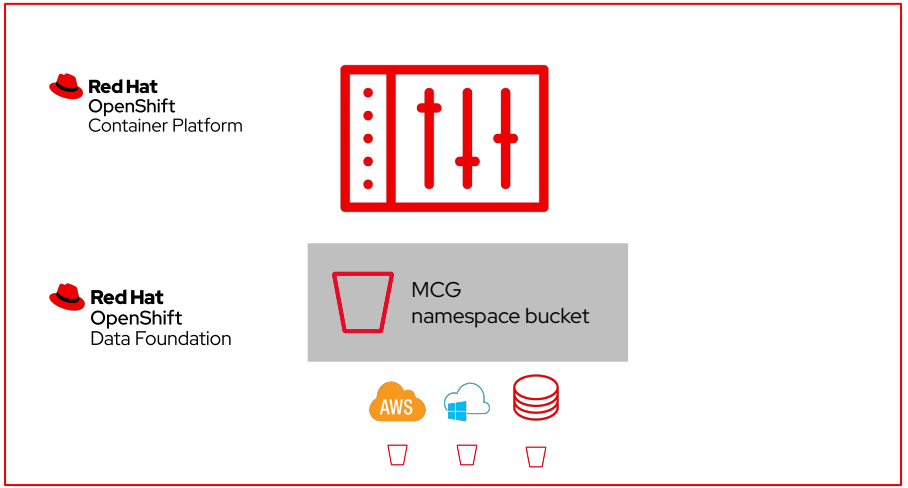
Two fold replication for the entire cluster



FUNCTIONALITY

Multicloud Object Gateway

UI option for MCG Namespace bucket class and backing store



What's new?

Red Hat OpenShift Data Foundation 4.8

Tech Preview

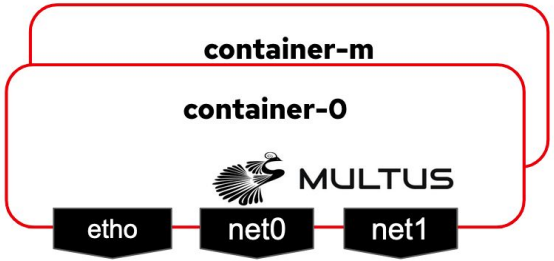
FUNCTIONALITY

MULTUS

Multi Network Plugin–Multus

Provides network isolation by enabling data plane and control plane separation

Ability to improve security and performance by isolating networks.



Red Hat
Data Services

FUNCTIONALITY

Recovery with a few commands

Supportability–recover from a full cluster failure event

Provides a way to recover quickly

Red Hat provides a job template containing simple instructions to help customers recover quickly

What's new?

Red Hat OpenShift Data Foundation 4.8

FUNCTIONALITY

SECURITY

PERFORMANCE

EFFICIENCY

What's new?

Red Hat OpenShift Data Foundation 4.8

Dev Preview

PERFORMANCE

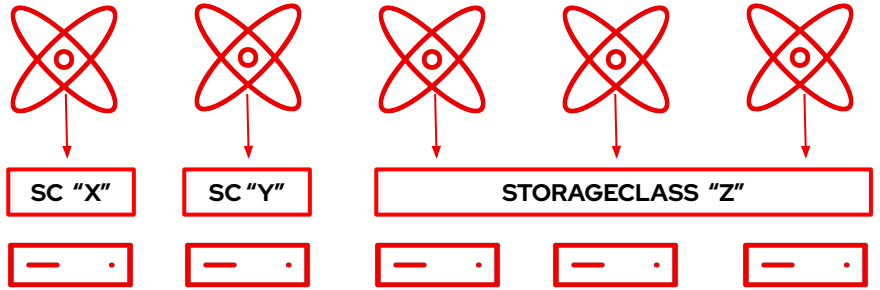
Enables for both security and resource fencing

Restricting workload access to specific physical disk device groups

Data segregation

Data segregation per group of hosts

Provides a way to isolate I/O between workloads using a specific node or nodes group and storageclass per workload



Red Hat
Data Services

What's new?

