Advanced Technology Group

Accelerate with ATG

A Ceph Primer

The difference between IBM Storage Ceph and Ceph in IBM Storage Fusion

John Shubeck – ATG Storage Technical Specialist

Date: July 25, 2023

© Copyright IBM Corporation 2023

Accelerate with ATG Technical Webinar Series

Advanced Technology Group experts cover a variety of technical topics.

Audience: Clients who have or are considering acquiring IBM Storage solutions. Business Partners and IBMers are also welcome.

To automatically receive announcements of upcoming Accelerate with IBM Storage webinars, Clients, Business Partners and IBMers are welcome to send an email request to accelerate-join@hursley.ibm.com.

2023 Upcoming Webinars – click on the link to register for the live event:

August 1 – Data Resiliency with IBM Storage Scale

August 22 – Introduction to IBM's newest Tape Storage, the IBM Diamondback Tape Library

August 29 – IBM Storage Virtualize 8.6 and Storage Sentinel Technical Update

Important Links to bookmark:

ATG Accelerate Support Site: https://www.ibm.com/support/pages/node/1125513

ATG MediaCenter Channel: https://ibm.biz/BdfEgQ

© Copyright IBM Corporation 2023

ATG-Storage Offerings

CLIENT WORKSHOPS

- IBM DS8900F Advanced Functions
- IBM Storage Point of View on Cyber Resiliency
- IBM FlashSystem and Storage Virtualize
- IBM Storage for Data and AI
- **IBM FlashSystem 9500 Deep Dive & Advanced Functions – August 2-3, 2023, in Raleigh, NC**
- IBM Storage Fusion

Please reach out to your IBM Rep or Business Partner for future dates and to be nominated.

TEST DRIVE / DEMO'S

- North America ATG Storage - IBM Storage Scale and Storage Scale System GUI
- North America ATG Storage - IBM Storage Virtualize Test Drive
- North America ATG Storage - IBM DS8900F Storage Management Test Drive
- North America ATG Storage - Managing Copy Services on the DS8000 Using IBM Copy Services Manager Test Drive
- North America ATG Storage - IBM DS8900F Safeguarded Copy (SGC) Test Drive
- North America ATG Storage - IBM Cloud Object Storage Test Drive - (Appliance based)
- North America ATG Storage - IBM Cloud Object Storage Test Drive - (VMware based)
- North America ATG Storage - IBM Storage Protect Live Test Drive
- North America ATG Storage - IBM Storage Protect Plus Live Test Drive
- North America ATG Storage - IBM Storage Ceph Test Drive - (VMware based)

Please reach out to your IBM Rep or Business Partner for more information.

© Copyright IBM Corporation 2023

Accelerate with ATG Technical Webinar Series - Survey

Please take a moment to share your feedback with our team!

You can access this 6-question survey via Menti.com with code 2243 3599 or

Direct link https://www.menti.com/albneqj15g57

Or

QR Code

© Copyright IBM Corporation 2023

Advanced Technology Group

Accelerate with ATG

A Ceph Primer

The difference between IBM Storage Ceph and Ceph in IBM Storage Fusion

John Shubeck – ATG Storage Technical Specialist

Date: July 25, 2023

© Copyright IBM Corporation 2023

Meet the Speakers

Shu Mookerjee is a Level 2 Certified Technical Specialist with over twenty years at IBM, working in a variety of roles including sales, management and technology. For the last decade, he has focused exclusively on storage and has been the co-author of four (4) Redbooks. Currently, Shu is part of the Advanced Technology Group where he provides education, technical guidance, Proofs of Concept and Proofs of Technology to IBMers, business partners and clients.

John Shubeck is an information technology professional with over 41 years of industry experience spanning both the customer and technology provider experience. John is currently serving as a Senior Storage Technical Specialist on IBM Object Storage platforms across all market segments in the Americas.

© Copyright IBM Corporation 2023

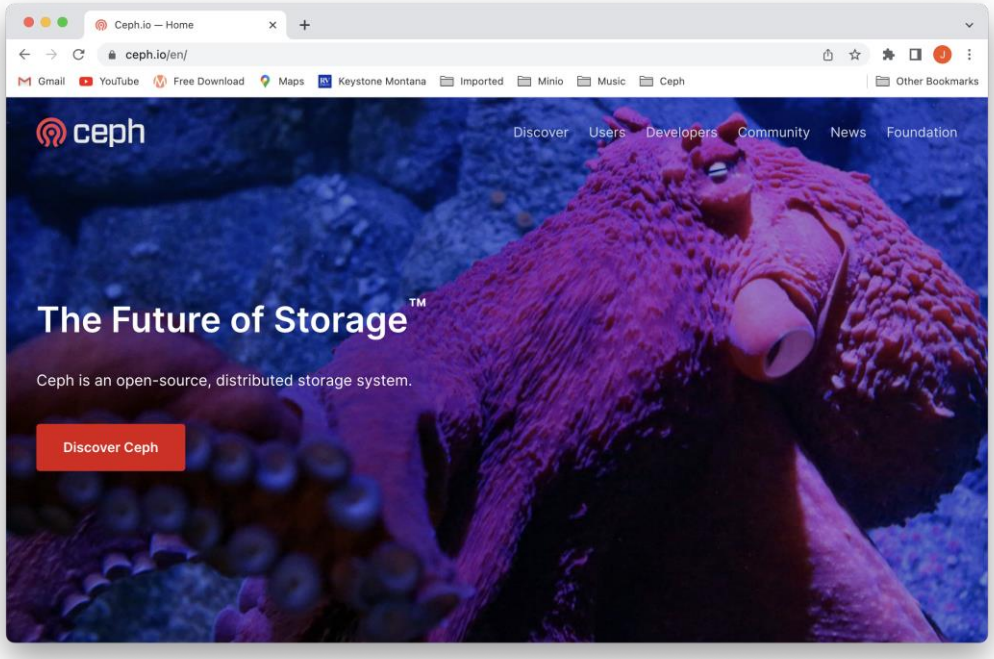
Agenda

- What is Ceph
- What is IBM Storage Ceph
- IBM Storage Ceph use cases
- IBM Storage Ceph Basics
- IBM Storage Ceph Deployment
- IBM Technology Zone Ceph Test Drive
- The flexibility of IBM Storage Ceph
- Why IBM Storage Ceph

© Copyright IBM Corporation 2023

What is Ceph?

