

Advanced Technology Group



Accelerate with ATG

Introduction to IBM Storage Ceph

John Shubeck – ATG Storage Technical Specialist

Date: January 10, 2023



Storage Ceph



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.



Important Links to bookmark:

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

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

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](https://www.menti.com) with code 2243 3599 or

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

Or

QR Code



Advanced Technology Group



Accelerate with ATG

Introduction to IBM Storage Ceph

John Shubeck – ATG Storage Technical Specialist

Date: January 10, 2023



Storage Ceph



About the Presenter



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.

Introducing our Panelists



JC Lopez is a Senior IT Specialist and ATG presenter, panelist, and developer of content, plus a member of the Spectrum Storage Technical Leadership Team



Norman Bogard is a Senior IT Specialist and ATG presenter, panelist, and developer of content, plus a member of the Spectrum Storage Technical Leadership Team

Introducing our panelists



Lloyd Dean is an IBM Principal Storage Technical Specialist in IBM Storage Solutions. Lloyd has held numerous senior technical roles at IBM during his 22 plus years at IBM. Lloyd most recently is leading efforts in the Advanced Technology Group as the IBM Storage for Red Hat OpenShift focal and as a Hybrid Cloud storage solution SME covering IBM Block, File and Object storage solutions and their use cases supporting IBM Cloud Paks.



Uday Boppana is the product management lead for Ceph at IBM. He has experience working in hybrid cloud, data center and storage solutions in different roles and with a variety of technologies. In prior roles, he worked in product management, technical marketing, solutions architecture, and in leadership and technical positions in engineering. He is a regular speaker at industry conferences and events related to AI/ML and hybrid cloud data services and solutions.

Introducing our panelists



Kyle Bader is a Senior Principal Architect, contributing expertise at the intersection of open source, networking, distributed storage, big data, and machine learning. He is currently engaged on cross functional Ceph projects spanning both the Red Hat and IBM organizations.

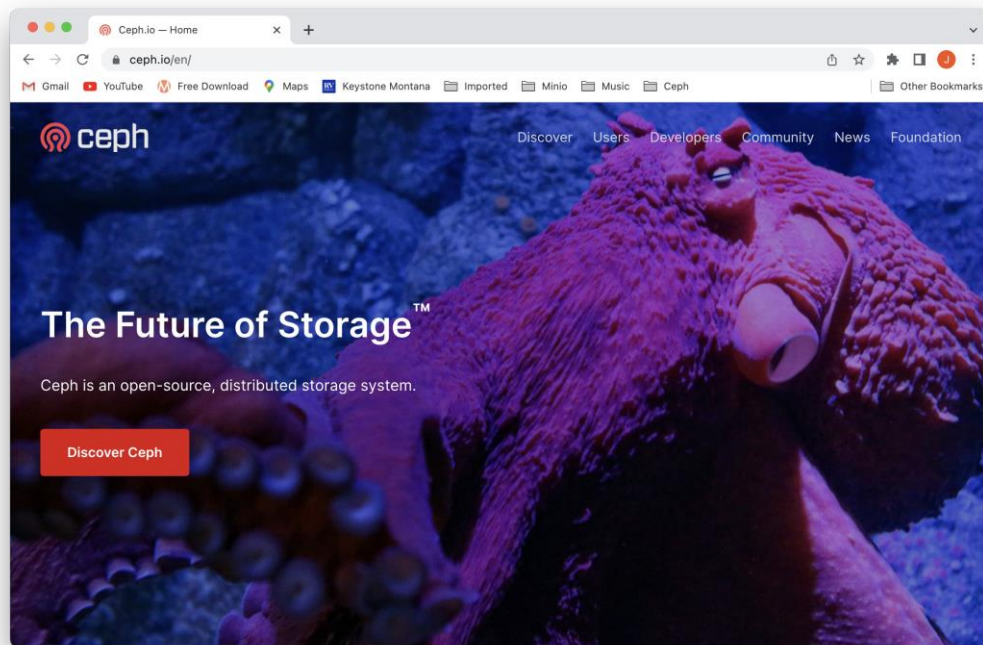
Agenda



- What is Ceph
- The history of Ceph
- What is IBM Storage Ceph
- IBM contributions to the Ceph community
- IBM Storage Ceph Basics
- IBM Storage Ceph Deployment
- IBM Technology Zone Ceph Test Drive
- The flexibility of IBM Storage Ceph
- Why IBM Storage Ceph

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

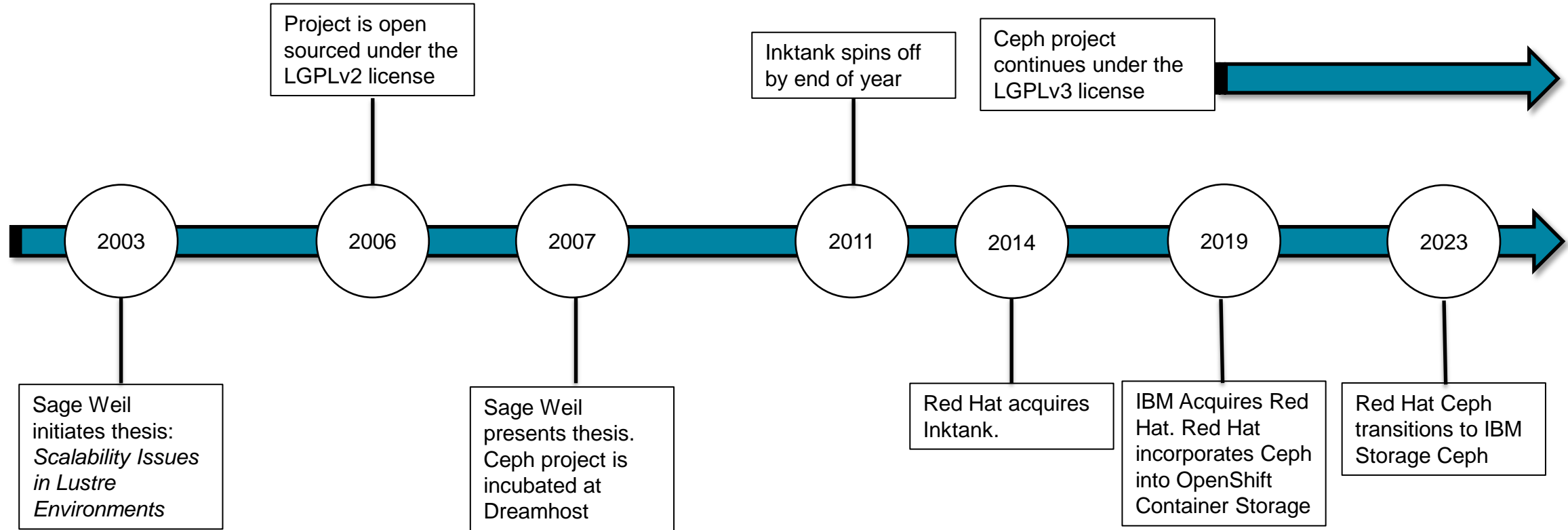
Flexible, scale-out architecture on clustered standard hardware

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

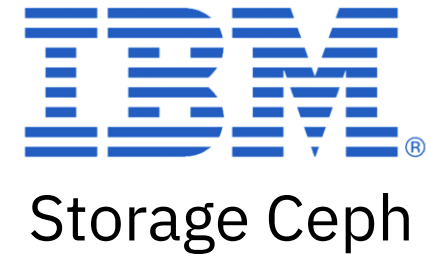
The History of Ceph



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 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



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

Ceph as an Open Source project

What is the meaning of the term Open Source?

- The term open source refers to something people can modify because its design is publicly accessible.
- The term originated in the context of software development to designate a specific approach to creating computer programs.
- Today, however, “open source” designates a broader set of values that we call “the open source way.”

Source: (opensource.com)

Ceph as an Open Source Project

The screenshot shows the GitHub repository page for `ceph/ceph`. The main content area displays the README for the project, which describes Ceph as a scalable distributed storage system. The README includes sections for contributing code, checking out the source, and licensing information. On the right side, there is a sidebar with repository statistics, including 649 watches, 5.2k forks, and 11.4k stars. Below the statistics, there are sections for releases (373 tags), packages (no packages published), and users (42 users).

Repository Statistics:

- Watch: 649
- Fork: 5.2k
- Star: 11.4k
- Commits: 642
- Tags: 373
- Users: 42

README Content:

Ceph - a scalable distributed storage system

Please see <https://ceph.com/> for current info.

Contributing Code

Most of Ceph is dual licensed under the LGPL version 2.1 or 3.0. Some miscellaneous code is under a BSD-style license or is public domain. The documentation is licensed under Creative Commons Attribution Share Alike 3.0 (CC-BY-SA-3.0). There are a handful of headers included here that are licensed under the GPL. Please see the file COPYING for a full inventory of licenses by file.

Code contributions must include a valid "Signed-off-by" acknowledging the license for the modified or contributed file. Please see the file SubmittingPatches.rst for details on what that means and on how to generate and submit patches.

We do not require assignment of copyright to contribute code; code is contributed under the terms of the applicable license.

Checking out the source

You can clone from github with

```
git clone git@github.com:ceph/ceph
```

Footer: Merge pull request #48843 from rhcs-dashboard/expose_slow_ops

IBM as a Ceph Project Contributor

Principles of the Open Source way

- Transparency
- Collaboration
- Release early and often
- Inclusive meritocracy
- Community



[The Open Source
Guidebook 2.0](#)

Source: (opensource.com)

IBM leaders, innovators and talented intellects behind Ceph

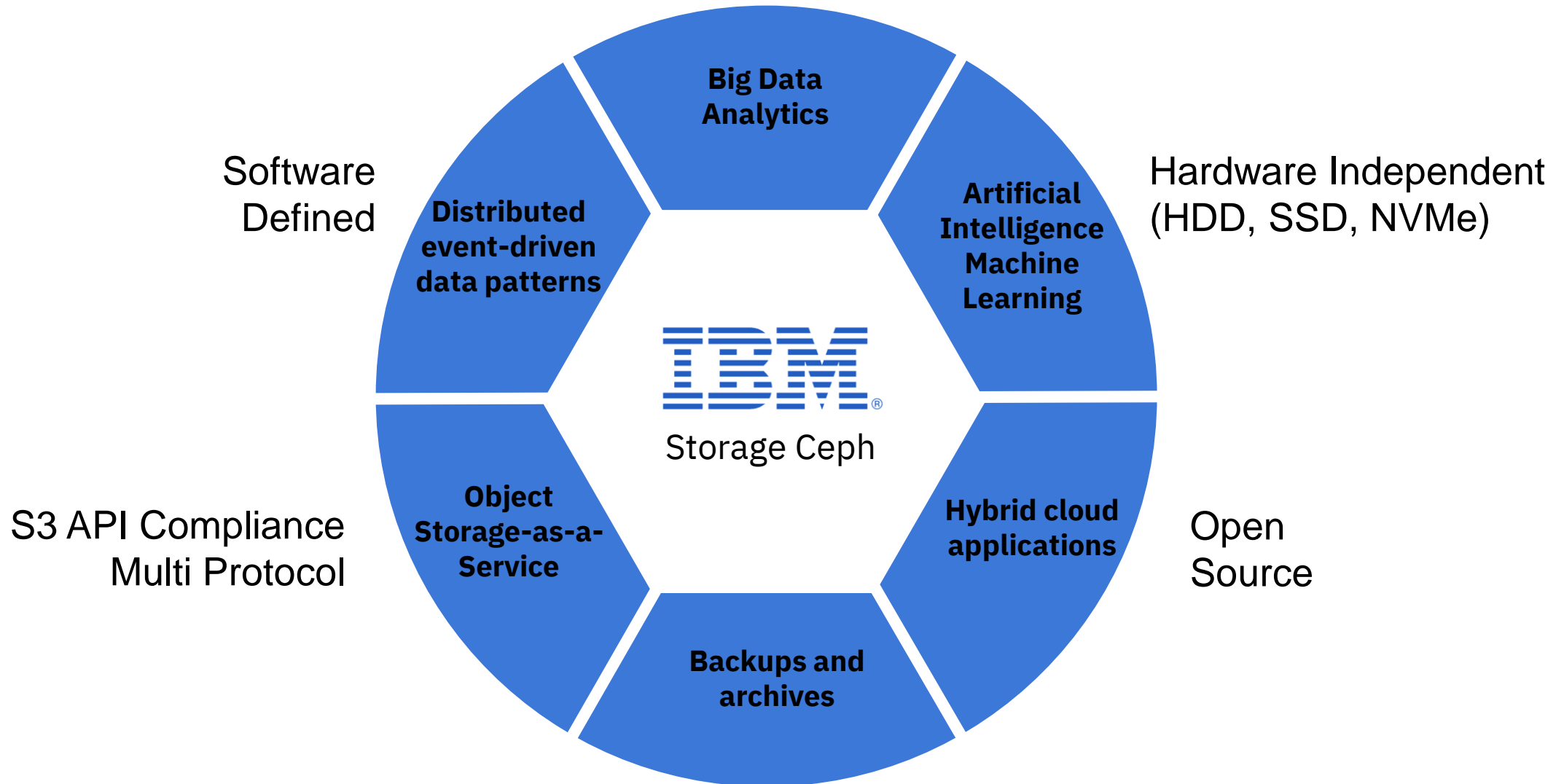
The Ceph Foundation Leadership Team is comprised of the key technical players who manage the community and oversee the advancement of Ceph.

