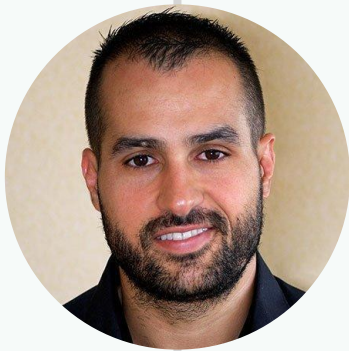




Taming a Cephalopod

Swarm: Multi-cluster Monitoring
Comes to the Ceph Dashboard

Ernesto Puerta - IBM



Ernesto Puerta

epuertat@ibm.com

- Principal Software Eng. IBM
- **24 years'** experience (Telefonica, Alcatel-Lucent, Bell Labs, Nokia, Red Hat)
- **8 years** with Ceph (user & developer)
- Former **Ceph Dashboard** Lead
- Recently involved in **Ceph NVMe over Fabrics**



Agenda

- Multi-cluster Monitoring
- RGW Multi-site Management
- Next Dashboard Features
- How to Contribute





ceph

Multi-cluster Monitoring & Management



Multi-cluster Ceph is a FACT



The average Ceph Cluster:*

	Median	Mean**
Capacity (TB)	24	495
OSDs/Cluster	18	72
Hosts/Cluster	3	12
Clusters/user	2	8



* Data from [Ceph Telemetry](#) and [Ceph User Survey \(2022\)](#).

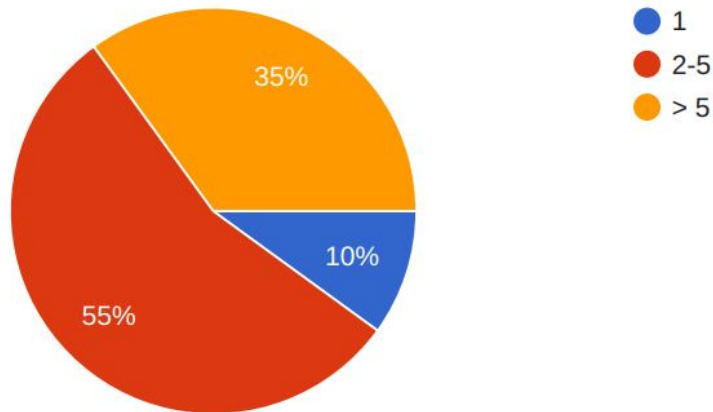
** Difference between median and mean can be explained by asymmetrical multimodal distribution.

Multi-cluster Ceph is a FACT



1. How many Ceph Clusters do you manage?

40 responses



* Data from [Ceph Multi-cluster Survey](#) (2022).

Why multiple Ceph Clusters?



How many of **you** run **multiple** Ceph clusters?

Why?

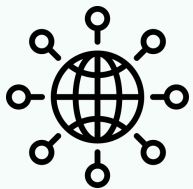
What for?

Why multiple Ceph Clusters?

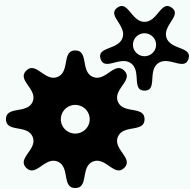
“A Ceph cluster *can* virtually grow indefinitely...”

... but users prefer to **break them down**.

Why?



Reduce
infrastructure
complexity



Reduce
operational
risk-impact



Meet **legal**
requirements



Why multiple Ceph Clusters?

“A Ceph cluster *can* virtually grow indefinitely...”

... but users prefer to **break them down**.

What for?