Red Hat OpenShift Data Foundation 4.8

FUNCTIONALITY

SECURITY

PERFORMANCE

EFFICIENCY

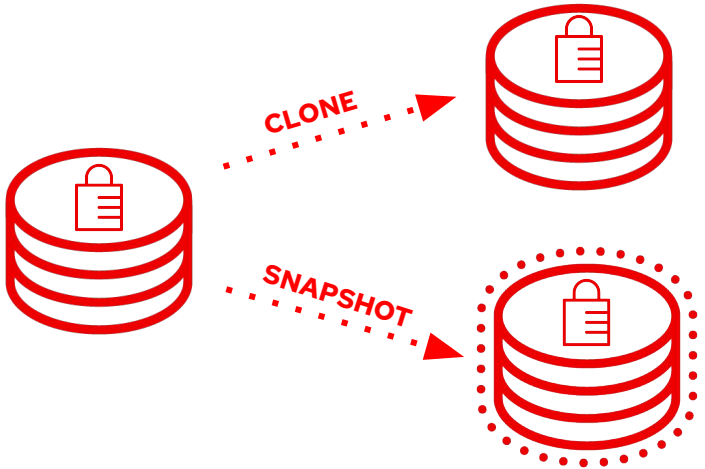
SECURITY

OpenShift Data Foundation 4.7
capability to encrypt PVs

OpenShift Data Foundation 4.8
supports encrypted snapshots and clones

Enhanced Block Device persistent volume encryption

Enhanced RBD PV encryption with the ability to clone the volume and take a snapshot



What's new?

Red Hat OpenShift Data Foundation 4.8

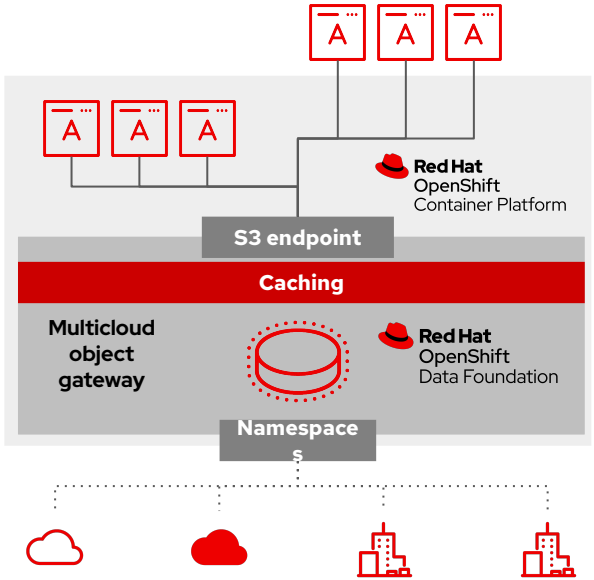
FUNCTIONALITY

SECURITY

PERFORMANCE

EFFICIENCY

EFFICIENCY



Multicloud object gateway (MCG)
Caching support

A caching object solution for customers where data gravity is required. This is particularly useful for those using artificial intelligence/machine learning (AI/ML) platforms.

What's new?

Red Hat OpenShift Data Foundation 4.8

Dev Preview

EFFICIENCY

Flexibility in components deployment

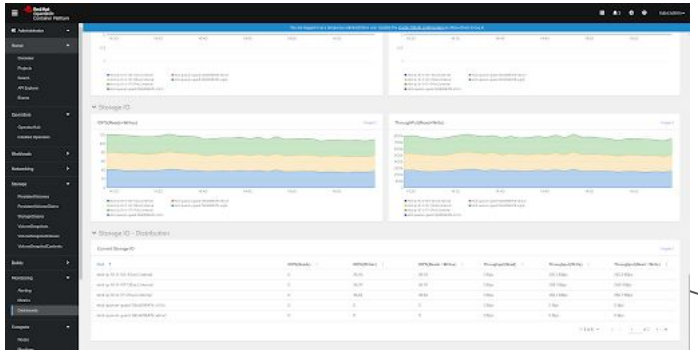
More flexibility in deployment,
choice for components to become installed.