Distributed, enterprise-grade universal storage, proven at scale



Source: Ceph Foundation (ceph.io)

Open source, massively scalable, software-defined storage based on Ceph

Flexible, scale-out architecture on clustered standard hardware

Single, efficient, unified storage platform for distributed object, block and file

User-driven storage life-cycle management with consistent API coverage between versions

Storage designed for modern workloads like cloud infrastructure, analytics, and AI/ML, data lakes

© Copyright IBM Corporation 2023

What is IBM Storage Ceph?

IBM Storage Ceph is an open source distributed software-defined storage solution that allows for data consumption through several interfaces, such as object, block and file.

It is packaged as two deployment options and is fully supported by IBM.

IBM Storage Ceph

© Copyright IBM Corporation 2023

IBM Storage Ceph Offering

IBM Storage Ceph and Red Hat Ceph packaging

IBM Storage Ceph
On-prem S3 storage at scale and performance

- Object storage
- Block storage
- File storage
- Presence at the on-prem object market at 10-Petabyte+ scale
- S3 compatibility with AWS

Red Hat OpenStack Platform

Red Hat Ceph Storage

Ceph for OpenStack
 # 1 in OpenStack storage

- Cinder block storage
- Nova ephemeral storage
- Glance image storage
- Swift object store
- Manila file storage
- Advanced integration
- Unified management
- Hyperconverged and Edge capabilities

IBM Storage Fusion
Ceph for OpenShift

- Self-managing storage powered by Red Hat Ceph Storage
- Automated by Rook and completed with Multicloud object gateway (MCG)
- Advanced integration, automation, ease of use
- Persistent storage for OpenShift stateful workloads

© Copyright IBM Corporation 2023

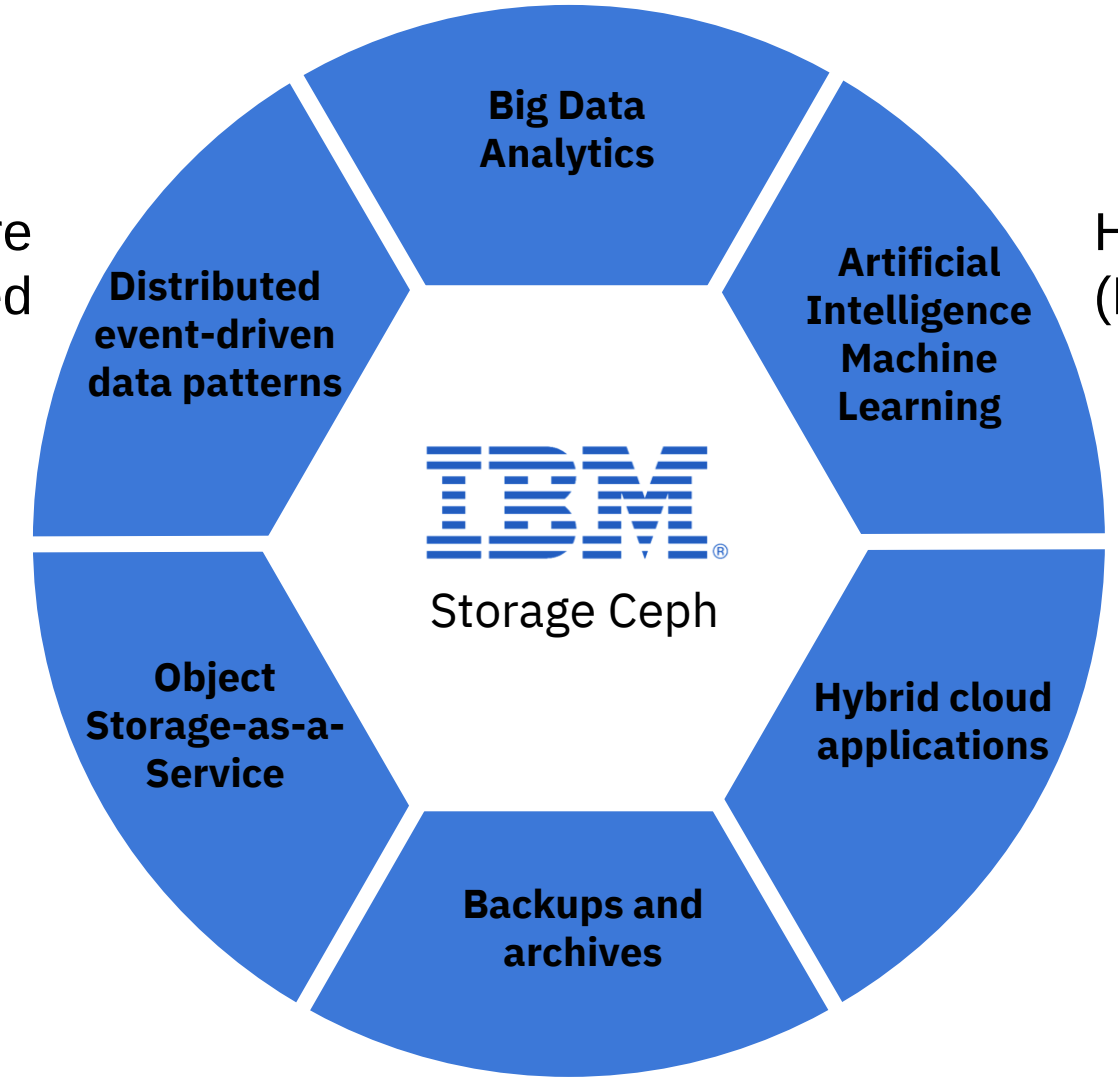
IBM Storage Ceph Use Cases

Software Defined

Hardware Independent (HDD, SSD, NVMe)

