Launch of Ceph User Council
April 2024

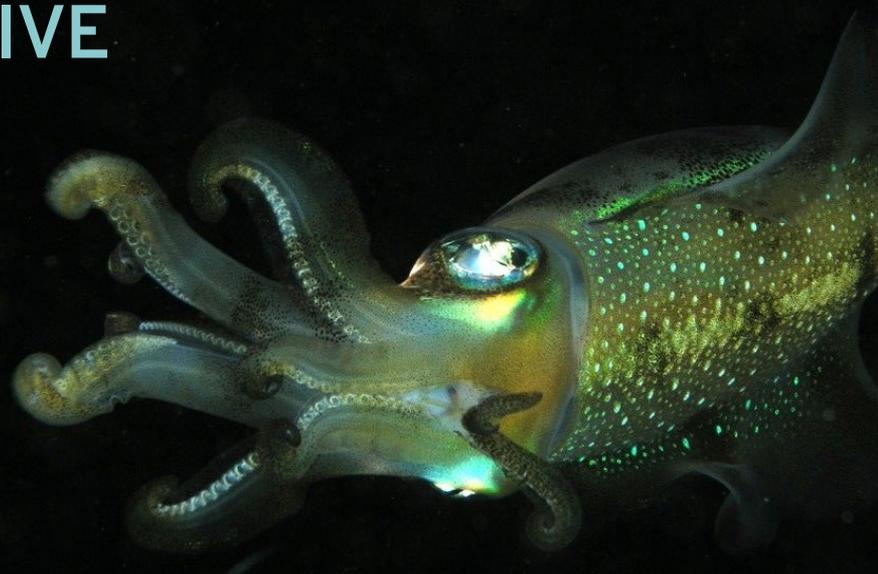**Vincent Hsu, Neha Ojha, Laura Flores**

Agenda

- Objective
- Workflow Overview
  - Organizing Focus Topics
  - Collecting User Feedback
  - Processing Feedback and Taking Action
- Initial Survey Feedback Highlights
  - Performance
  - Orchestration & Deployment
- Next Steps
- Questions

OBJECTIVE

Objective

- Maintainers of the Ceph project value user feedback!
- Several formal initiatives to capture user community over the years:
  - Ceph User Survey
  - Ceph User + Dev Monthly Meetup
- The Ceph Foundation is launching a **new initiative** to elicit more concerted user feedback.
- Targets include users, companies, and organizations who use Ceph in their production environments.
- Feedback will be collected via focused surveys covering broad themes to capture feature requirements and understand pain points.
- Ceph developers will be able to take action based on the feedback we collect.
- Goal is to grow a **user council** with champions leading various pillars of Ceph.