Lowering resources allocation and subscription cost tied to required resources

Block only

https://access.redhat.com/articles/6106741

EFFICIENCY



TOP utility–viewing pods I/O metrics

Ability to drill down when there is a load or overload situation on a system

Pods level performance information helps finding "noisy" applications

What's new?
Red Hat OpenShift Data Foundation 4.8

SUMMARY

General Available √

- Compact Mode (for Edge)
- VMWare IPI provisioning
- Block encryption extended with snap and clone
- Easy pools management
- Multicloud object gateway User Interface option (new) and caching feature (TP in 4.7)
- Supportability—recover from a full cluster failure event
- TOP IO metrics for pods

Tech Preview

- Metro-DR stretch cluster
- Multi Network Plugin (Multus)
- Object Storage Daemon Weight option

Dev Preview

- Block Device thick provisioning
- Regional-DR (for RBD)
- VMware thick storageclass
- Replica-2 for the entire cluster (RBD and CephFS)
- Data segregation
- Flexible component deployment

Red Hat
Data Services

What's New in Red Hat Ceph Storage 5.0

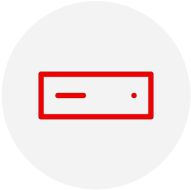
Red Hat
Data Services

Red Hat Data Services

Red Hat
Ceph Storage

- Software defined storage for on-premise cloud buildout
- Massively scalable to support tens of petabytes of data
- Delivers solid reliability and data durability
- Storage with industry-standard x86 servers
- Multi-site aware and disaster-recovery enabled

Flexibility to meet the demands of tomorrow



Uses industry standard hardware

- Runs on industry standard servers and TCP/IP networks
- HDDs, SSDs, NVMe

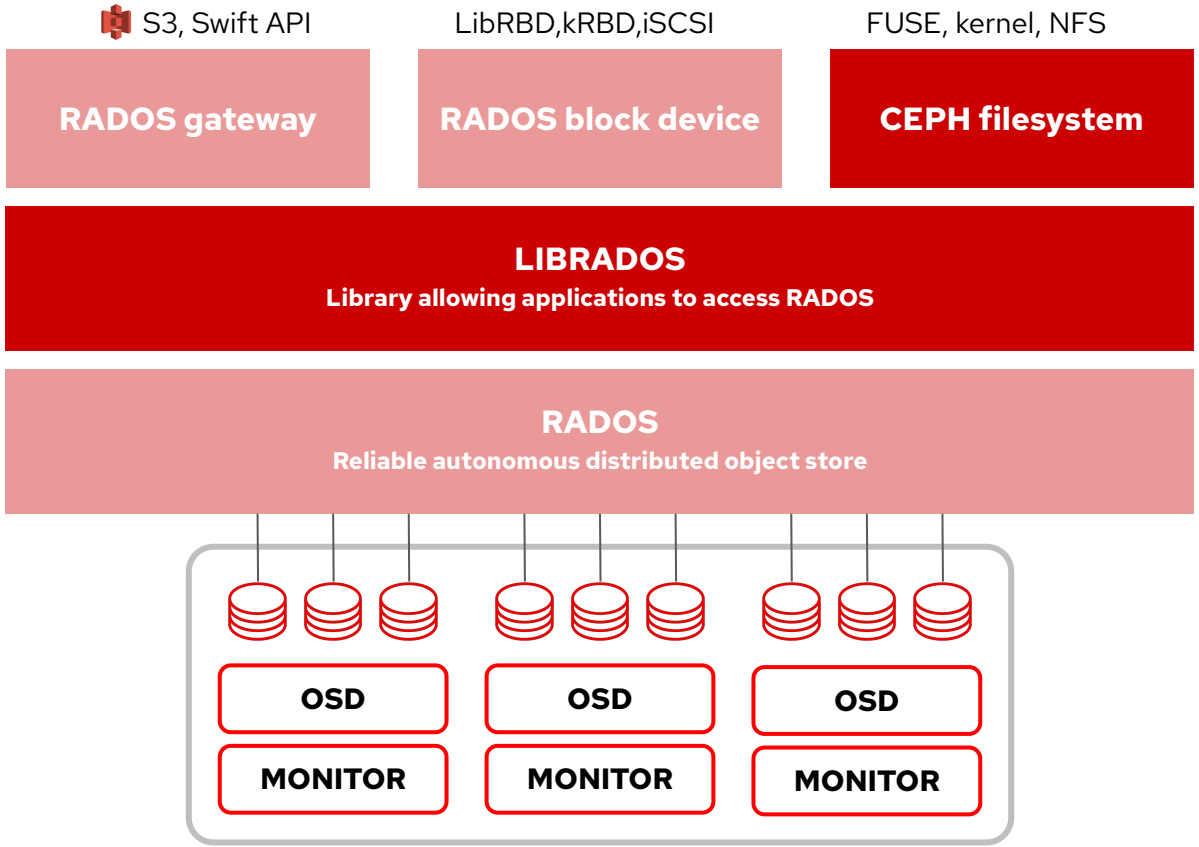
Delivers reliability

- Fully distributed, no single point of failure
- Ensure data durability via replication or erasure coding
- Expand or shrink clusters as required
- Federate multiple clusters across sites with asynchronous replication and disaster recovery capabilities