S3 API Compliance Multi Protocol

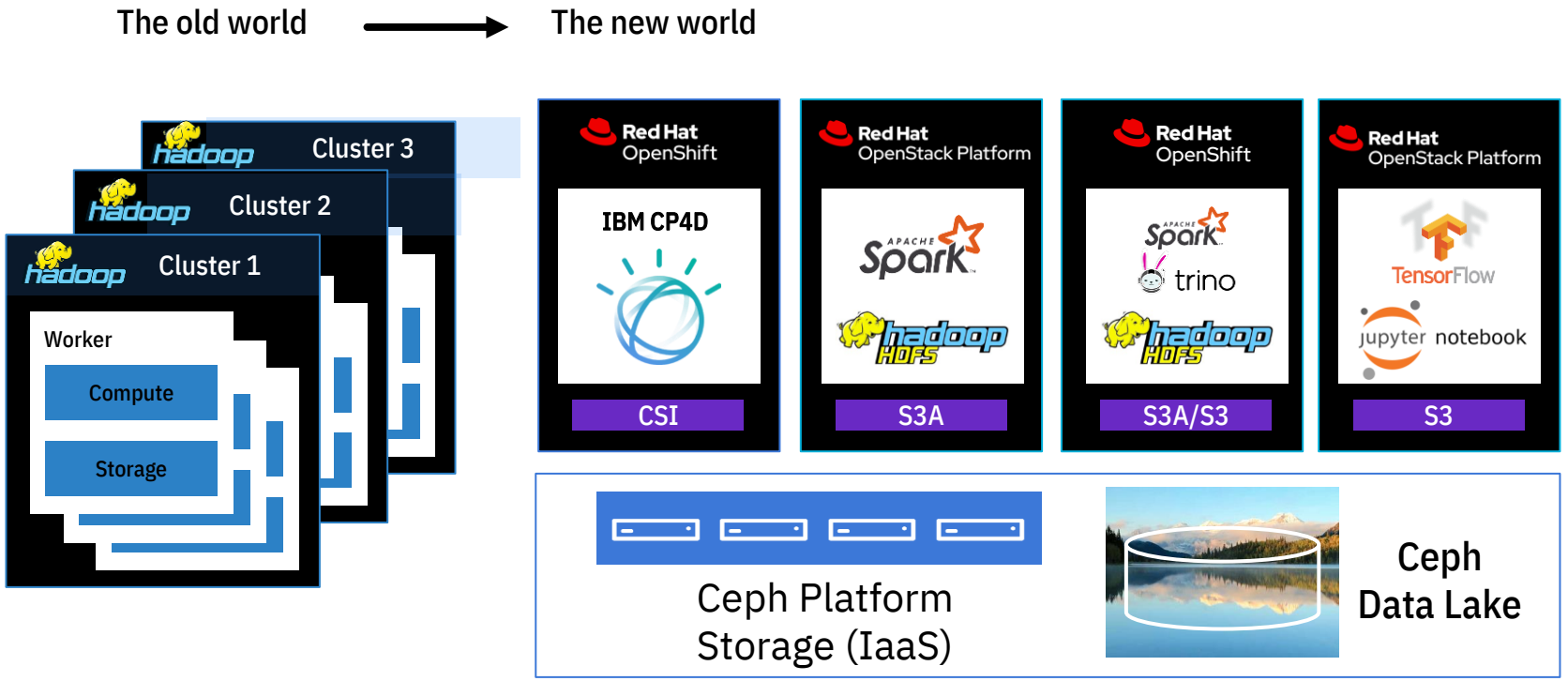
Open Source



© Copyright IBM Corporation 2023

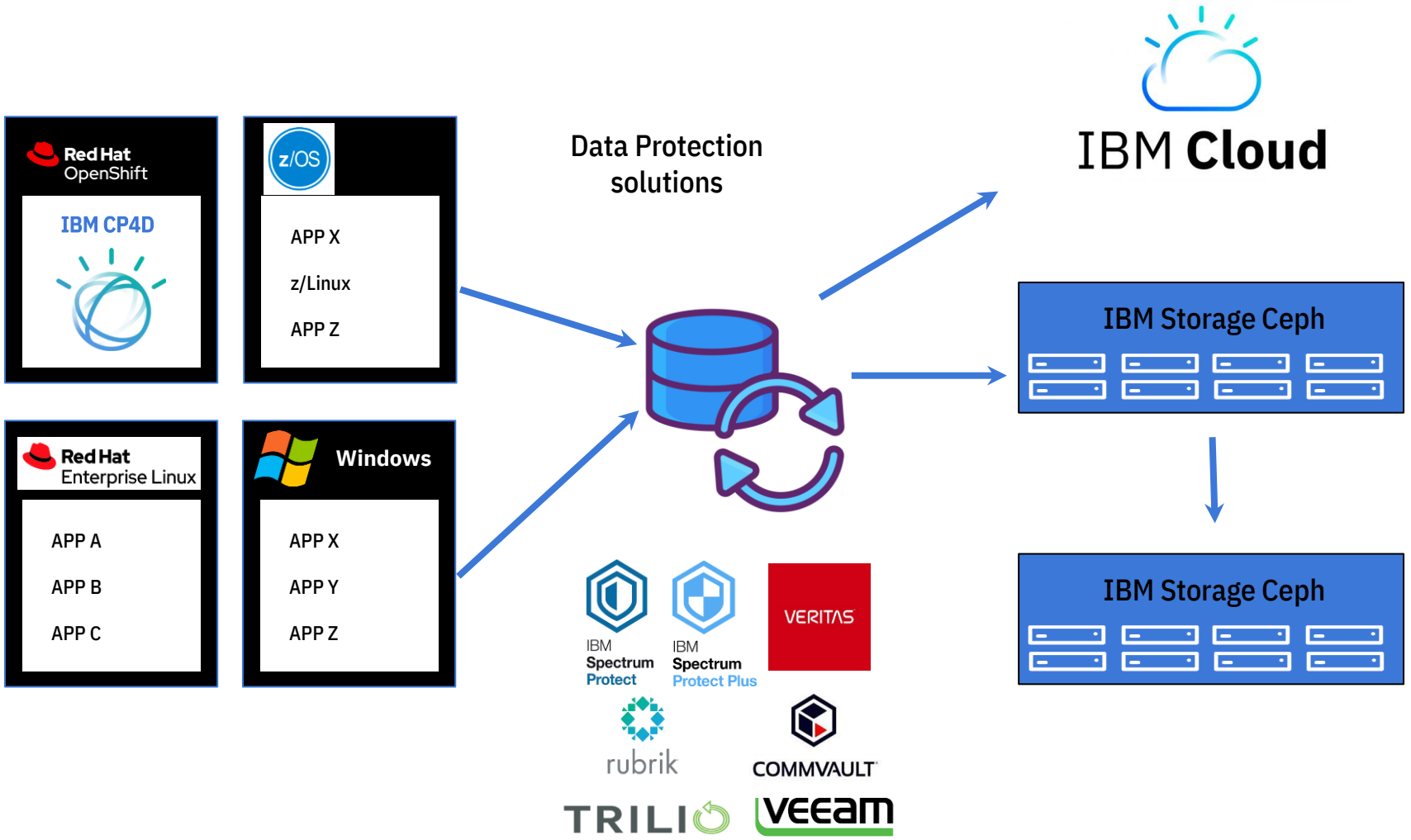
Object storage for live Data Analytics and AI/ML Infrastructure