Pillars of Ceph
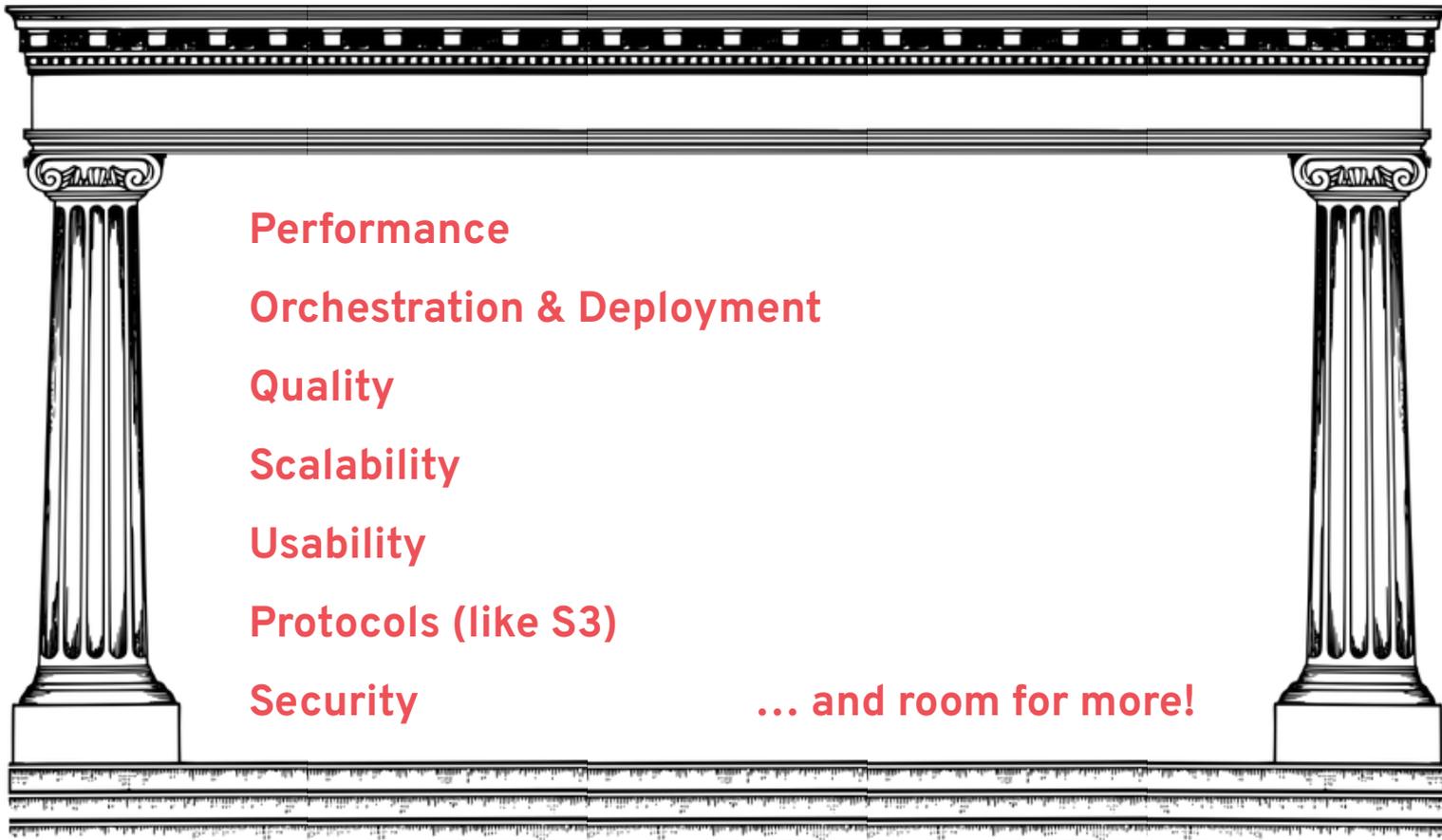
Performance

Orchestration & Deployment

Quality

Scalability

Usability

Protocols (like S3)

Security   … and room for more!

WORKFLOW OVERVIEW

Organizing Focus Topics

- To keep things streamlined, we will start with the following pillars:
  - **Performance** - championed by Vincent Hsu
  - **Orchestration & Deployment** - championed by Matthew Leonard
- Champions will help guide discussion and organize community talking points related to their respective pillars.
- Champions will work with meeting coordinators Neha Ojha and Laura Flores to organize the monthly agenda.
- In the future, we plan to include more champions (volunteers from the community) to lead additional pillars
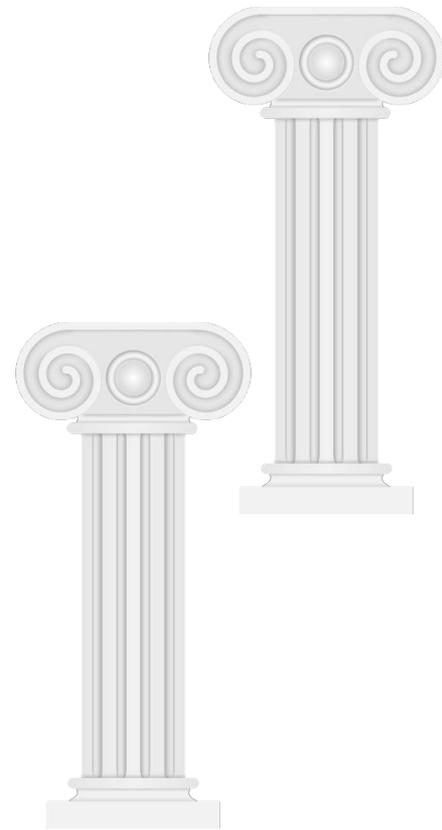
Collecting User Feedback

- A survey will be created in Google Forms (or similar) to gather user feedback.
- Champions will help design the survey questions around their respective pillars.
- Survey will be sent out to community forums:
  - Mailing lists
  - Slack
  - Social media
  - Word of mouth
  - … etc.
- The larger the audience, the better. Any and all Ceph users are encouraged to participate!

Processing Feedback and Taking Action

- Survey feedback will be processed in data visualizations and by analyzing raw results.
- Champions will use background knowledge to review feedback and identify pain points.
- Survey results will also be correlated with data sent by users via telemetry.
- Insights will be discussed at the User + Dev meetings where champions, users, and developers will work together to create action items.
- Users are encouraged to attend to provide additional real-time feedback.