Improves versatility

- A single cluster can support object, block, and file workloads
- Scale out within a cluster for capacity/speed
- Add or remove hardware while system is online —even if it's under load
- Apply updates without interrupting service

Ceph architecture



What's new?

Red Hat Ceph Storage 5

FUNCTIONALITY

New integrated control plane
Stable management API
NFS filesystem support

SECURITY

WORM (object lock)
FIPS 140-2 Cryptography
Key management integration

PERFORMANCE

80% increase in block performance for virtual machine and container hosting

EFFICIENCY

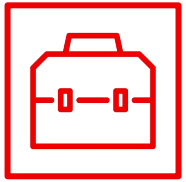
Reduced resource consumption for small file
Complete set of data reduction options

Red Hat
Data Services

What's new?

Red Hat Ceph Storage 5

FUNCTIONALITY



SECURITY

PERFORMANCE

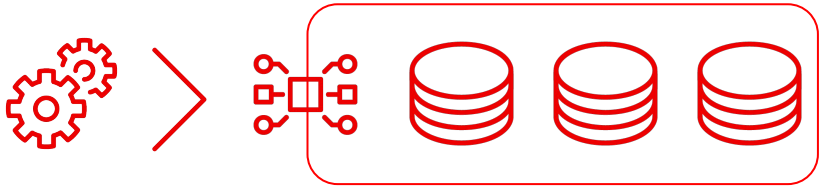
EFFICIENCY

FUNCTIONALITY

Manageability features

Integrated control plane (Cephadm)
Stable management API
OSD replacement workflows (UI and CLI)
Object multi-site monitoring

Management scripts written for a major version of Red Hat Ceph Storage will continue to operate unchanged for the version's life cycle.