Multi-tenant workload isolation with shared data context



© Copyright IBM Corporation 2023

IBM Storage Ceph as an operational backup target



© Copyright IBM Corporation 2023

IBM Storage Ceph Basics

© Copyright IBM Corporation 2023

IBM Storage Ceph high level overview

MGR / MON

- Red Hat Enterprise Linux
- RHEL native containers
- Monitoring
- Cluster health
- Ceph Dashboard
- Alerts, Notifications, Reports

OSDs

- Red Hat Enterprise Linux
- RHEL native containers
- Object Storage Device
- Storage Daemon
- Usually 1:1 OSD:Disk
- Scalable to 1000’s

RBD / MDS / RGW

- Red Hat Enterprise Linux
- RHEL native containers
- Block device (RBD)
- CephFS (MDS)
- S3 API (RGW)
- Cluster host/node
- Dedicated or shared nodes

Red Hat Enterprise Linux | **IBM Storage Fusion**

Any 3rd party x86 Appliance | **Hardware Resource Guidelines**

Hewlett Packard Enterprise, CISCO, SUPERMICRO, IBM, Lenovo, DELL

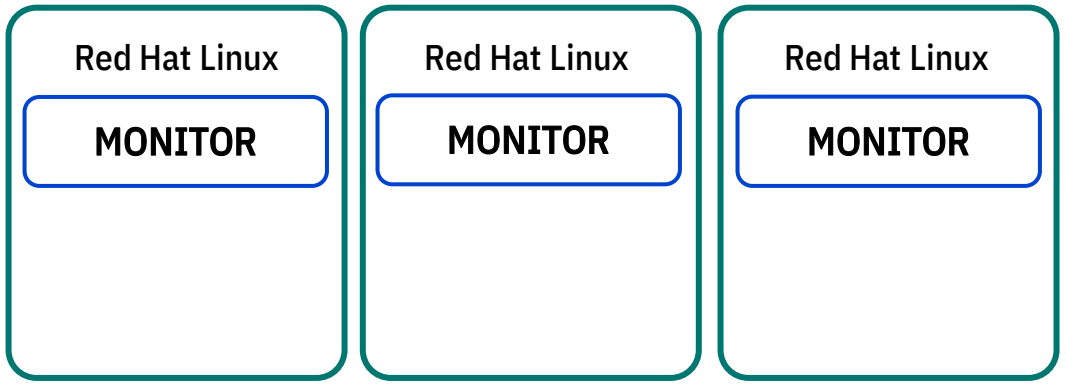
© Copyright IBM Corporation 2023

IBM Storage Ceph basics

MONITOR PROCESS

Ceph Monitors . . .

- Maintain the Ceph cluster map
- Make decisions based on consensus
- Operate in a small and odd number of instances (e.g. 3 or 5 nodes)
- Run as a containerized process



X86 Servers
Red Hat Enterprise Linux

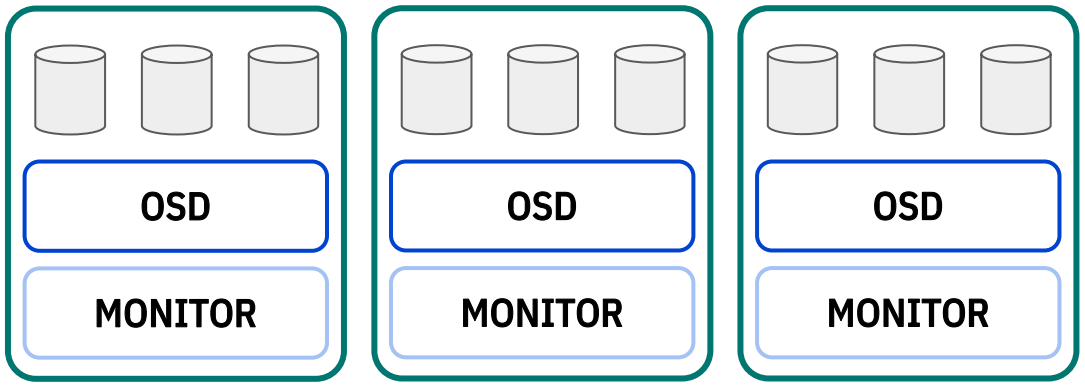
© Copyright IBM Corporation 2023

IBM Storage Ceph basics

OSD PROCESS

Ceph Object Storage Daemons (OSD) . . .

- Provide direct access to physical storage devices (i.e. HDD, SSD, etc.)
- Manage the layout of data on media
- Aggregated into pools
- Coordinates data protection, distribution, integrity checking, and data recovery peer to peer



X86 Servers with built-in storage
Red Hat Enterprise Linux

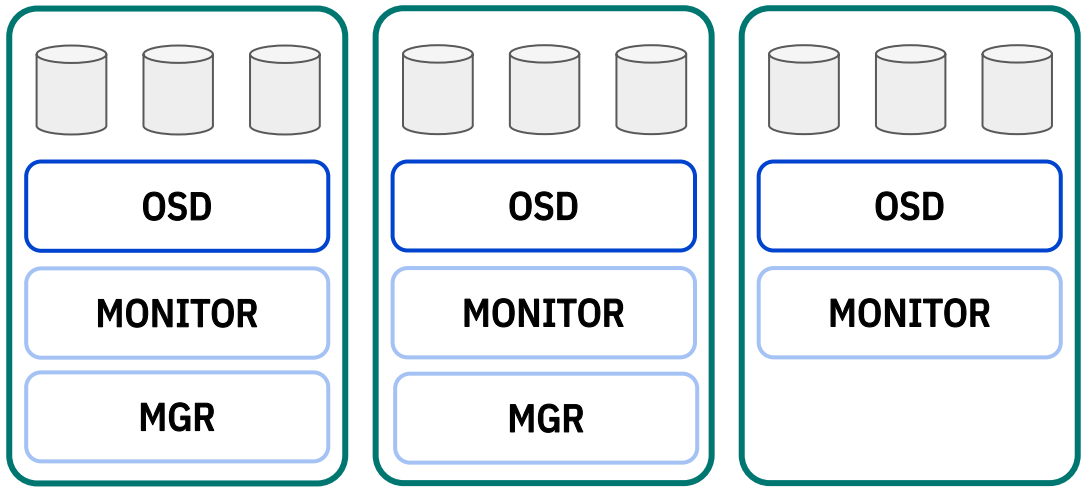
© Copyright IBM Corporation 2023

IBM Storage Ceph basics

MONITOR PROCESS | OSD PROCESS | MANAGER PROCESS

A basic Ceph cluster is composed of Monitors and OSD daemons

- Easy command-line interface (Cephadm CLI) and user interface (Ceph Dashboard UI) setup
- A minimal instance contains 3 nodes
  - IBM Recommends 4 nodes
- OSDs can scale to 10,000s in a cluster
- Can tune for performance, capacity, or cost