By **Access**
Protocol: Block,
File, Object



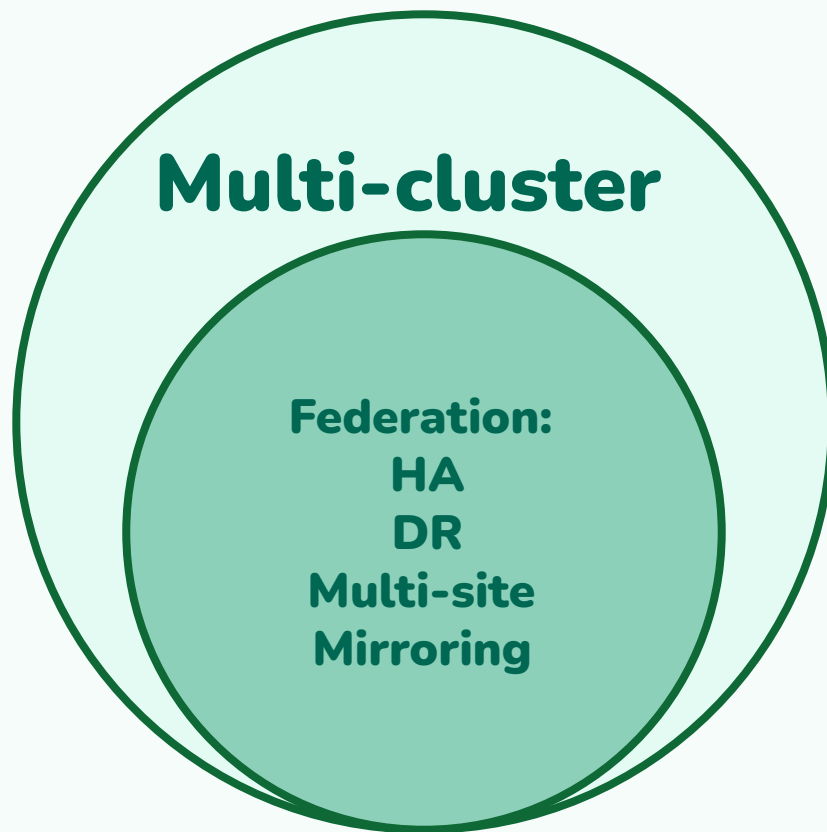
By **Use-case:**
IaaS (OpenStack),
PaaS (k8s)



Other:
Geographic,
Ceph version, ...



Multi-cluster vs. Federation



Multi-cluster with the Ceph CLI



Requires:

```
# /etc/ceph/ceph.conf
```

```
[global]
mon host = [v2:172.20.0.5:3300,v1:172.20.0.5:6789]

[client]
keyring = /etc/ceph/keyring
```

```
# /etc/ceph/ceph-2.conf
```

```
[global]
mon host = [v2:172.20.0.7:3300,v1:172.20.0.7:6789]

[client]
keyring = /etc/ceph/keyring-2
```

```
ceph -c /etc/ceph/ceph.conf -s
```

```
cluster:
  id:      3082be38-b70c-44da-83a3-2373ee451f3a
  health: HEALTH_OK
...
```

```
ceph -c /etc/ceph/ceph-2.conf -s
```

```
cluster:
  id:      931af5e6-34c8-4ca8-a6cc-a09ed0f34a8d
  health: HEALTH_OK
...
```

Multi-cluster with the Ceph CLI: Downsides



- **Security** challenges:
 - Requires monitor connection (3300, 6789).
 - Coarse-grained permissions (“allow *”).
- **Poor UX:**
 - Multiple ceph.conf and keyrings.
 - Clumsy, error-prone syntax (-c <cluster_conf>).
 - Poll-based monitoring.
 - Does not scale well beyond 2-3 clusters.

Multi-cluster with Ceph Dashboard: Goals



- **Operator** friendly (reduce cognitive load):
 - Integrated user experience.
 - Alert-based vs. polling
 - Simplify multi-site/mirroring configuration.
- More **secure**:
 - HTTPS based.
 - Fine-grained credentials.
 - Short-lived configurable authorization (JWT-based).

A Brief History of Multi-cluster in Dashboard



- **Mimic-Nautilus:**
 - RBD mirroring functionality [#20103](#) (2018)
- **Reef:**
 - Multi-cluster support to Prometheus alerts [#47633](#) (2022)
- **Squid:**
 - RGW multi-site configuration [#49953](#) (2023)
 - RGW multi-site migrate from single to multi-site: [#50806](#) (2023)
 - RGW overview dashboard [#52317](#) (2023)
- **Tentacle:**
 - Multi-cluster authentication [#55099](#) (2024)
 - Multi-cluster management using context switcher [#55213](#) (2024)

Multi-cluster: Wizard



The screenshot shows the Ceph web interface. On the left is a dark sidebar with a menu. The 'Multi-Cluster' option is highlighted with a red box, and its sub-items 'Overview' and 'Manager Cluster' are also visible. The main content area has a light gray background. At the top right of the main area, there are links for 'English', a notification bell, a help icon, a settings gear, and a user profile icon. In the center of the main area, there is a large teal wrench icon. Below the icon, the text 'Connect Cluster' is displayed in bold. Underneath this, a paragraph reads: 'Upgrade your current cluster to a multi-cluster setup effortlessly. Click on the "Connect Cluster" button to begin the process.' At the bottom of this section, there is a teal button with a white plus sign and the text 'Connect Cluster', which is also highlighted with a red box.