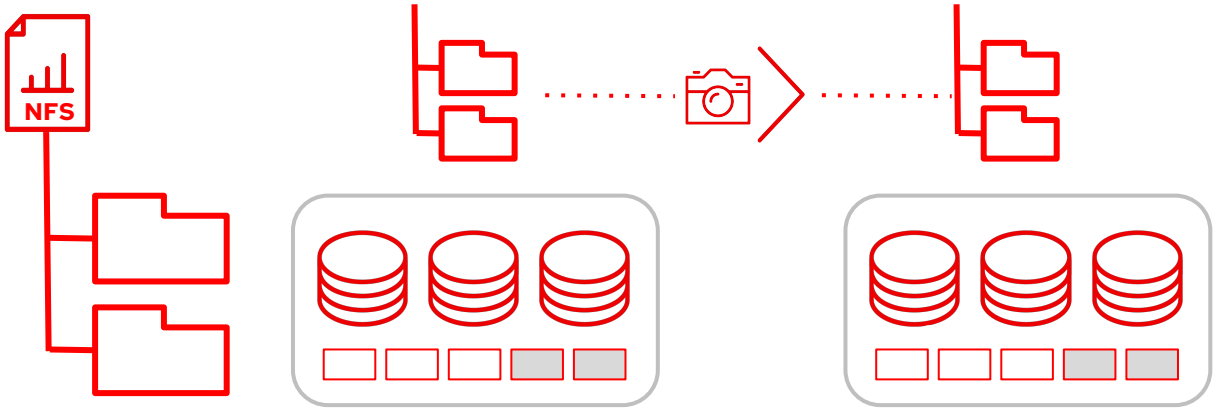
FUNCTIONALITY

New Ceph filesystem capabilities

NFS access option

Erasure code option

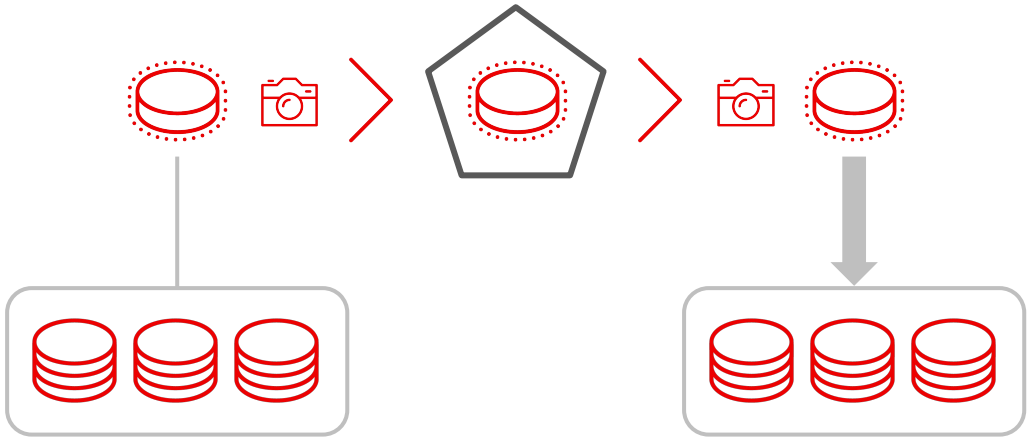
Snapshot based geo replication



FUNCTIONALITY

New RADOS block device capabilities

RBD snapshot-based migration across clusters



What's new?

Red Hat Ceph Storage 5

FUNCTIONALITY

SECURITY

PERFORMANCE

EFFICIENCY

PERFORMANCE

Improved performance

Dramatic boost for virtual machines:
Improved block performance by 80%

LibRBD data path optimization

Improvements shown in OpenStack benchmark

Red Hat OpenStack Platform

PERFORMANCE

Improved performance
Overhauled cache architecture

- New read-only large object cache
  Offloads reads of objects from the cluster
- Improved in-memory write-around cache
  No longer serves reads (page cache), dedicated to batching writes
- Optane write-back cache
  Dramatically reducing latency

Tech Preview

PERFORMANCE

Improved performance

New object benchmark HDD test results:
>80 GB/s aggregate throughput

Improved scale

10+ billion objects in RADOS object gateway
1.5 billion objects per cluster node

"We were able to achieve a staggering 79.6 GiB/s aggregate throughput from the 10-node Ceph cluster utilized for our testing."
That's 85.5GB/sec from a disk-based data set composed of 350 million objects.

Kyle Bader, Red Hat

SEAGATE

Sources:
ceph.io blog article by Kyle Bader
blocksandfiles.com article (March 31, 2021)
scientific-computing.com article (April 1, 2021)

PERFORMANCE

IO500

Virtual Institute for IO reference link

Improved performance– CephFS ephemeral pinning

Metadata scalability enhancement

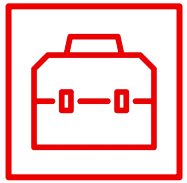
Improves the ability of multiple MDS servers to balance the load with a temporary setting that distributes the load "round robin" across all metadata servers.

Manual operator intervention, where a volume or directory is identified as a hotspot. The intervention is one-time instead of continuous in the current model.

What's new?

Red Hat Ceph Storage 5

FUNCTIONALITY



SECURITY

PERFORMANCE

EFFICIENCY

SECURITY

Object security and governance

S3 object lock provides read-only capability

S3 object lock to store objects using a write-once-read-many (WORM) model.

Object lock can help prevent objects from being deleted or overwritten for a fixed amount of time or indefinitely.