X86 Servers with built-in storage
Red Hat Enterprise Linux

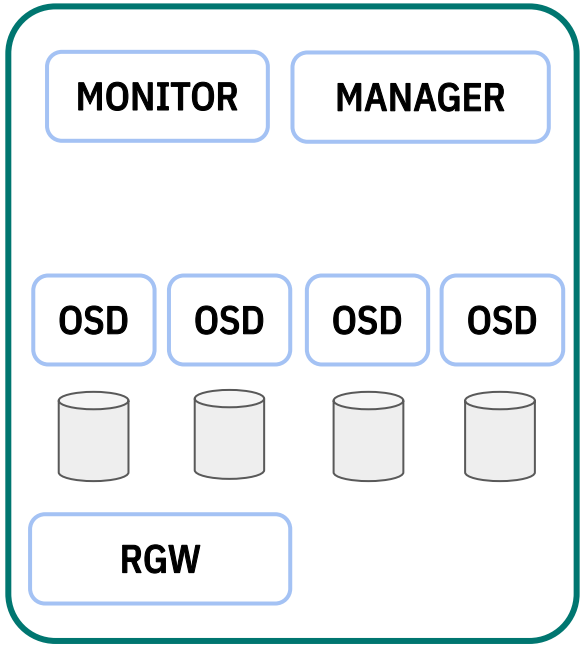
© Copyright IBM Corporation 2023

IBM Storage Ceph basics

MONITOR PROCESS | OSD PROCESS | MANAGER PROCESS

A personal workstation and virtualization software can host a single node Ceph sandbox

- Hardware resources (i.e. CPU, RAM, Disk)
- Prefers x86 architecture
- Virtualization software
- Open source Ceph or IBM Storage Ceph
- Community support