INITIAL FEEDBACK HIGHLIGHTS
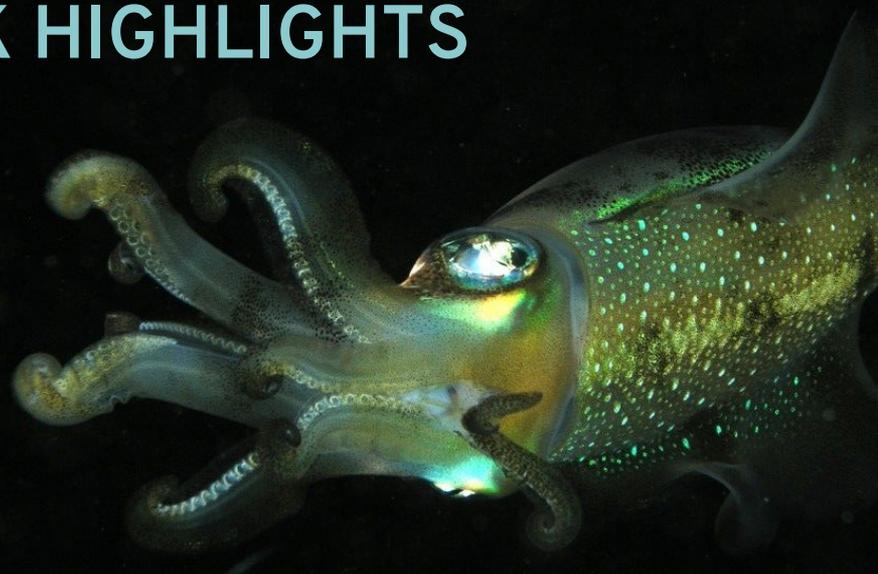
Survey Overview

- What did we want to gain from this survey?
  - Focus on two pillars: **Performance** and **Orchestration & Development**
  - Learn about general performance requirements and pain points
  - Learn about orchestration & deployment use cases and pain points
- Champions (Vincent and Matthew) helped formulate targeted questions.
- Survey was created on Google Forms; open for 1 week.
- We sent it out to Ceph mailing lists:
  - User list: ceph-users@ceph.io
  - Developer list: dev@ceph.io
- Population:
  - ~149 active participants from the user list
  - ~25 active participants from developer list
  - Some overlap between lists
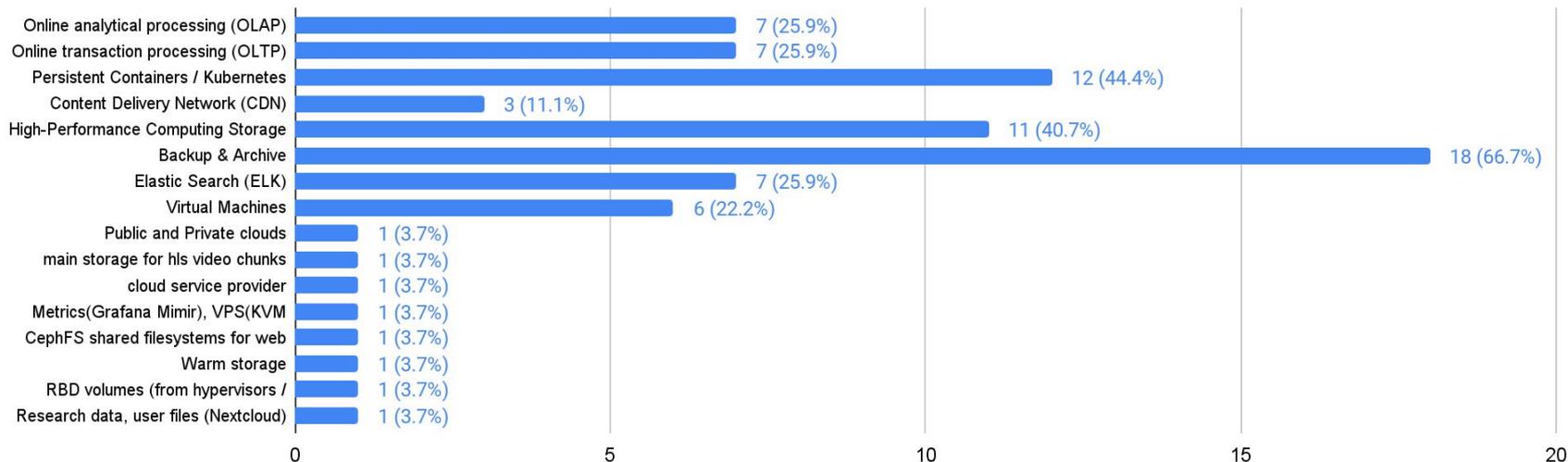- Total responses:
  - 27

Workload Highlights

- **Backup & Archive** was the most popular Ceph workload.
- **Persistent Containers / Kubernetes** placed second.