Standard WORM certification is planned for the coming year

SECURITY

Federal Information Processing Standard
FIPS 140-2 cryptography

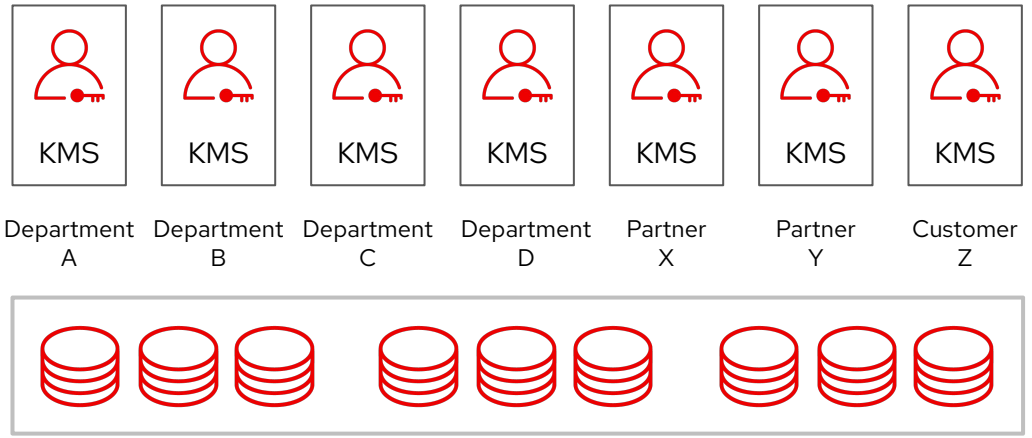
Red Hat Ceph Storage can use FIPS 140-2 validated cryptography modules when run on Red Hat Enterprise Linux 8.1

Newer versions are certified on DISA's schedule

SECURITY

External key manager integration

Key management service integration with Hashicorp Vault, IBM SGKLM, OpenStack Barbican, OpenID Connect (OIC) identity support



What's new?

Red Hat Ceph Storage 5

FUNCTIONALITY

SECURITY

PERFORMANCE

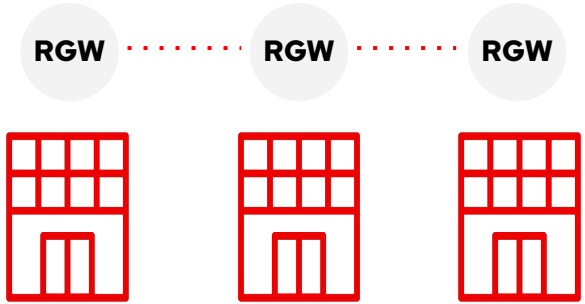
EFFICIENCY

EFFICIENCY

Multi-site capabilities

RADOS object gateway across sites

Currently scalable in the tens of Petabytes

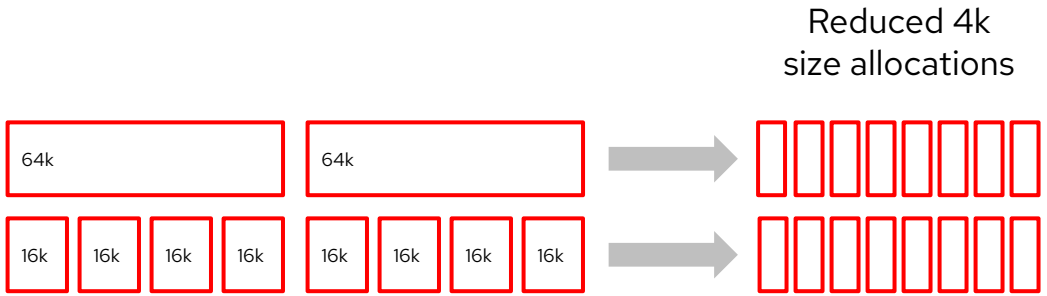


EFFICIENCY

Resource consumption

Improved space utilization for small objects

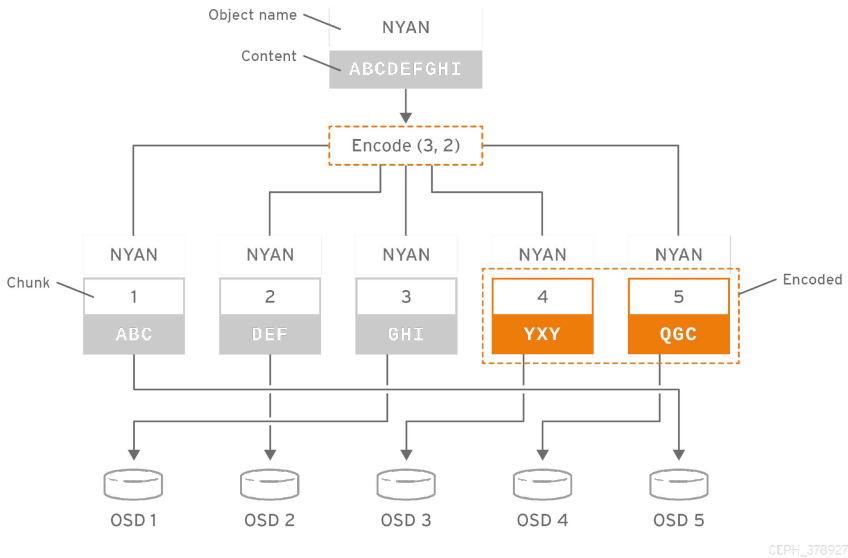
Bluestore 4k allocation size replacing the current 64k (HDD) and 16k (SSD) when managing small objects, significantly reducing overhead for storage of small objects.