- Core Team
- Maintainers
- Component Leads
- Members



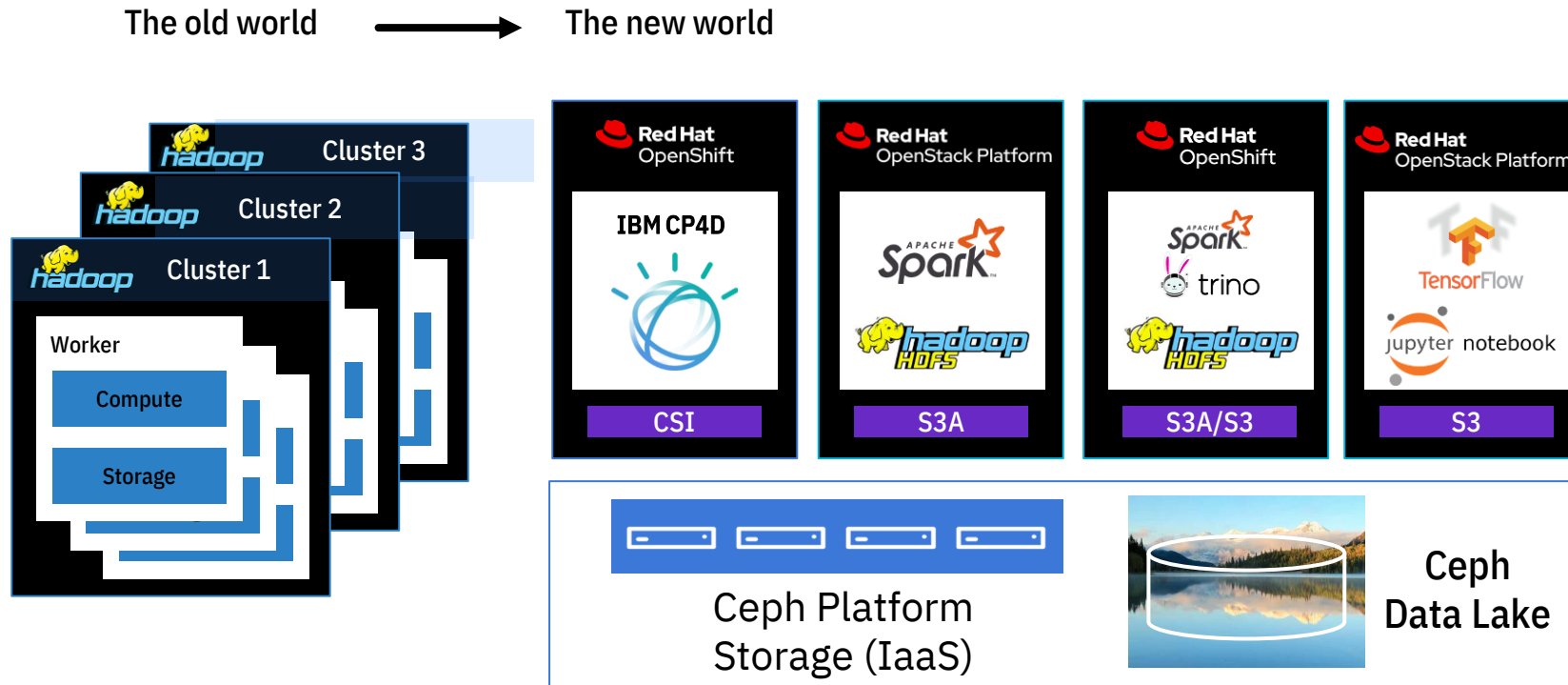
Source: Ceph Foundation Team (<https://ceph.io/en/community/team/>)

IBM Storage Ceph Use Cases

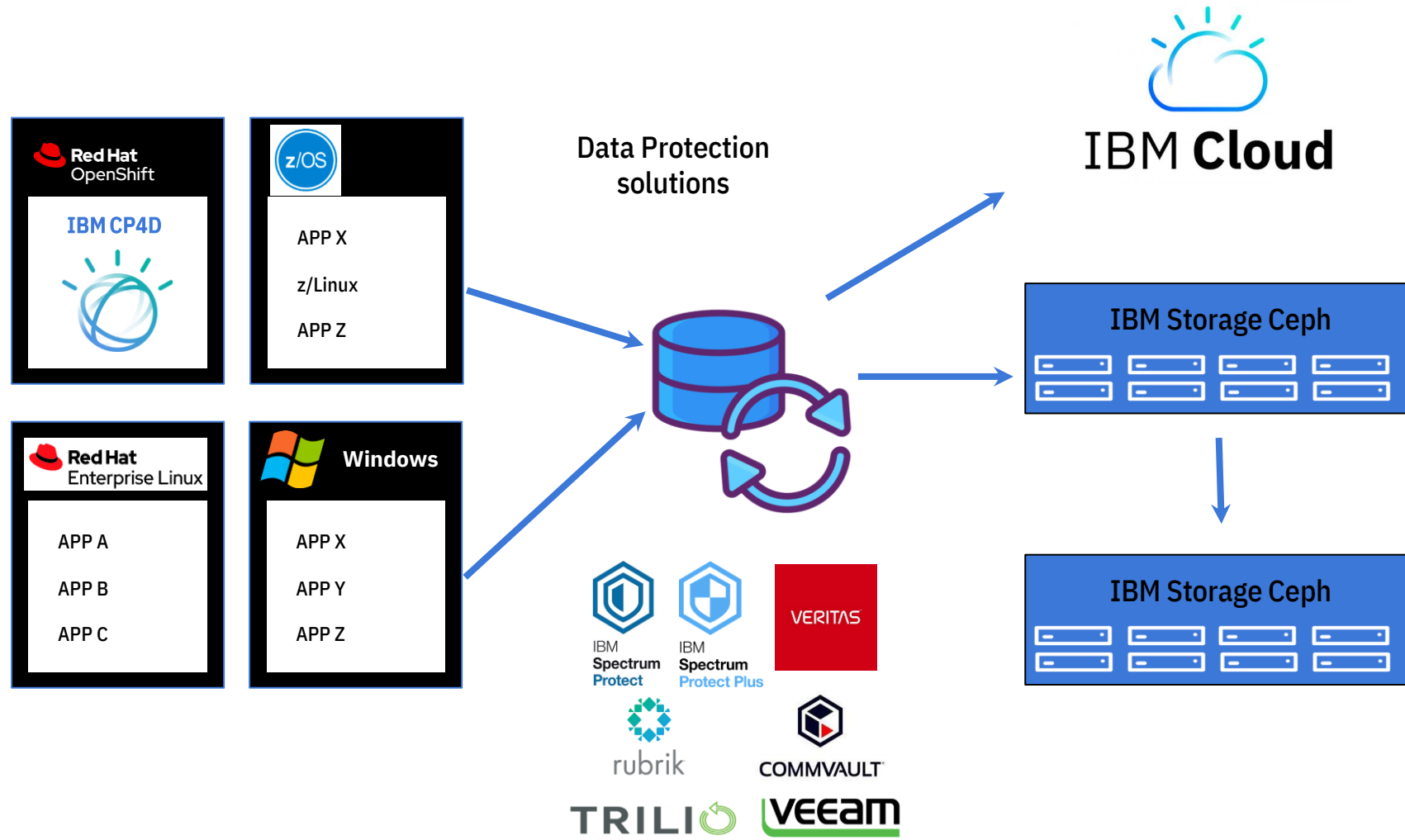


Object storage for live Data Analytics and AI/ML Infrastructure

Multi-tenant workload isolation with shared data context



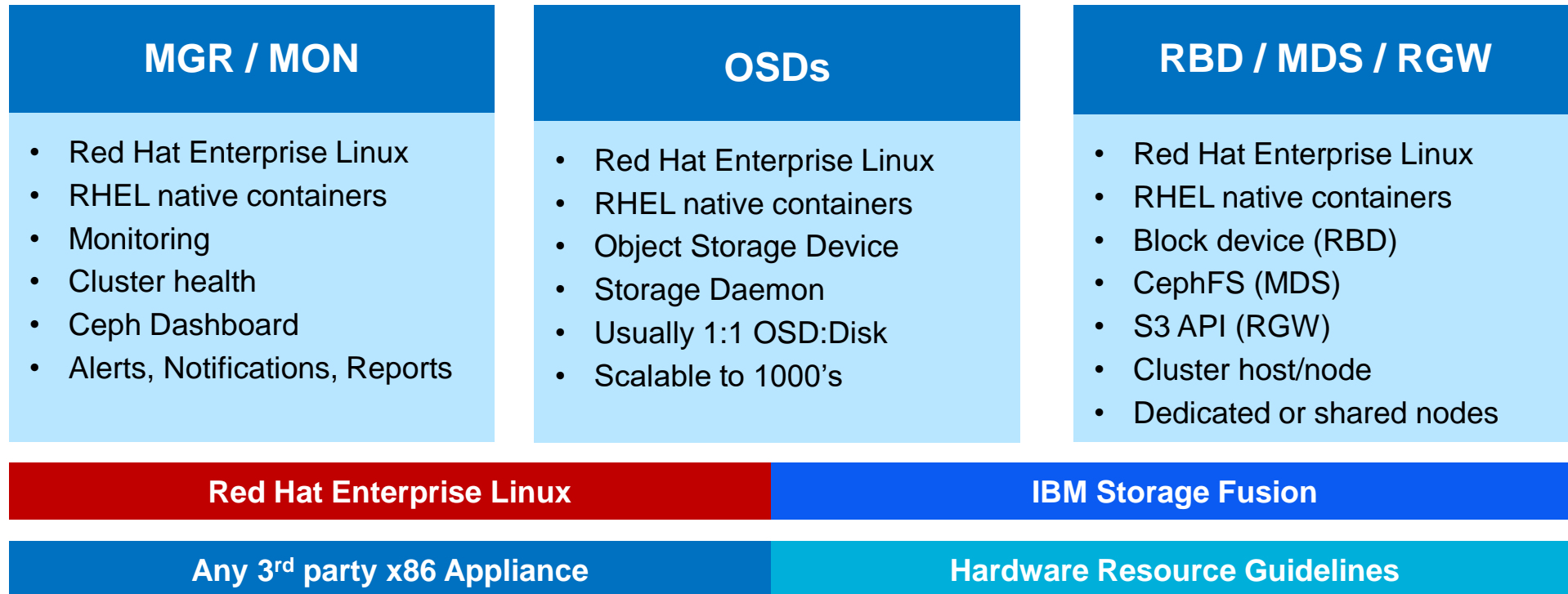
IBM Storage Ceph as an operational backup target