VirtualBox or VMware

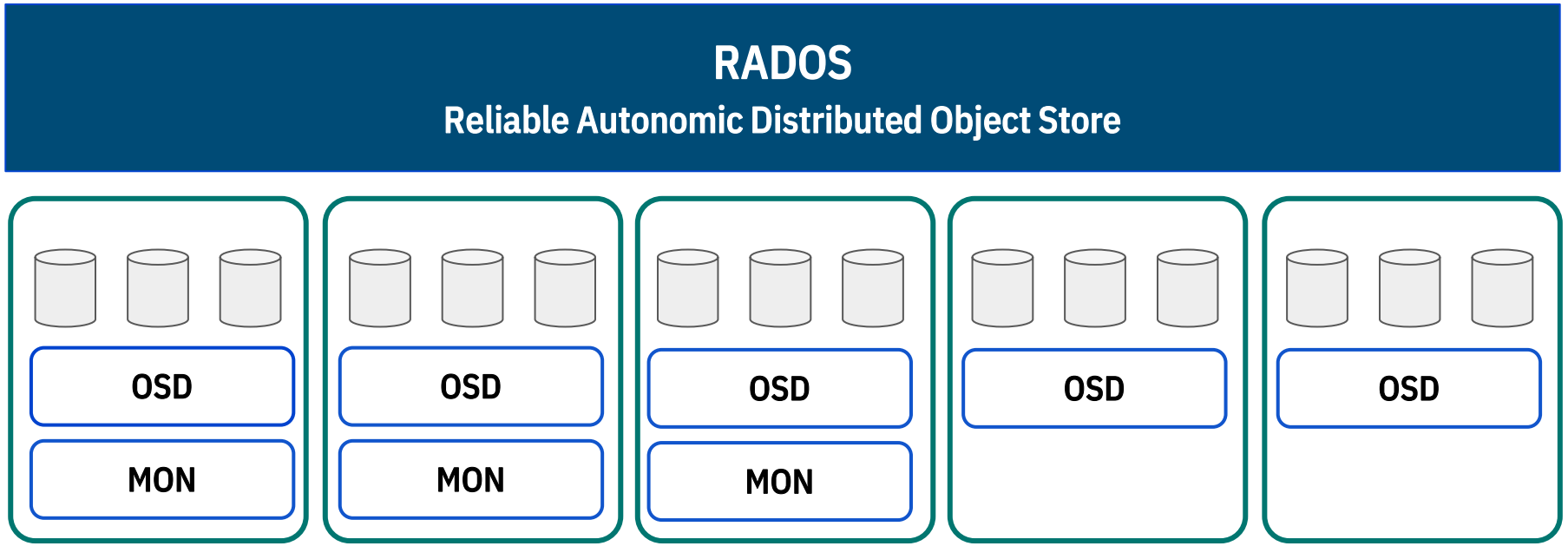
© Copyright IBM Corporation 2023

IBM Storage Ceph Data Access

© Copyright IBM Corporation 2023

IBM Storage Ceph data access

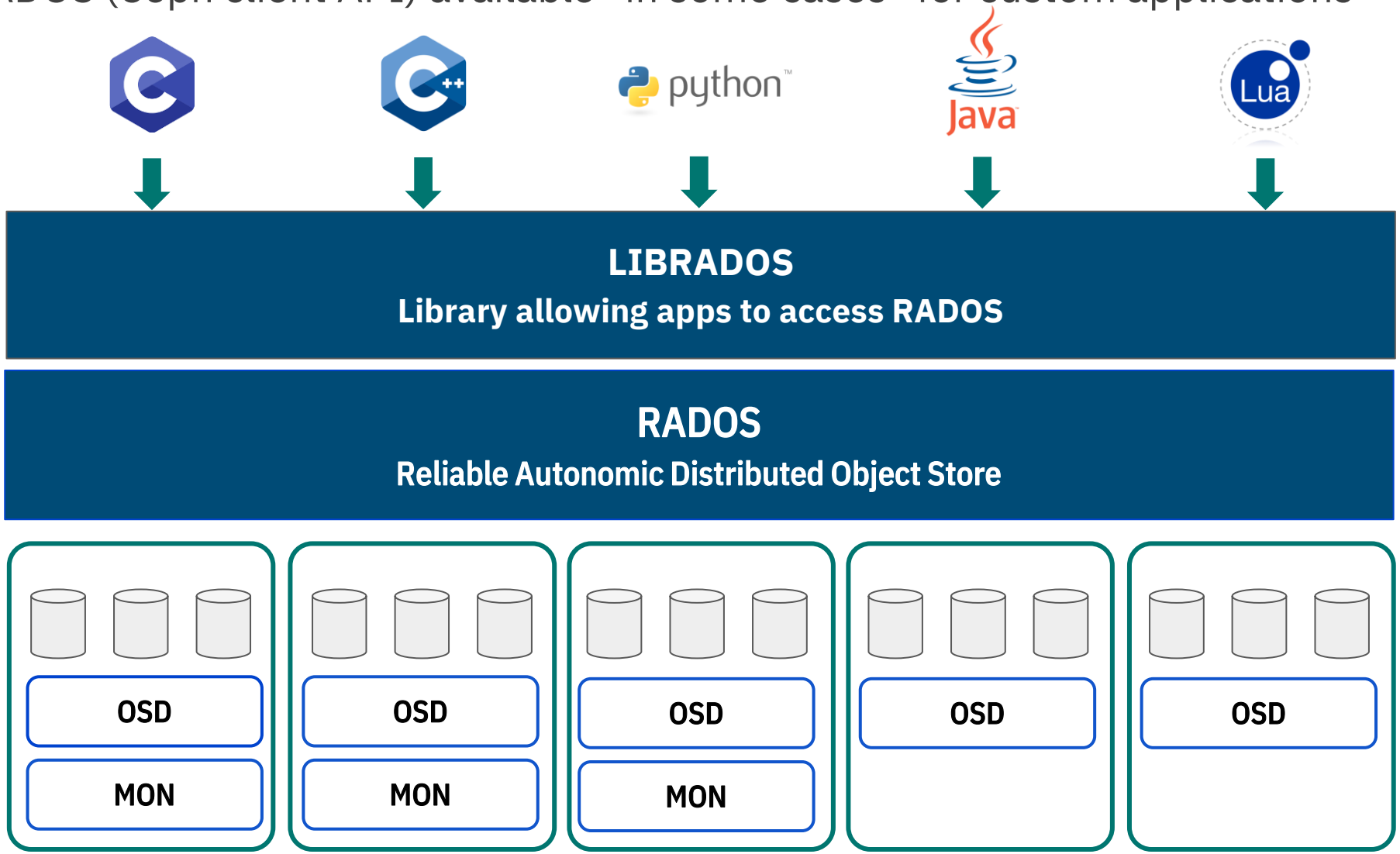
Ceph RADOS (Reliable Autonomic Distributed Object Storage)
Provides a data access abstraction layer



© Copyright IBM Corporation 2023

IBM Storage Ceph data access

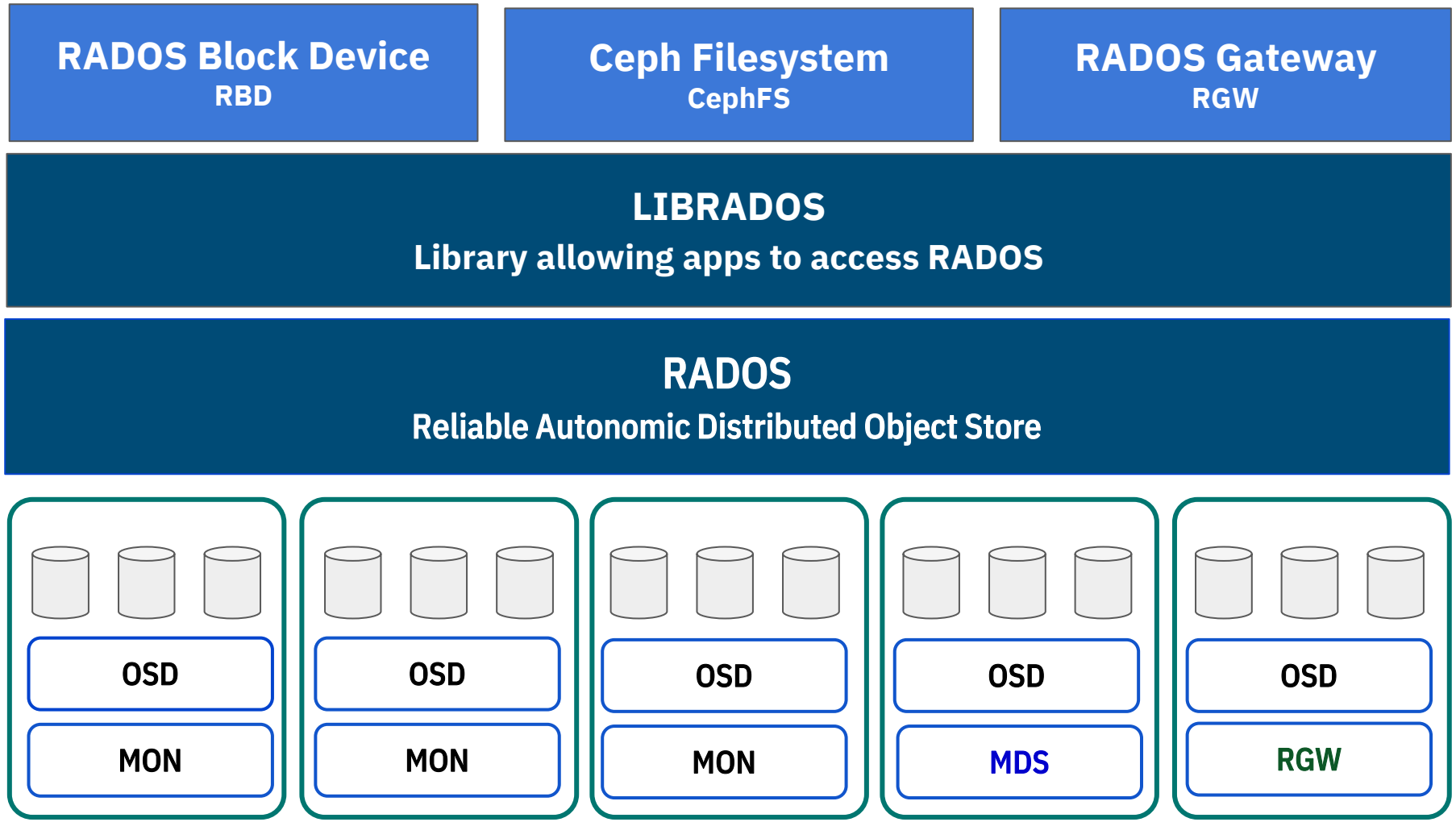
Ceph LIBRADOS (Ceph client API) available “in some cases” for custom applications