× ceph English ▾ 🔔 ? ⚙️ 👤

- Dashboard
- Multi-Cluster** ^
 - Overview
 - Manager Cluster
- Cluster ▾
- Block ▾
- Object ▾
- File ▾
- Observability ▾
- Administration ▾

Connect Cluster

Upgrade your current cluster to a multi-cluster setup effortlessly. Click on the "Connect Cluster" button to begin the process.

+ Connect Cluster

Multi-cluster: New Cluster



× ceph English ▾ 🔔 ? ⚙️ 👤

- Dashboard
- Multi-Cluster
- Overview
- Manager Cluster
- Cluster
- Block
- Object
- File
- Observability
- Administration

Connect Cluster

Cluster API URL ? * ✓

Alias Name * ✓

Username * ✓

Password * ✓ 👁️ 📋

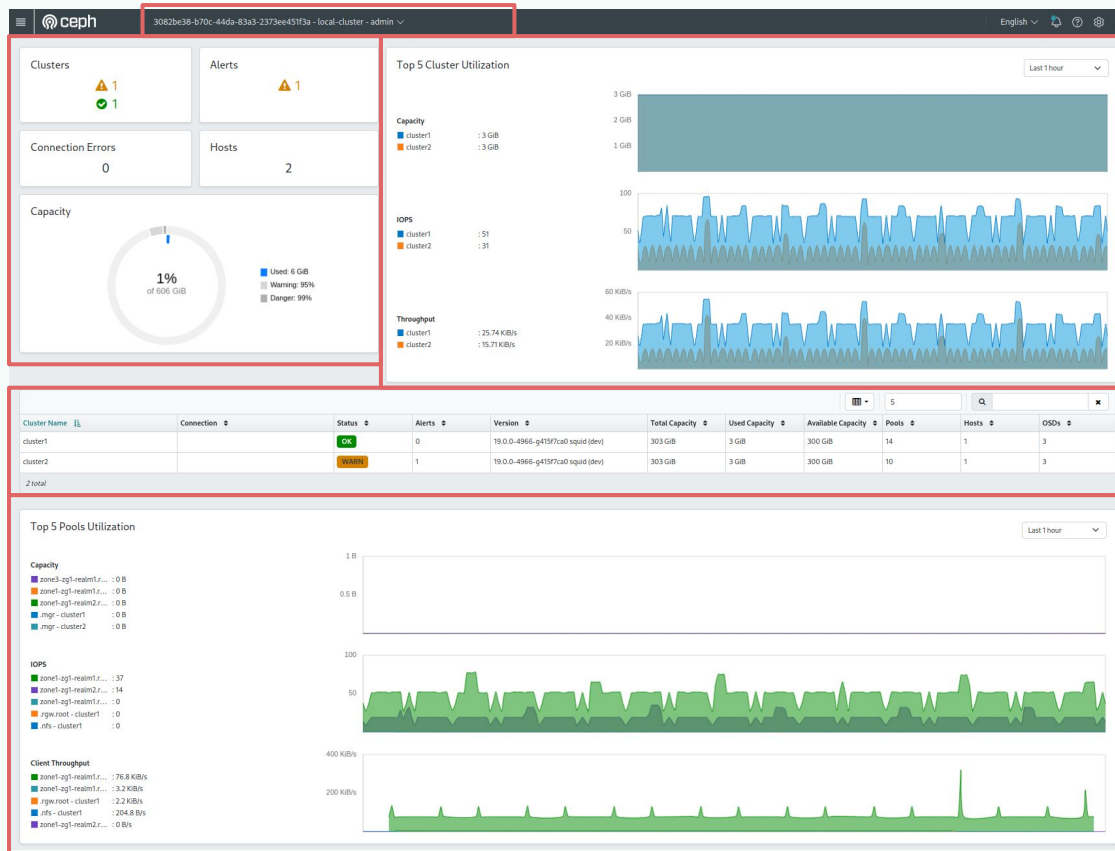
Login Expiration ▾

☐ SSL

Cancel Connect Cluster

...t Cluster" button to

Multi-cluster: Dashboard



Multi-cluster: Cluster context selector



3082be38-b70c-44da-83a3-2373ee451f3a - local-cluster - admin ^

931af5e6-34c8-4ca8-a6cc-a09ed0f34a8d - ceph-cluster-2 - admin

Multi-cluster: Dashboard - Status



Clusters



Alerts



Connection Errors

0

Hosts

2

Capacity



■ Used: 6 GiB
■ Warning: 95%
■ Danger: 99%

Multi-cluster: Dashboard - Utilization

Top 5 Cluster Utilization

Last 1 hour

Capacity

cluster1 : 3 GiB
cluster2 : 3 GiB



IOPS

cluster1 : 63
cluster2 : 38



Throughput

cluster1 : 32.84 KiB/s
cluster2 : 19.13 KiB/s



						5	Q				×
Cluster Name	Connection	Status	Alerts	Version	Total Capacity	Used Capacity	Available Capacity	Pools	Hosts	OSDs	
cluster1		OK	0	19.0.0-4966-g415f7ca0 squid (dev)	303 GiB	3 GiB	300 GiB	14	1	3	
cluster2		WARN	1	19.0.0-4966-g415f7ca0 squid (dev)	303 GiB	3 GiB	300 GiB	10	1	3	
2 total											

Multi-cluster: Dashboard - Top 5 Pools



Top 5 Pools Utilization

Last 1 hour

Capacity

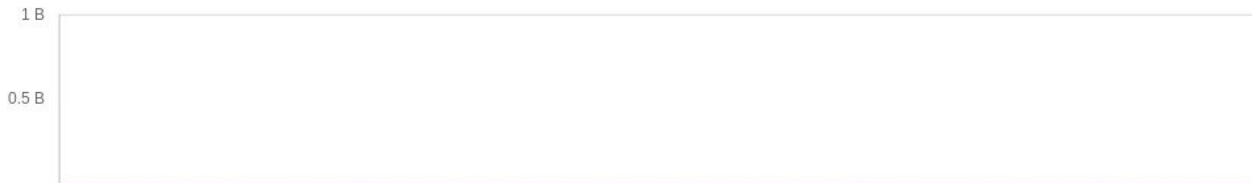
zone3-zg1-realm1.r...	: 0 B
zone1-zg1-realm1.r...	: 0 B
zone1-zg1-realm2.r...	: 0 B
.mgr - cluster1	: 0 B
.mgr - cluster2	: 0 B

IOPS

zone1-zg1-realm1.r...	: 77
zone1-zg1-realm2.r...	: 19
zone1-zg1-realm1.r...	: 0