IBM Storage Ceph Basics



IBM Storage Ceph high level overview

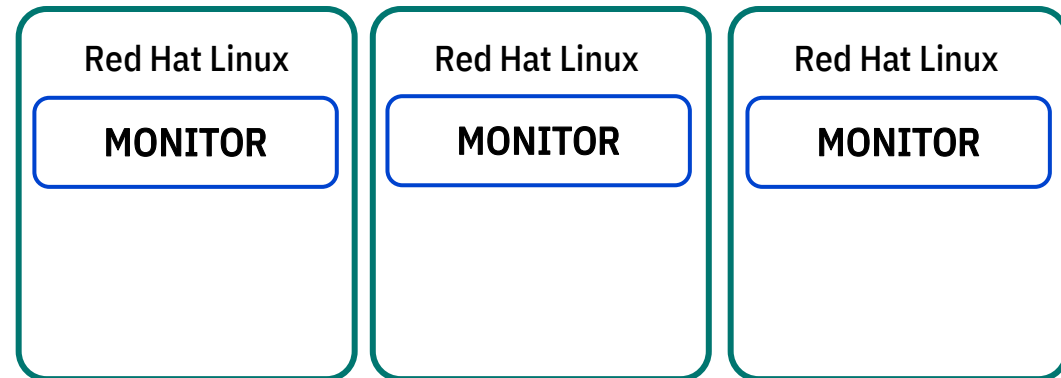


IBM Storage Ceph basics

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



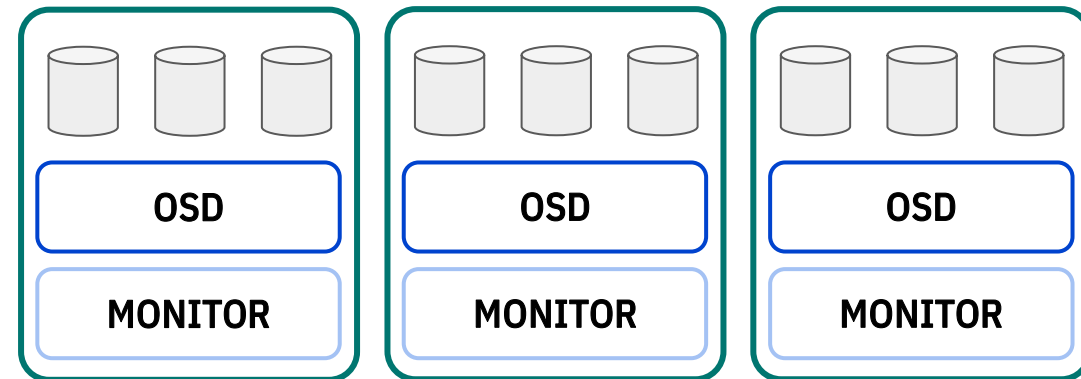
X86 Servers
Red Hat Enterprise Linux

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

IBM Storage Ceph basics

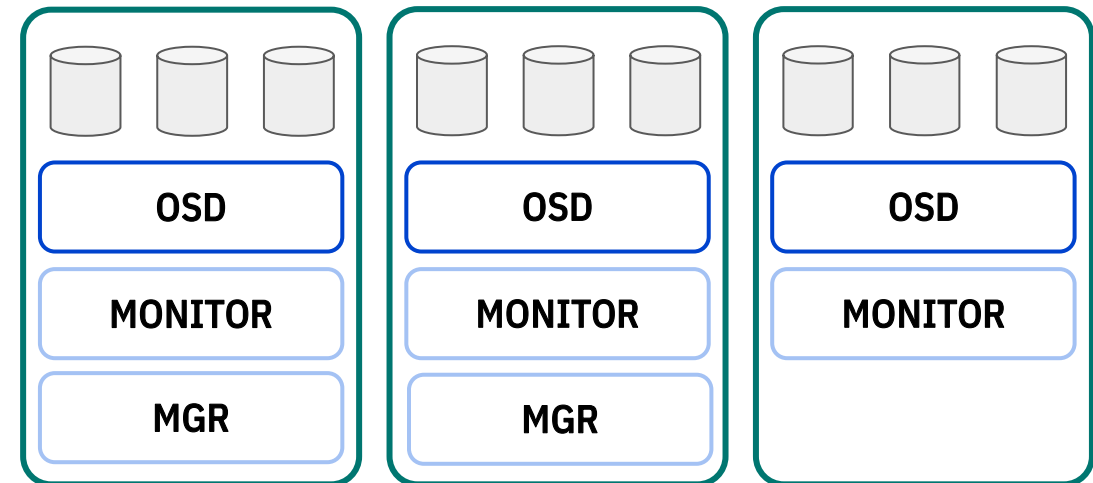
MONITOR PROCESS

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

IBM Storage Ceph basics

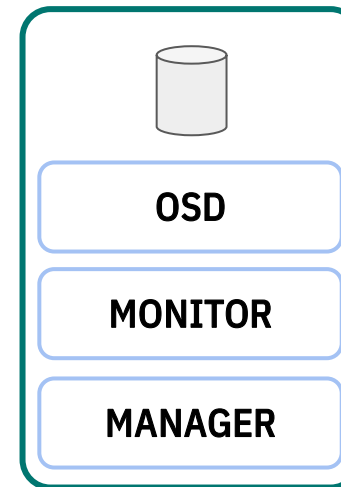
MONITOR PROCESS

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

“cephadm” – The Command Line Interface



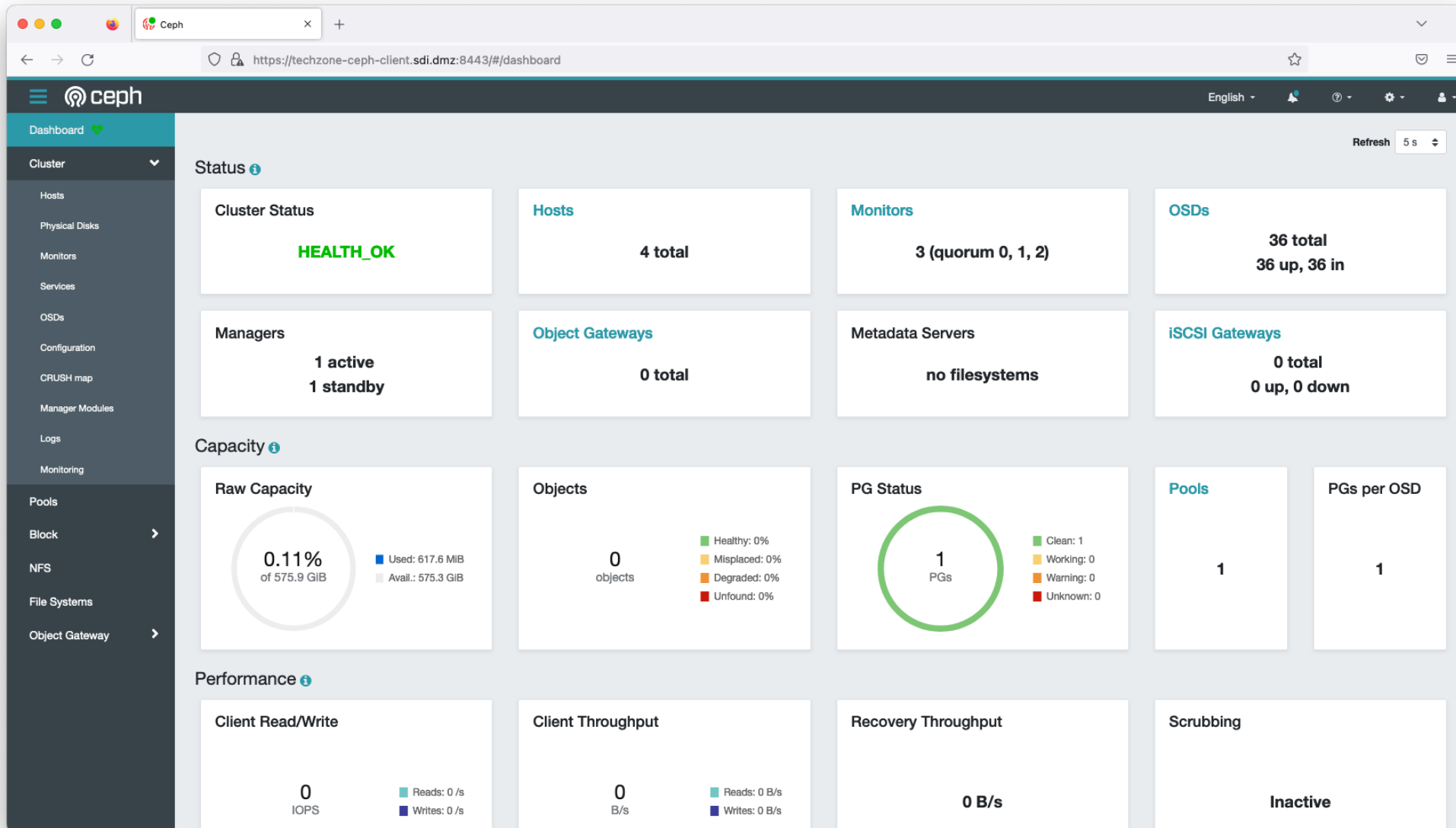
```

IBM Ceph TechZone Demo - November 2022.sh
UNREGISTERED

446 #
447 # Create a zone group
448 # Create the zone
449 # Commit the changes.
450 #
451 [ceph: root@ceph-mon01 /]# radosgw-admin zonegroup create --rgw-zonegroup=default --master --default
452 [ceph: root@ceph-mon01 /]# radosgw-admin zone create --rgw-zonegroup=default --rgw-zone=test_zone --master --default
453 [ceph: root@ceph-mon01 /]# radosgw-admin period update --rgw-realm=test_realm --commit
454 #
455 #
456 # Deploy the RGW daemons now
457 #
458 [ceph: root@ceph-mon01 /]# ceph orch apply rgw test --realm=test_realm --zone=test_zone --placement="2 proxy01.example.com ceph-mon03.example.com"
459 #
460 #
461 # Verification tasks
462 #
463 [ceph: root@ceph-mon01 /]# ceph -s
464 [ceph: root@ceph-mon01 /]# ceph orch ls
465 [ceph: root@ceph-mon01 /]# ceph orch ps
466 #
467 #
468 # Next command on the Proxy01
469 #
470 [root@proxy01 ~]# netstat -tuln | grep radosgw
471 # tcp        0      0 0.0.0.0:80          0.0.0.0:*          LISTEN        6970/radosgw
472 # tcp6      0      0 :::80              :::*                LISTEN        6970/radosgw
473 #
474 #
475 # Back on the Ceph Monitor
476 #
477 [root@ceph-mon01 ~]# curl http://192.168.56.24:80
478 # <?xml version="1.0" encoding="UTF-8"><ListAllMyBucketsResult xmlns="http://s3.amazonaws.com/doc/2006-03-01/"><Owner><ID>anonymous</ID><DisplayName></DisplayName></Owner>
479 [root@ceph-mon01 ~]# rados lspools
480 # device_health_metrics
481 # .rgw.root
482 # test_zone.rgw.log
483 # test_zone.rgw.control
484 # test_zone.rgw.meta
485 #
486 #
487 # Create a RADOS Gateway (RGW) user for S3 access
488 #
489 [ceph: root@ceph-mon01 /]# radosgw-admin user create --uid='user1' --display-name='First User' --access-key='S3user1' --secret-key='S3user1key'
490 [ceph: root@ceph-mon01 /]# radosgw-admin user info --uid='user1'
491 #
492 #
493 # Set up the Ceph RGW front end dashboard Front End now
494 #
495 [ceph: root@ceph-mon01 /]# radosgw-admin user create --uid=rgw-admin --display-name=rgw-admin --system
496 #
497 #
498 #
499 #
500 #
501 #
502 # If you lose the access key, use the next command to list it
503 #