© Copyright IBM Corporation 2023

IBM Storage Ceph data access

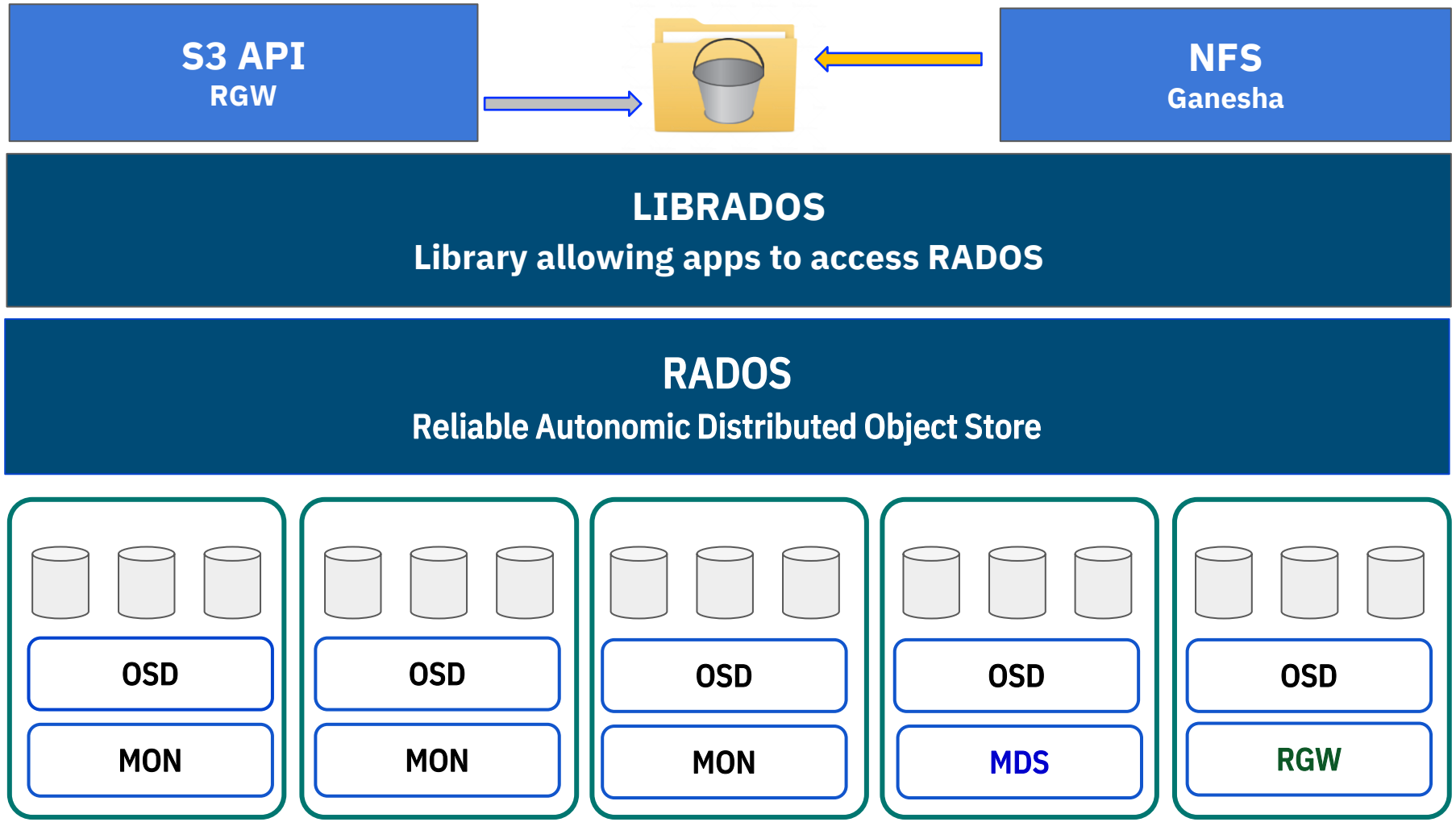
Ceph Data Services (RBD, CephFS, RGW) for mainstream application use cases