.rgw.root - cluster1	: 0
.nfs - cluster1	: 0

Client Throughput

zone1-zg1-realm1.r...	: 76.6 KiB/s
zone1-zg1-realm1.r...	: 3.2 KiB/s
.rgw.root - cluster1	: 2.2 KiB/s
.nfs - cluster1	: 0 B/s
cephfs.a.data - clust...	: 0 B/s



Multi-cluster: Dashboard - Cluster Manager



3082be38-b70c-44da-83a3-2373ee451f3a - local-cluster - admin ▾

English ▾



Multi-Cluster » Manage Clusters

Clusters List

+ Connect ▾



10



Alias	Connection	FSID	URL	User	Token expires
ceph-cluster-2	CONNECTED	931af5e6-34c8-4ca8-a6cc-a09ed0f34a8d	https://172.20.0.7:11000	admin	14 days 16 hours 23 minutes
local-cluster	CONNECTED	3082be38-b70c-44da-83a3-2373ee451f3a	https://localhost:11000	admin	N/A

0 selected / 2 total

- **One-click set-up** for:
 - RGW multi-site
 - RBD mirroring



RGW

Multi-site



RGW Multi-site Management: Overview



Object Gateway » Overview

Inventory

2 Daemons	2
2 Realms	2
4 Zonegroups	4
8 Zones	8
1 Bucket	1
1 User	1
963 Objects	963

Used Capacity

7.5 MiB

Avg Object Size

7.9 KiB

RGW Multi-site Management: Wizard



Dashboard 🏠

Cluster >

Pools

Block >

NFS

File Systems

Object Gateway ▾

Daemons

Users

Buckets

Multisite



Multi-site not configured

Multi-site provides disaster recovery and may also serve as a foundation for content delivery networks

Add Multi-site Configuration

Import Multi-site Configuration

RGW Multi-site Management: Topology-view



Object Gateway » Multisite



Information

You need to enable the rgw module to access the import/export feature. [Enable RGW Module](#)

+ Create Realm

Import

Export

Topology Viewer

▼ realm2

▼ zg2-realm2

zone1-zg2-realm2 master

zone2-zg2-realm2

▼ zg1-realm2 master

zone2-zg1-realm2

zone1-zg1-realm2 master

▼ realm1 default

▼ zg1-realm1 default master

zone2-zg1-realm1

zone1-zg1-realm1 default master

▼ zg2-realm1

zone2-zg2-realm1

zone1-zg2-realm1 master

RGW Multi-site Management: Migration



Migrate Single Site to Multi-site ?