```

Ceph Dashboard – The Browser Interface

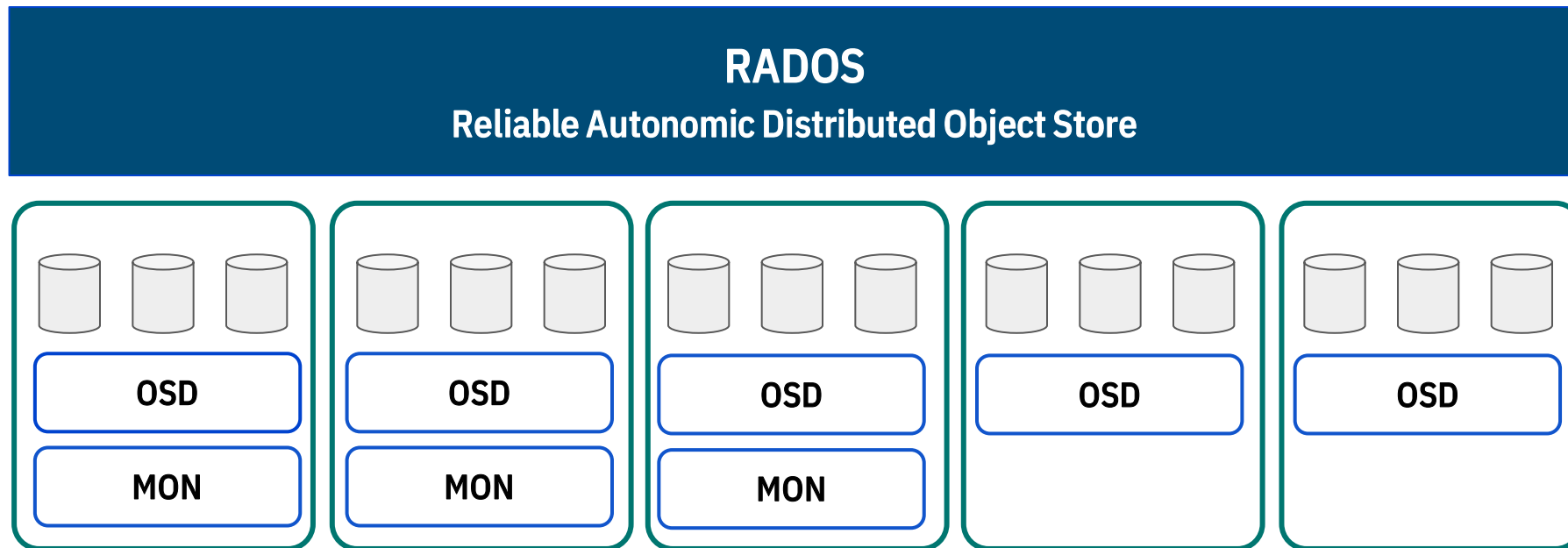


IBM Storage Ceph Data Access



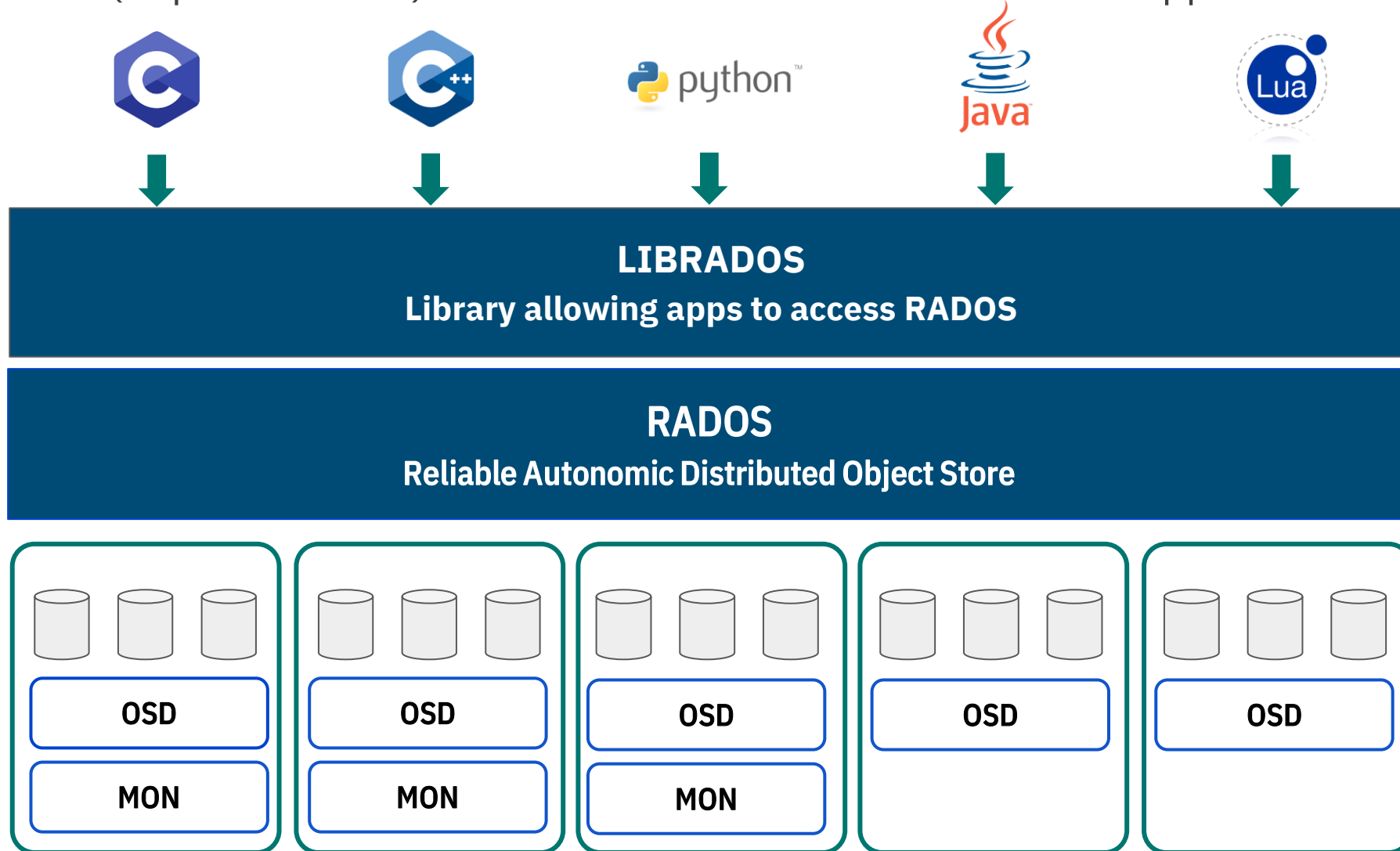
IBM Storage Ceph data access

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



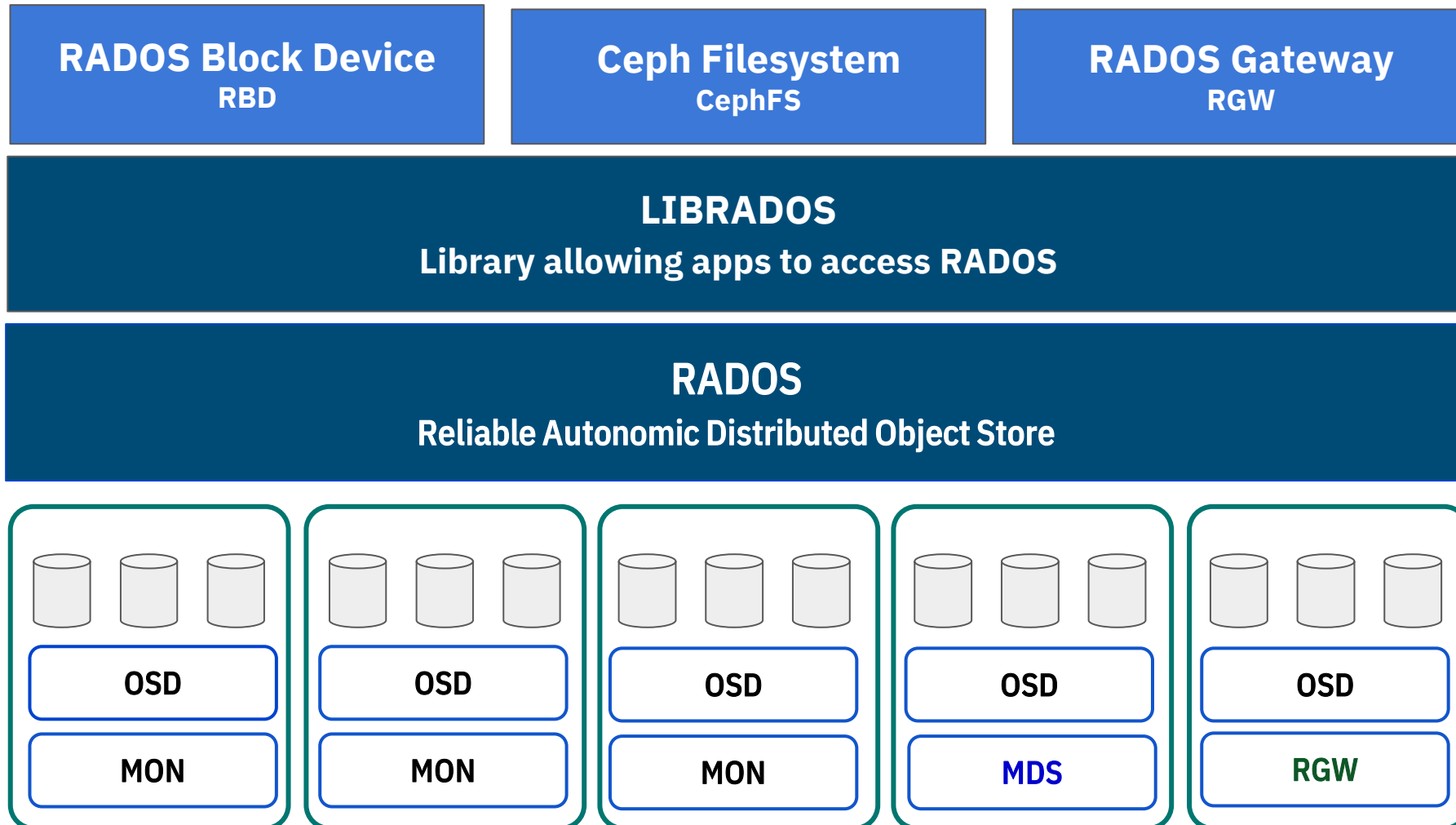
IBM Storage Ceph data access

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



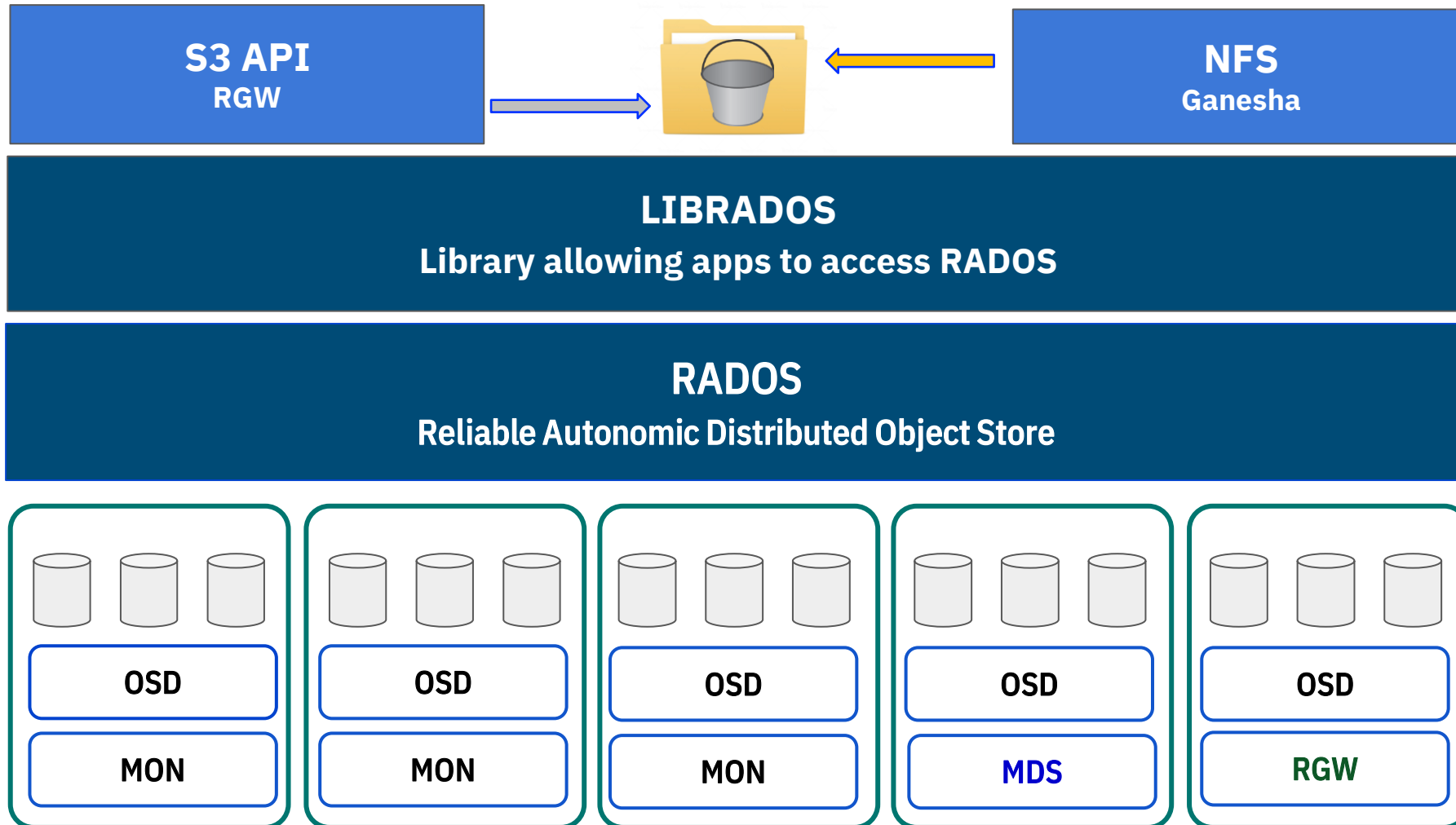
IBM Storage Ceph data access

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



IBM Storage Ceph data access