EFFICIENCY

Improved reliability

Faster erasure coding recovery



Red Hat
Data Services

Introducing
Red Hat Openshift Data Science

Red Hat
Data Services

Based on Open Data Hub and Operate First

Upstream code enhanced with operational excellence

Open Data Hub

Community driven upstream meta-project demonstrating AI/ML platform on Red Hat OpenShift comprised of open source projects

Operate First

Subset Open Data Hub operated at scale for community and university audiences to infuse operational excellence

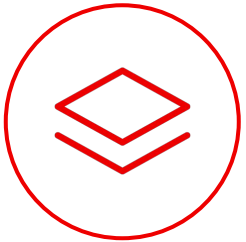
Red Hat OpenShift Data Science

Subset of Operate First delivered as a cloud service on Red Hat OpenShift Managed on Amazon Web Services with optional ISV offerings

Open Data Hub: https://opendatahub.io
Operate First: https://openinfralabs.org/ and https://www.operate-first.cloud/

Depth and scale without lock-in

Capabilities delivered through the combination of Red Hat and partner ecosystem

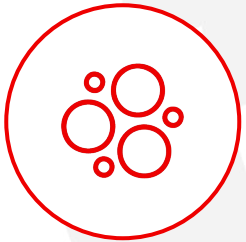


Managed cloud platform

Deployed on Red Hat OpenShift and managed on Amazon Web Services providing access to compute and accelerators based on your workload

Red Hat portfolio and services

Complement common data science tools in Red Hat OpenShift Data Science with other Red Hat products and cloud services



Partner ecosystem

Access specialized capabilities by adding certified ISV ecosystem products and services from Red Hat Marketplace

Key features of Red Hat OpenShift Data Science

Addressing AI/ML experimentation and integration use cases on a managed platform

Cloud Service

Available on Red Hat OpenShift Dedicated (AWS) and Red Hat OpenShift Service on AWS

Core data science workflow

Provides data scientists and intelligent application developers the ability to build, train, and deploy ML models

Increased capabilities/collaboration

Combines Red Hat components, open source software, and ISV certified software available on Red Hat Marketplace

Rapid experimentation use cases

Model outputs are hosted on the Red Hat OpenShift managed service or exported for integration into an intelligent application

Red Hat OpenShift Data Science

Tools and capabilities

Building on the foundations of data science

Jupyter notebooks

Conduct exploratory data science in JupyterLab with access to core AI / ML libraries and frameworks including TensorFlow and PyTorch.

Source-to-image (S2I)

Publish models as end points via S2I for integration into intelligent apps. Rebuild and redeploy based on changes to the source code.

GPU Acceleration (post-GA)

Accelerate your data science experiments through the use of GPU acceleration on the Red Hat OpenShift Dedicated platform.

Timeline

Beta

- Red Hat Summit announcement and demo
- Start Beta program with small group of customers
- Announce aligned consulting services

Future (Enhancements TBD)

- Support other cloud platforms
- Enhance on-premise automation
- Additional AI / ML ISVs
- Enhance MLOps: KFServing, Kubeflow Pipelines

April 2021 *Early customer adoption*

2H 2021

TBD

First release: Version 1

- As a cloud service add-on to OpenShift Dedicated (on AWS initially) or Red Hat OpenShift Service on AWS
- Initial AI / ML ISV partners
- Available via Red Hat Marketplace

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.


linkedin.com/company/red-hat


youtube.com/user/RedHatVideos


facebook.com/redhatinc


twitter.com/RedHat

Red Hat
Data Services