Migrate from a single-site deployment with a default zonegroup and zone to a multi-site system

Rename default zonegroup *

Zonegroup name...

Zonegroup Endpoints *

e.g. http://ceph-node-00.com:80

Rename default zone *

Zone name...

Zone Endpoints *

e.g. http://ceph-node-00.com:80

Cancel

Create



Next Ceph Dashboard Features



Next Ceph Dashboard features (Squid & T)



- **CephFS**

- Subvolume/Subvolume group
- Snapshots and clone
- Snapshots Scheduler
- Access management

- **RGW**

- Advanced Workflows
- User roles/policies
- Bucket policies
- Lifecycle policies
- Multi-site Workflow
- Realm, Zone group and Zone management
- Import and export of token

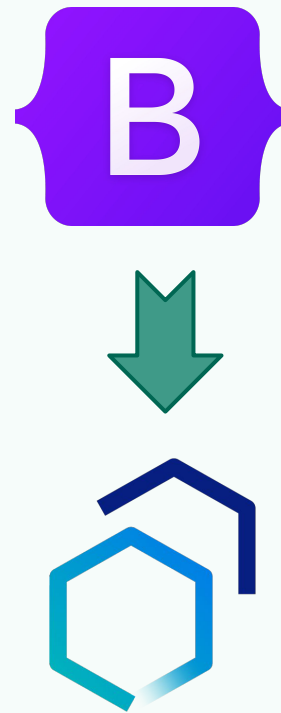
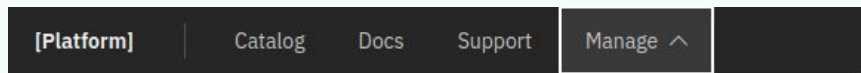
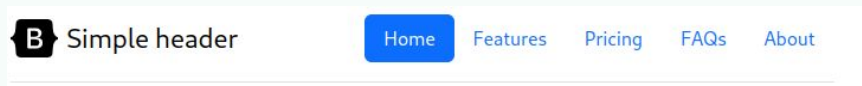
- **NVMe-oF**

- **UI/UX Improvements**

- UI framework replacement: from Bootstrap to IBM Carbon
- More changes and improvements to Landing page

New Features: Carbon Design System

- Replaces Twitter's **Bootstrap**
- IBM-backed **Open Source** project
- **Native Angular** components
- **Advanced** UI patterns: charts, tables, tree views, ...
- Sleek & sober (e.g.: flat square buttons):



New Features: Carbon Design System



Dashboard

Multi-Cluster

Cluster

Block

Object

File

Observability

Administration

English



Cluster Utilization

Last 30 minutes

Used Capacity (RAW)

■ 3 GiB used of 303 GiB



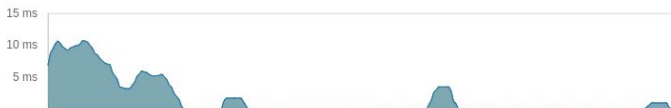
IOPS

■ Reads : 33
■ Writes : 5



OSD Latencies

■ Apply : 0 ms
■ Commit : 0 ms

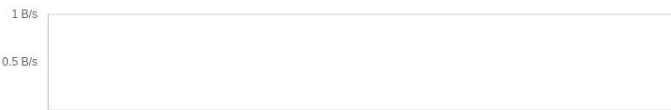


Client Throughput

■ Reads : 18.82 KiB/s
■ Writes : 0 B/s



Recovery Throughput



How to contribute



- Virtual **meetings**:
 - Ceph **Developer Monthly** (Wednesday)
 - Ceph **Users + Devs** Monthly (Thursdays)
- As a **user**:
 - Getting help:
 - ceph-users@lists.ceph.com,
 - Slack [#ceph-dashboard](#)
 - Reporting issues or requesting features: tracker.ceph.com
- As a **developer**:
 - dev@ceph.io
 - github.com/ceph



Questions?



Thank you!