Multi protocol S3 and NFS access for data migration



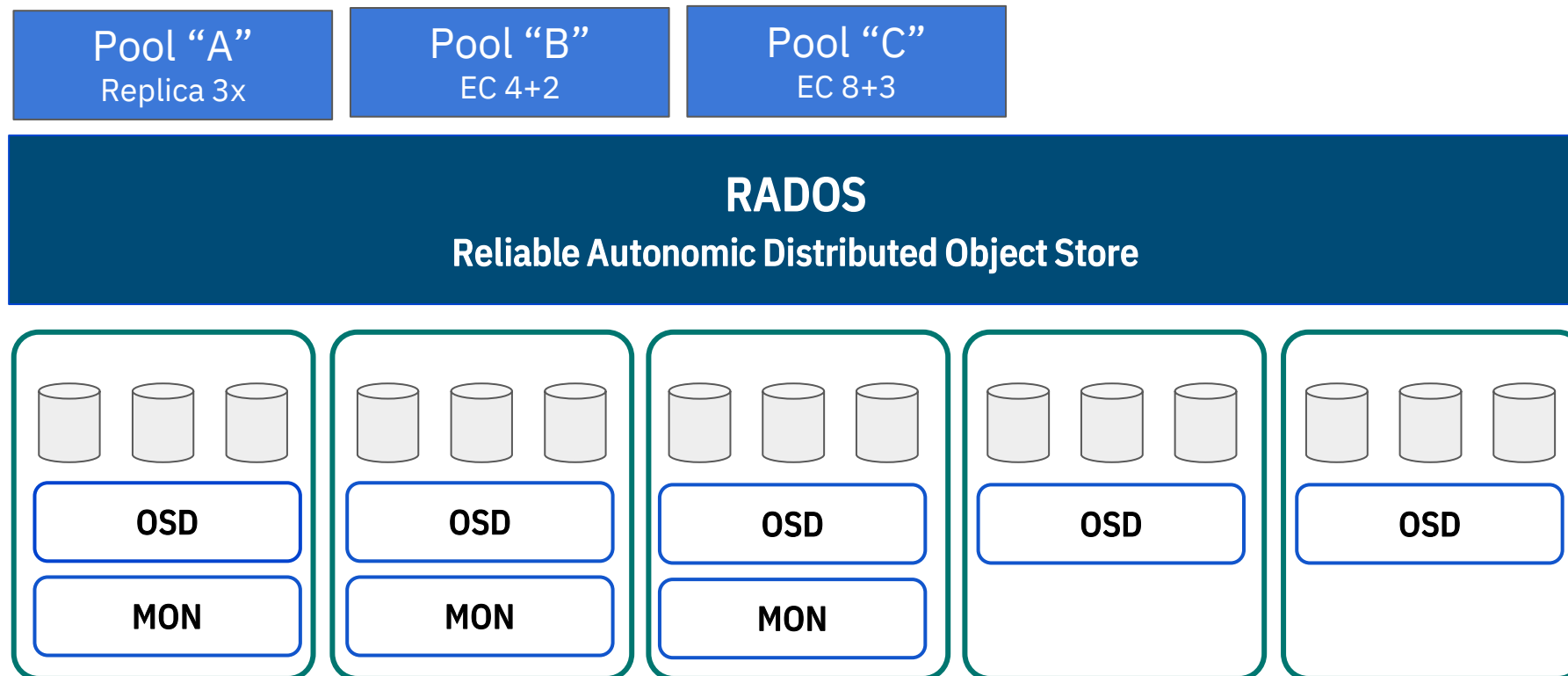
IBM Storage Ceph Data Protection



IBM Storage Ceph data protection

Ceph *storage* Pools . . .

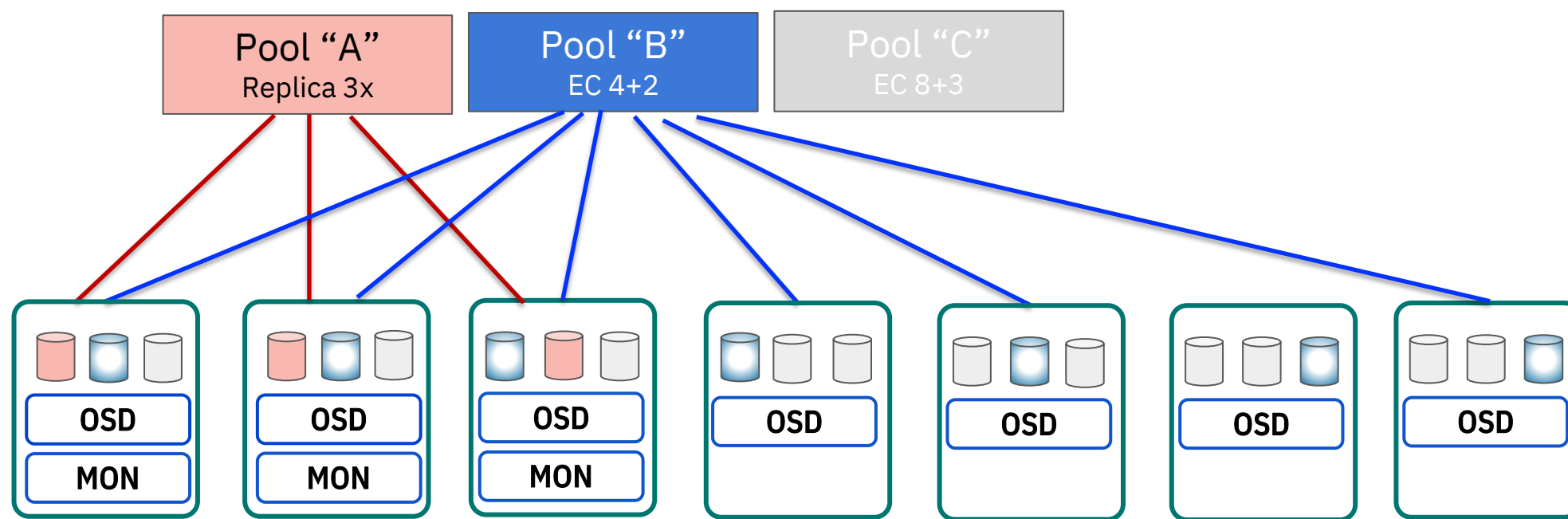
- Provide a storage access entity for each access method (file, block and object)
- Each pool has an assigned data redundancy scheme (e.g. replicated, erasure-coded)
- Storage pools are usually thin provisioned
- Each data object *lives* in only one (1) Ceph pool



IBM Storage Ceph data protection

Ceph *storage* Pools . . .