© Copyright IBM Corporation 2023

IBM Storage Ceph data access

Multi protocol S3 and NFS access for data migration



© Copyright IBM Corporation 2023

IBM Storage Ceph Live Demo

© Copyright IBM Corporation 2023

“cephadm” – The Command Line Interface

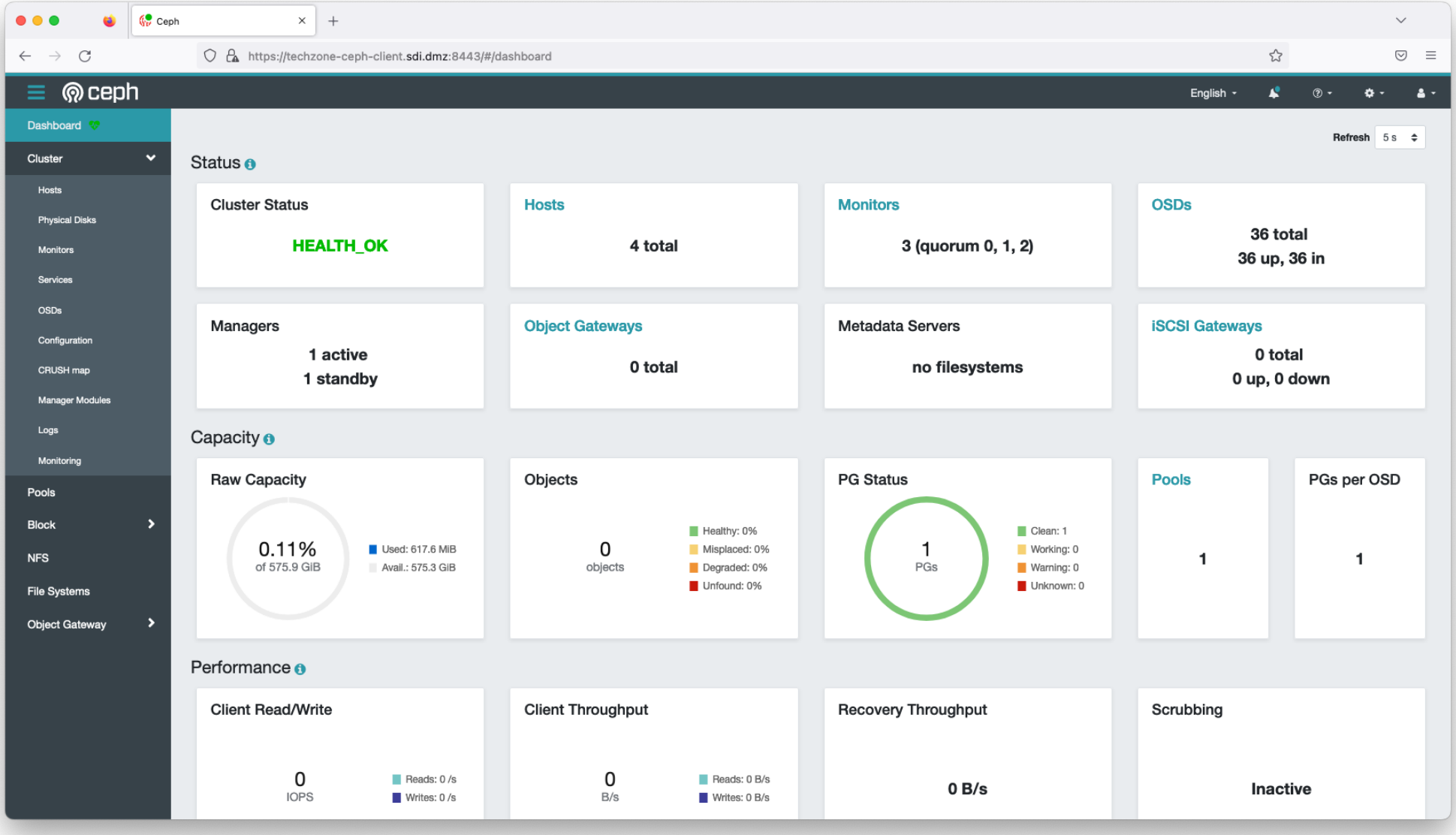
IBM Ceph TechZone Demo - November 2022.sh — UNREGISTERED

```
#
# Create a zone group
# Create the zone
# Commit the changes.
#
[ceph: root@ceph-mon01 /]# radosgw-admin zonegroup create --rgw-zonegroup=default  --master --default
[ceph: root@ceph-mon01 /]# radosgw-admin zone create --rgw-zonegroup=default --rgw-zone=test_zone --master --default
[ceph: root@ceph-mon01 /]# radosgw-admin period update --rgw-realm=test_realm --commit

#
# Deploy the RGW daemons now
#
[ceph: root@ceph-mon01 /]# ceph orch apply rgw test --realm=test_realm --zone=test_zone --placement="2 proxy01.example.com ceph-mon03.example.com"

#
# Verification tasks
#
[ceph: root@ceph-mon01 /]# ceph -s
[ceph: root@ceph-mon01 /]# ceph orch ls
[ceph: root@ceph-mon01 /]# ceph orch ps

#
# Next command on the Proxy01
#
[root@proxy01 ~]# netstat -tulpn | grep radosgw
# tcp        0      0 0.0.0.0:80              0.0.0.0:*               LISTEN      6970/radosgw
# tcp6       0      0 :::80                   :::*                    LISTEN      6970/radosgw

#
# Back on the Ceph Monitor
#
[root@ceph-mon01 ~]# curl http://192.168.56.24:80
# <?xml version="1.0" encoding="UTF-8"?><ListAllMyBucketsResult xmlns="http://s3.amazonaws.com/doc/2006-03-01/"><Owner><ID>anonymous</ID><DisplayName></DisplayName></Owner
[root@ceph-mon01 ~]# rados lspools
# device_health_metrics
# .rgw.root
# test_zone.rgw.log
# test_zone.rgw.control
# test_zone.rgw.meta

#
# Create a RADOS Gateway (RGW) user for S3 access
#
[ceph: root@ceph-mon01 /]# radosgw-admin user create --uid='user1' --display-name='First User' --access-key='S3user1' --secret-key='S3user1key'
[ceph: root@ceph-mon01 /]# radosgw-admin user info --uid='user1'

#
# Set up the Ceph RGW front end dashboard Front End now
#
[ceph: root@ceph-mon01 /]# radosgw-admin user create --uid=rgw-admin --display-name=rgw-admin --system

            "user": "rgw-admin",
            "access_key": "42VCS0WDXDOP85NBTU60",
            "secret_key": "mtYiM1TNEzkbUdJl8WgsYPJhMecm3rW5dzJSPnD1"

#
# If you lose the access key, use the next command to list it
#
```

keyring — Find | Find Prev | Find All — 6 matches — Spaces: 4 — Bash

© Copyright IBM Corporation 2023

Ceph Dashboard – The Browser Interface

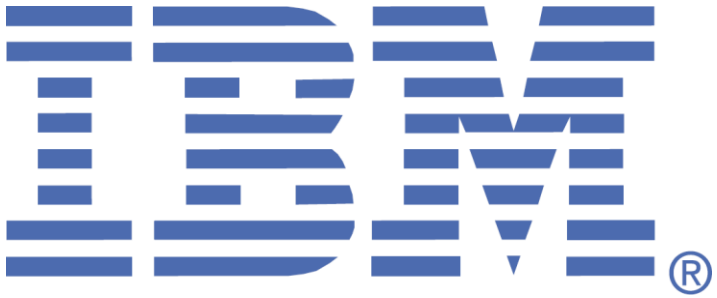


© Copyright IBM Corporation 2023

Cluster guidelines for starter configurations

Cluster Type	Minimum Recommended Nodes
Minimal Ceph cluster	Three (3) Node cluster Not recommended for production deployments
Starter IBM Storage Ceph cluster	Four (4) OSD Nodes per cluster Supports Replica 3x data protection plus node redundancy
Balanced cost/capacity cluster	Seven (7) OSD Nodes per cluster Supports EC 4+2 data protection plus node redundancy
Performance cluster	10 OSD Nodes per cluster Supports multiple Replica 3x data protection plus node redundancy
Capacity cluster	12 OSD Nodes per cluster Supports EC 8+3 data protection plus node redundancy

© Copyright IBM Corporation 2023



© Copyright IBM Corporation 2023