- Provide a storage access entity for each access method (file, block and object)
- Each pool has an assigned data redundancy scheme (e.g. replicated, erasure-coded)
- Storage pools are usually thin provisioned
- Each data object *lives* in only one (1) Ceph pool

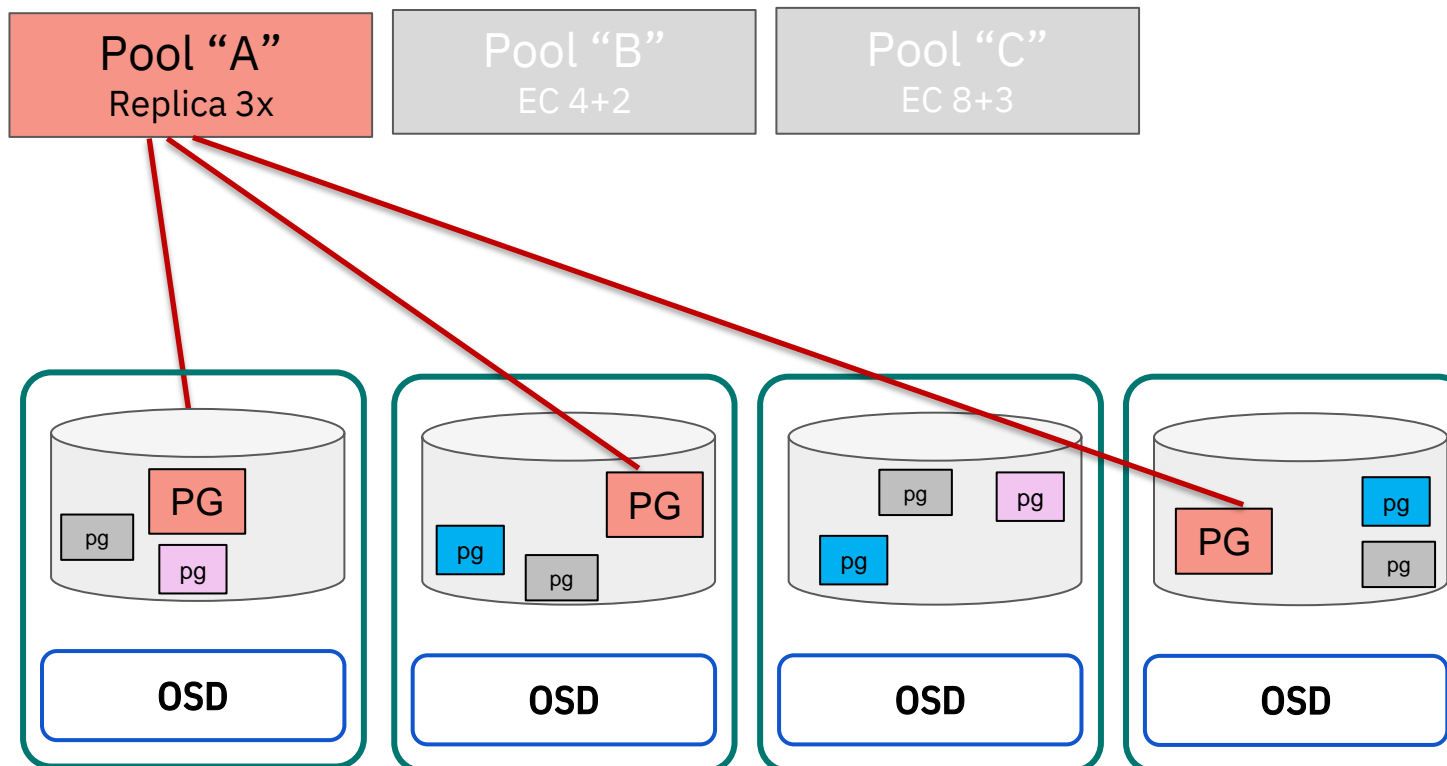


Note: For illustrative purposes. Nodes and full disks are not dedicated to any single pool

IBM Storage Ceph data protection

Placement Groups . . .

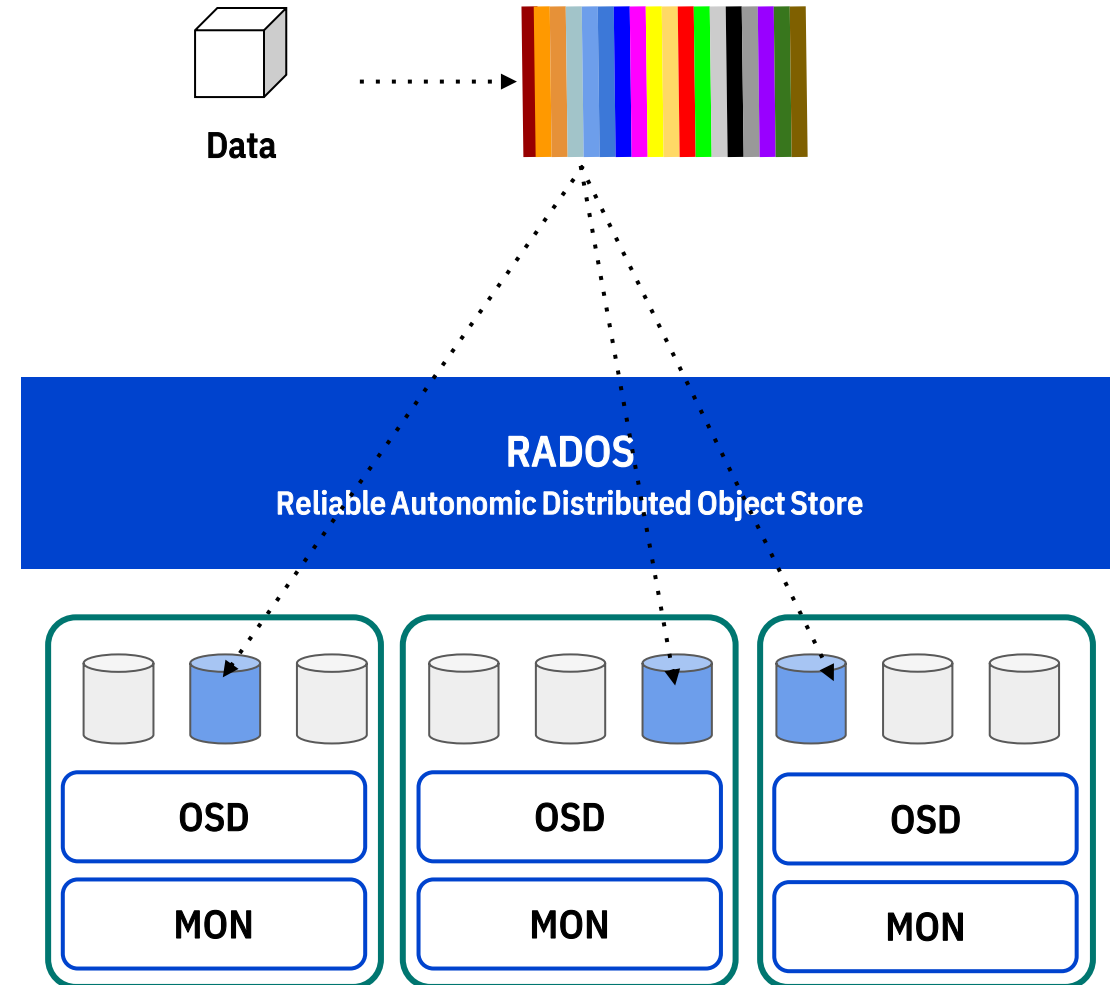
- Placement Groups (PGs) is the method to manage data redundancy at the OSD level
- In Replica x 3 pools for example, each object is stored on 3 different OSDs.
- If an OSD fails, there is a process triggered to find a replacement for that failed OSD and create a third data copy again. This is managed at the PG level)



IBM Storage Ceph data protection

CRUSH (Controlled Replication Under Scalable Hashing)

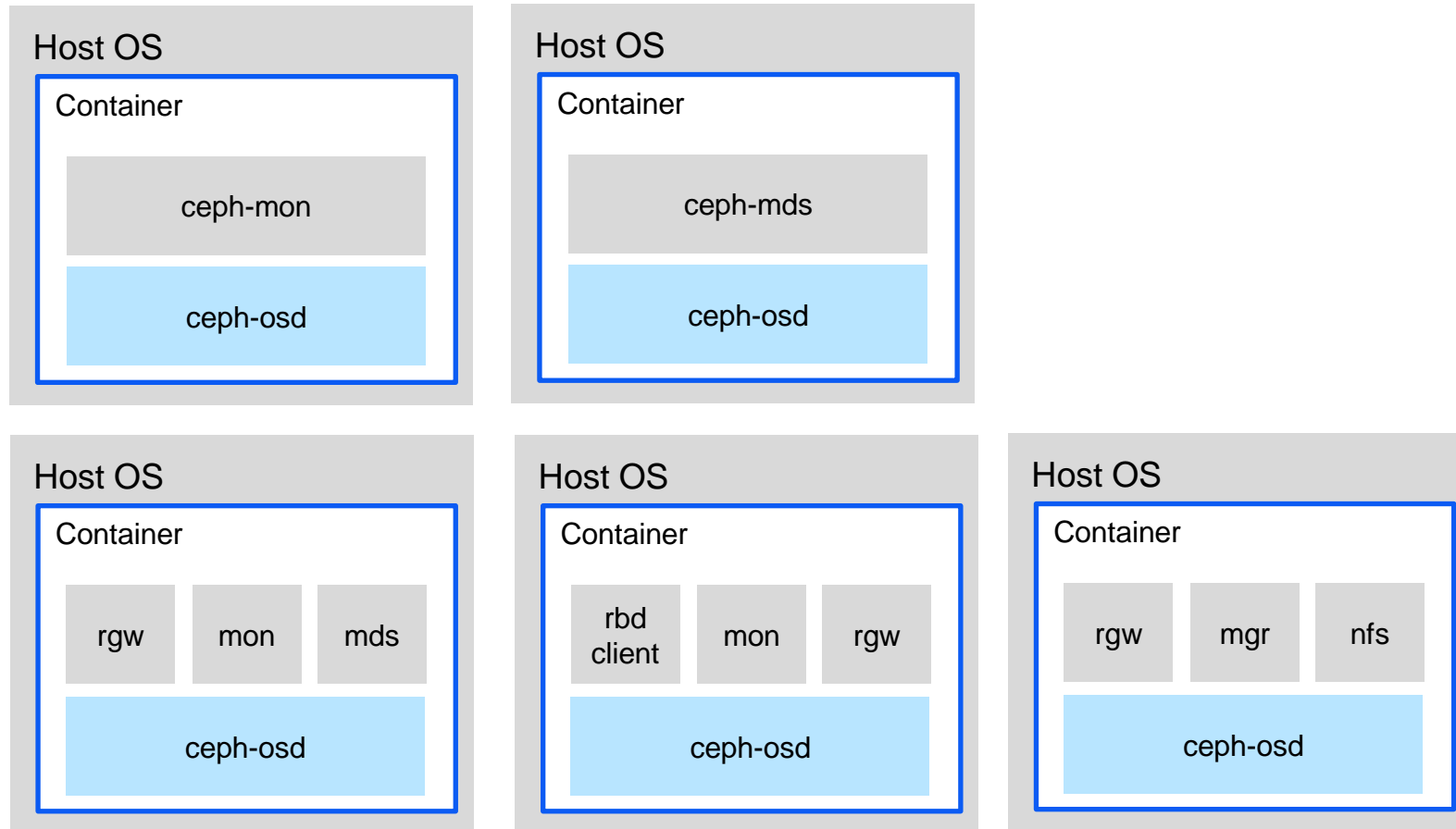
- Pseudo-random placement algorithm
- Fast calculation, no lookup, no gateways
- Repeatable and deterministic
- Rule-based and adjustable
- Weighting
- Can be Data Center, Room, and Rack aware



IBM Storage Ceph Deployment

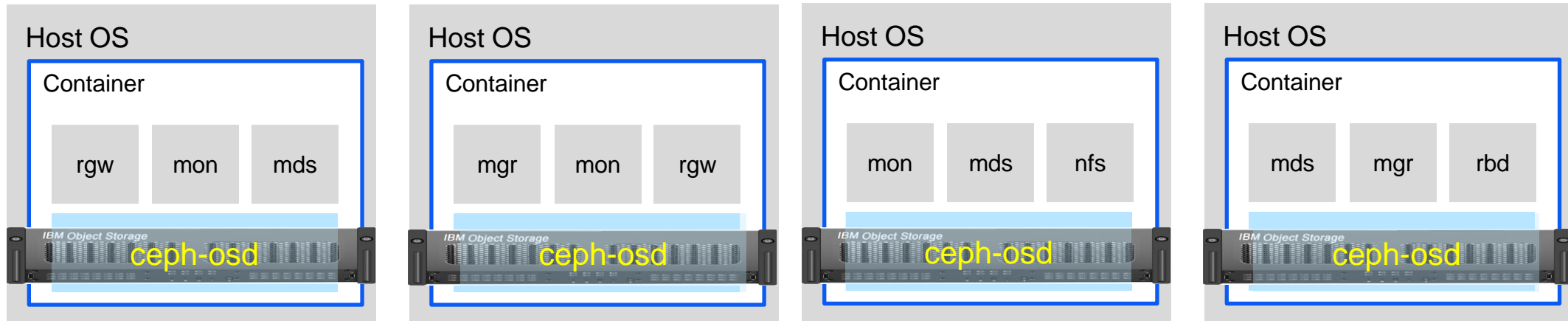


IBM Storage Ceph deployment - colocated daemons

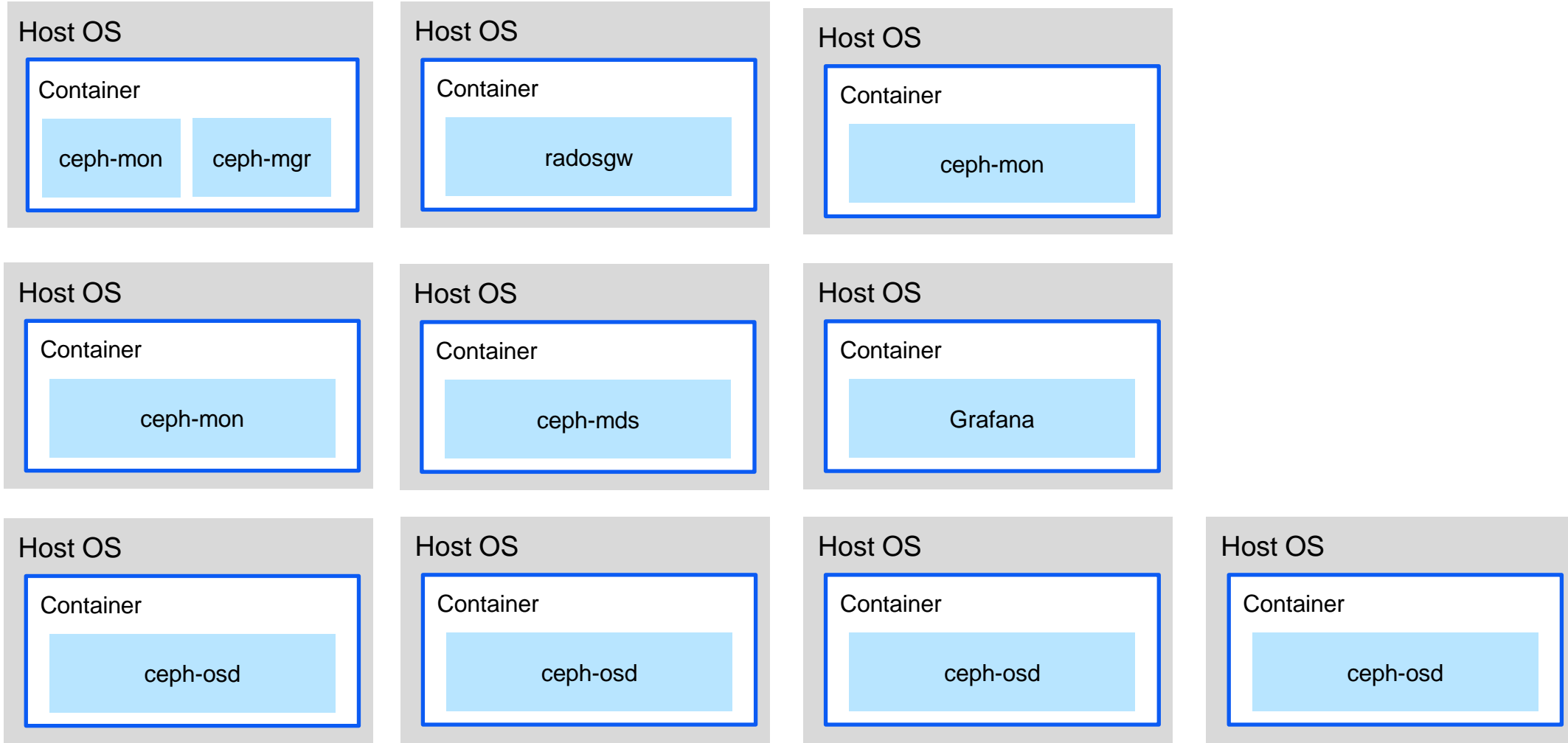


Note: For illustrative purposes only. Not recommended as a best practice.

IBM Storage Ceph deployment – colocated daemons in 4 nodes



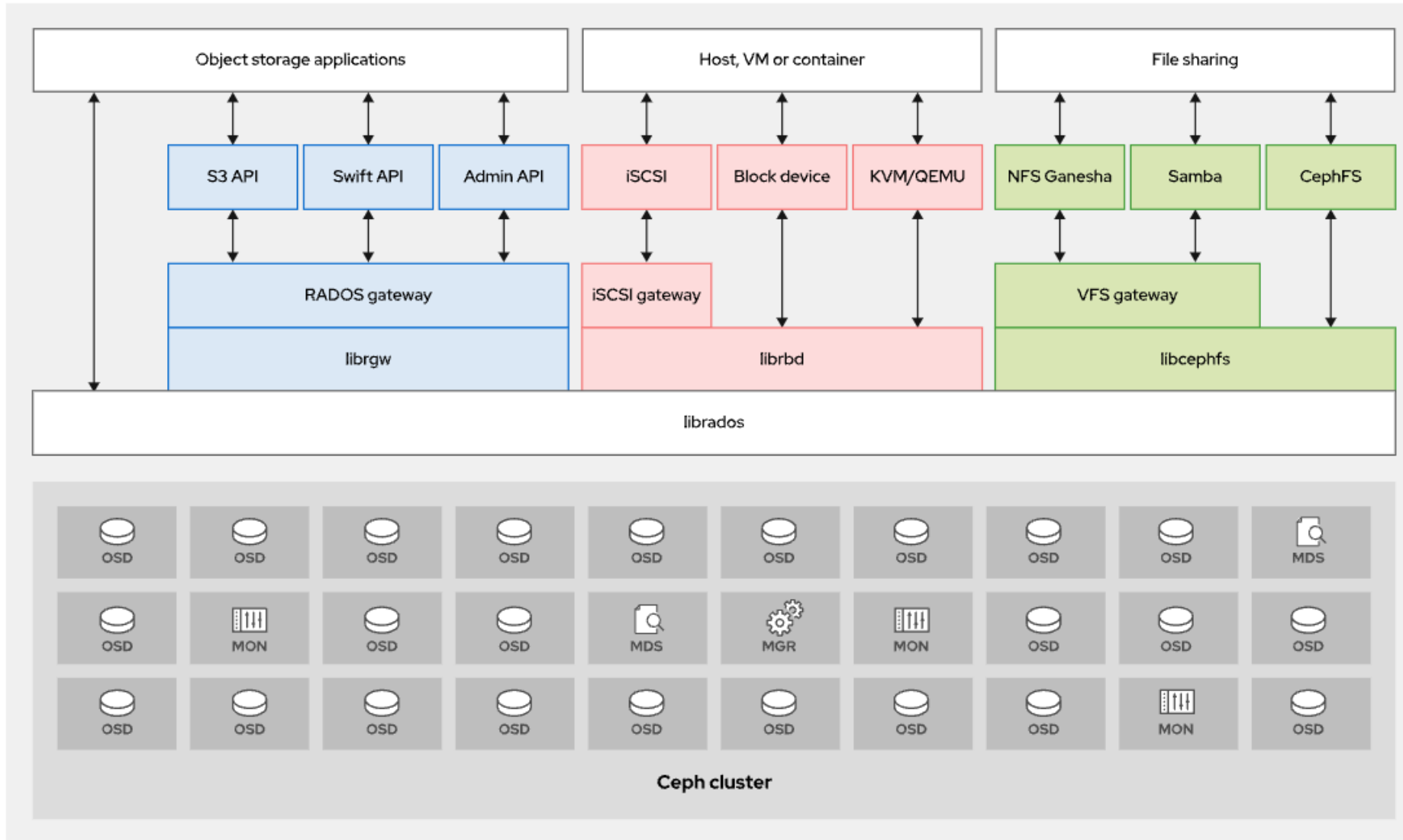
IBM Storage Ceph data protection – non colacated daemons



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

IBM Storage Ceph deployment and data services - putting it all together



IBM Storage Ceph Demo



IBM Technology Zone Test Drive

The screenshot shows a web browser window displaying the IBM Technology Zone page for a test drive. The page title is "North America ATG Storage - IBM Storage Ceph Test Drive - (VMware based)". The page features a dark header with the IBM logo and the slogan "Build. Show. Share.". Below the header, there is a section for "Business value" which describes the test drive as a virtual instance of IBM Storage Ceph installed in a VMware vSphere cluster. The page also includes a sidebar with navigation options like "Business value", "Authors", "Resources", "Environments", "Metadata", and "Comments".

North America ATG Storage - IBM Storage Ceph Test Drive - (VMware based)

Visibility IBMers, Business Partners
Status Active

☆☆☆☆☆ (0) Rate this resource

Business value

The IBM Storage Ceph (VMware based) offers IBMers, Business Partners, and Clients, hands-on access into an operational Ceph storage cluster.

The TechZone Ceph Test Drive is a virtual instance of IBM Storage Ceph installed in a VMware vSphere cluster. This cluster is located in the Advanced Technology Group (ATG) lab in Herndon, Virginia. The IBM Storage Ceph cluster is set up to show proof points of object storage features such through the RADOS Gateway (RGW). The key elements of the object gateway Realm, Zonegroups, Zones, Placement Targets, and Storage Pools are available for discovery and exploration. The end to end client experience via the S3 API as well as the multi-protocol access via NFS is available.

The baseline configuration can also serve as a learning resource, a self-demo, or a customer demo. This system is in a shared lab environment but is treated with the operational care of a production system. Therefore, after each demo, it could be be manually restored to a known good starting point if possible. In the alternate, VMware snapshots can be used to restore the system to a known good state.

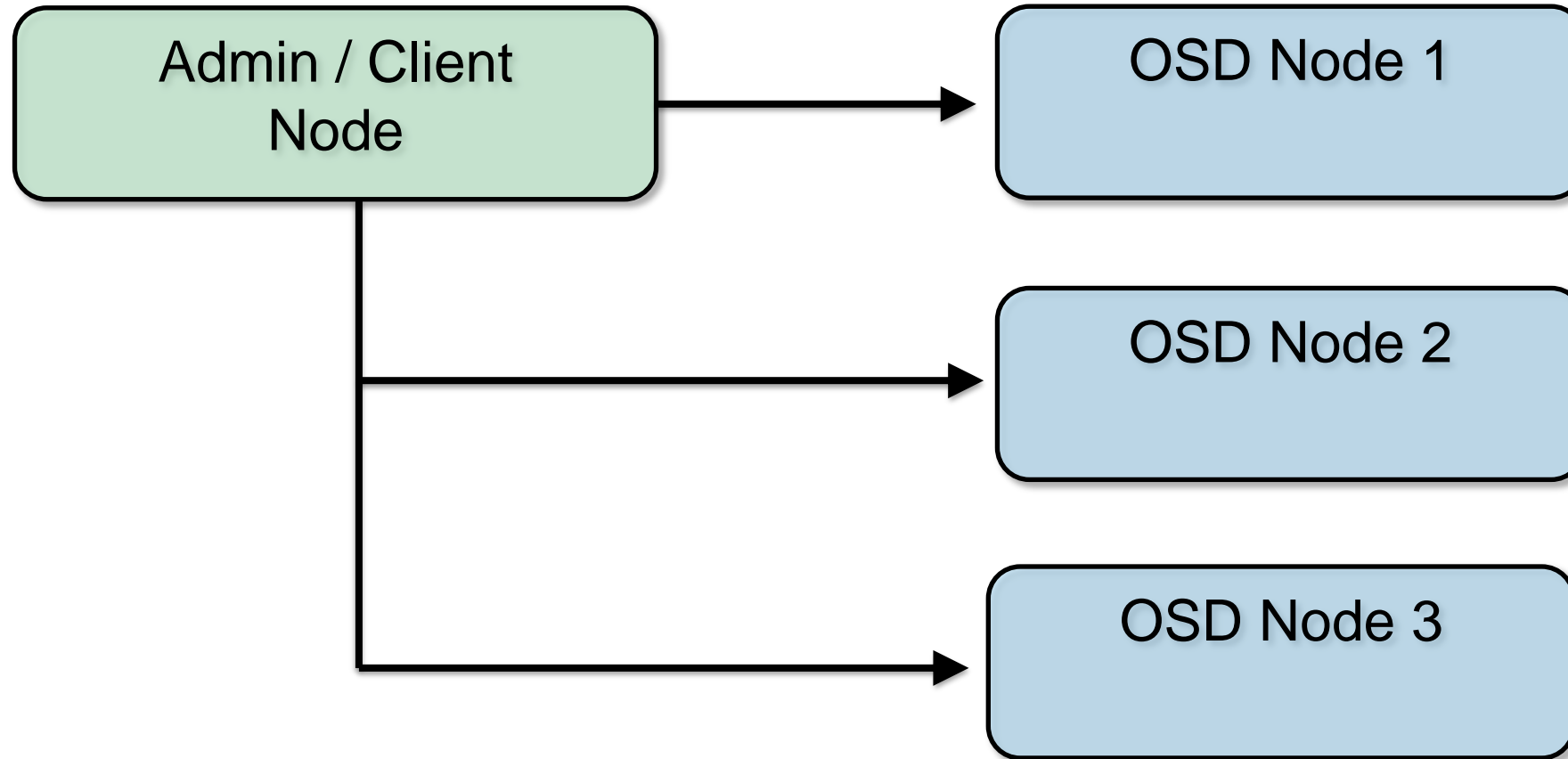
Please Note: This demo will go live as a supported test drive on January 3, 2023. If you reserve this demo in December, please reach out to John Shubeck at jshubeck@us.ibm.com for assistance.

IBM Technology Zone Test Drive

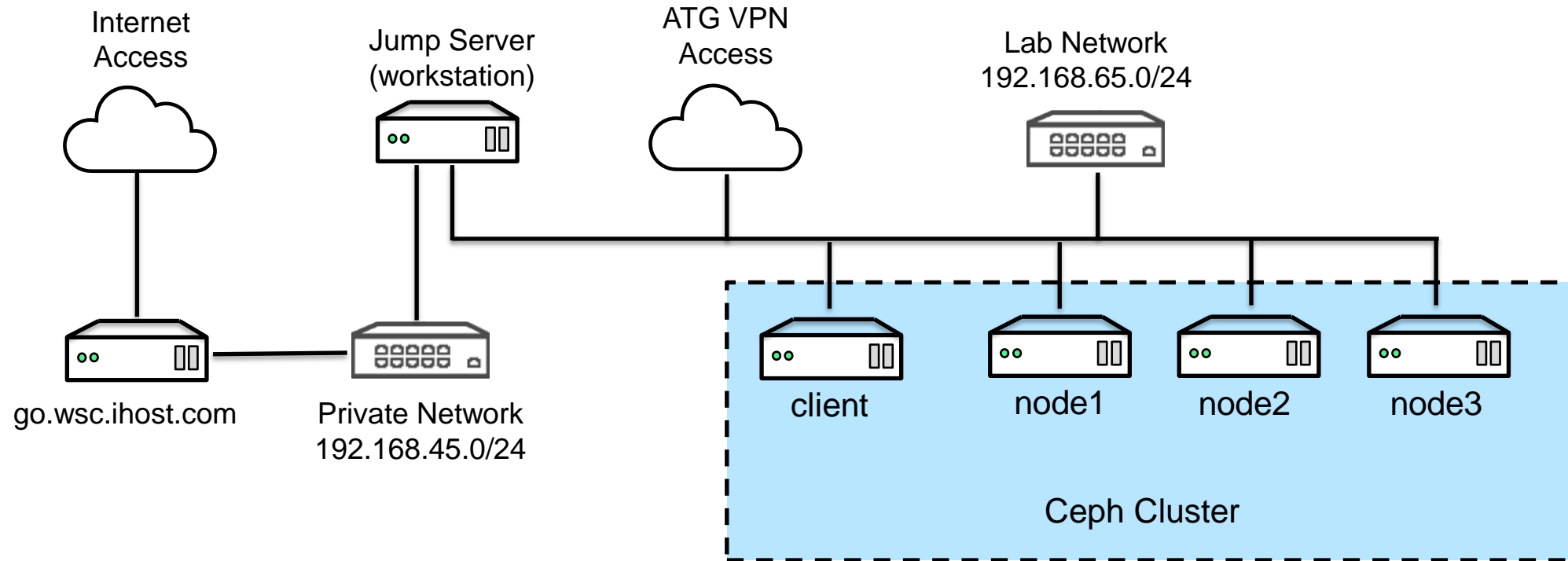
The screenshot shows a web browser window displaying the IBM Technology Zone page for a specific test drive. The browser tabs are 'North America ATG Storage - IB' and 'Ceph Test Drive'. The address bar shows the URL: <https://techzone.ibm.com/collection/north-america-atg-storage-ibm-storage-ceph-test-drive-vmware-based>. The page header includes 'IBM Technology Zone', 'My library', and 'Help'. The main content area features a large title: 'North America ATG Storage - IBM Storage Ceph Test Drive - (VMware based)'. Below the title, it indicates 'Visibility IBMers, Business Partners' and 'Status Active'. There is a 'Bronze' badge in the top right corner. A rating section shows five stars and '(0) Rate this resource'. A sidebar on the left lists navigation options: 'Business value', 'Authors', 'Resources', 'Environments', 'Metadata', and 'Comments'. The main content area is titled 'Resources' and contains three resource cards, each dated 'Jan 7, 2023'.

Resource Title	Description	Visibility
Ceph Test Drive Guided Labs	A guided tour through a learning experience around IBM Storage Ceph with emphasis on the S3 API object storage service.	IBMers
cIBM Storage Ceph Sales Kit on Seismic	Seismic landing page for the IBM Storage Ceph Sales and Technical Sales	IBMers, Business Partners
Ceph Test Drive topology, IP addresses, logins	These documents contains the network map, IP addresses and login accounts needed to access the demo system.	IBMers, Business Partners

IBM Technology Zone: Test drive cluster layout



IBM Technology Zone: Test drive network



IBM Storage Ceph guided labs

Ceph Test Drive

IBM Technology Zone
IBM Storage Ceph Test Drive

Block Storage Lab

File Storage Lab

Object Storage Lab

S3 Client Access Lab

Overview

The TechZone Ceph Test Drive is a virtual instance of IBM Storage Ceph installed in a VMware vSphere cluster. This cluster is located in the Advanced Technology Group (ATG) lab in Herndon, Virginia. The IBM Storage Ceph cluster is set up to show proof points of object storage features such through the RADOS Gateway (RGW). The key elements of the object gateway Realm, Zonegroups, Zones, Placement Targets, and Storage Pools are available for discovery and exploration. The end to end client experience via the S3 API as well as the multi-protocol access via NFS is available.

The baseline configuration can also serve as a learning resource, a self-demo, or a customer demo. This system is in a shared lab environment but is treated with the operational care of a production system. Therefore, after each demo, it could be manually restored to a known good state if possible. In the alternate, VMware snapshots can be used to restore the system to a baseline starting point.

*** NEWS ***

Join us for an ATG Accelerate presentation on January 10, 2023 entitled, "Introduction to IBM Storage Ceph". Details and a registration follow.

Information and registration here: <https://ibm.biz/BdPSGE>

About the author

John Shubeck is an information technology professional with over 41 years of industry experience spanning both the client and technology provider experience. John is currently serving as a Senior Storage Technical Specialist in the ATG (Advanced Technology Group) team, with focus as an IBM Object Storage platform specialist across all industries.

IBM Storage Ceph Flexibility and Completeness



The Flexibility of IBM Storage Ceph – Aligning to diverse customer requirements

High capacity



High speed



Economical



Commodity hardware



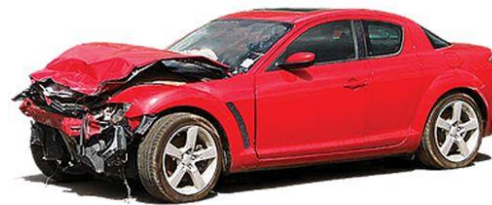
Flexible configurations



Fits any data center



Test drive



IBM support

24/7/365 “roadside” assistance

The Completeness of IBM Storage Ceph – Aligning to enterprise requirements



Efficiency

- Full data reduction option range
- 16X better space use on HDD small file
- 4X better space use on SDD small file



Security

- Write once, read many (WORM) object lock application programming interface (API)
- FIPS 140-2 cryptography
- Interoperate with key management interoperability protocol (KMIP) key managers
- Messenger v2.1 backplane encryption



Performance

- Optimized Librados block device (LibRBD) data path: 80% faster
- Overhauled cache architecture
- 10+ billion objects in RADOS gateway (RGW)
- Ceph file system (CephFS) “Top” tool



Manageability

- New integrated control plane—Cephadm
- Integrated monitoring and management dashboard
- OSD replacement workflow (CLI and UI)
- RGW multisite monitoring



APIs and protocols

- Management API
- CephFS + network file system (NFS)
- CephFS geo-replication

Why IBM Storage Ceph Storage?

REQUIREMENT	IBM VALUE
Innovation Maturity	<ul style="list-style-type: none">- Open Source project, Ceph Foundation- 20 year legacy of software defined storage
S3 Compliance	<ul style="list-style-type: none">- Latest S3 API Features (e.g. Object Lock, S3 Select)- Mature S3 API Implementation (ISV support)
Flexibility	<ul style="list-style-type: none">- Single site Ceph deployment starter- Data center topology mapping
Economics	<ul style="list-style-type: none">- Granular deployment and expansion, storage tiering
Relationship	<ul style="list-style-type: none">- Deep knowledge of open source, block, file, object protocols- IBM proven experience with object storage at massive scale

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](https://www.menti.com) with code 2243 3599 or

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

Or

QR Code