## Workloads

(participants could indicate more than one choice)

Workload	Count
Online analytical processing (OLAP)	7 (25.9%)
Online transaction processing (OLTP)	7 (25.9%)
Persistent Containers / Kubernetes	12 (44.4%)
Content Delivery Network (CDN)	3 (11.1%)
High-Performance Computing Storage	11 (40.7%)
Backup & Archive	18 (66.7%)
Elastic Search (ELK)	7 (25.9%)
Virtual Machines	6 (22.2%)
Public and Private clouds	1 (3.7%)
main storage for hls video chunks	1 (3.7%)
cloud service provider	1 (3.7%)
Metrics(Grafana Mimir), VPS(KVM	1 (3.7%)
CephFS shared filesystems for web	1 (3.7%)
Warm storage	1 (3.7%)
RBD volumes (from hypervisors /	1 (3.7%)
Research data, user files (Nextcloud)	1 (3.7%)

Performance Highlights

- Most participants use both **replication** and **erasure coding**.

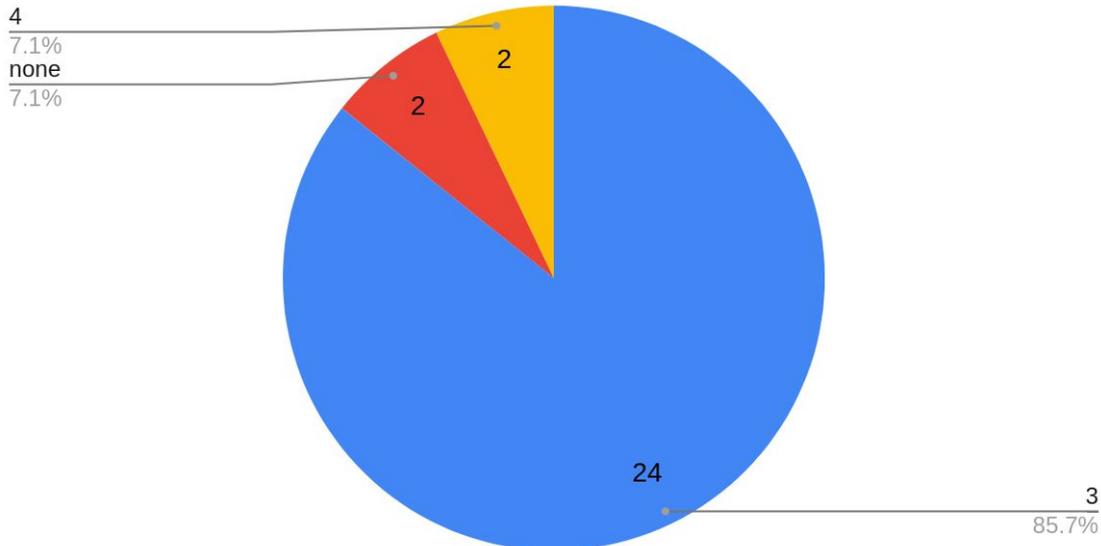Data Protection Technique

Replication
18.5%

5

Erasure Coding
7.4%

2

Both
74.1%

20
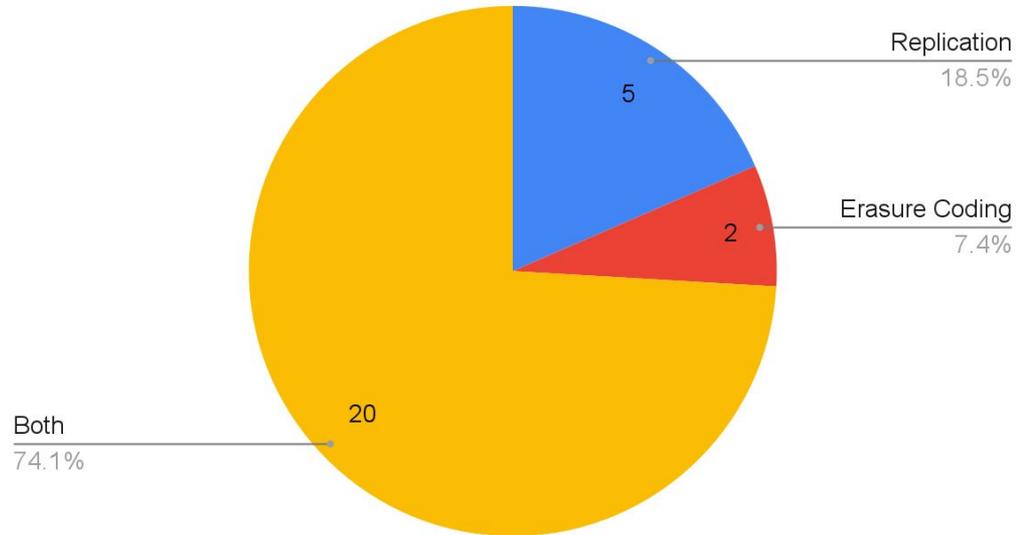
- Most participants who use replication have **3** replicas.
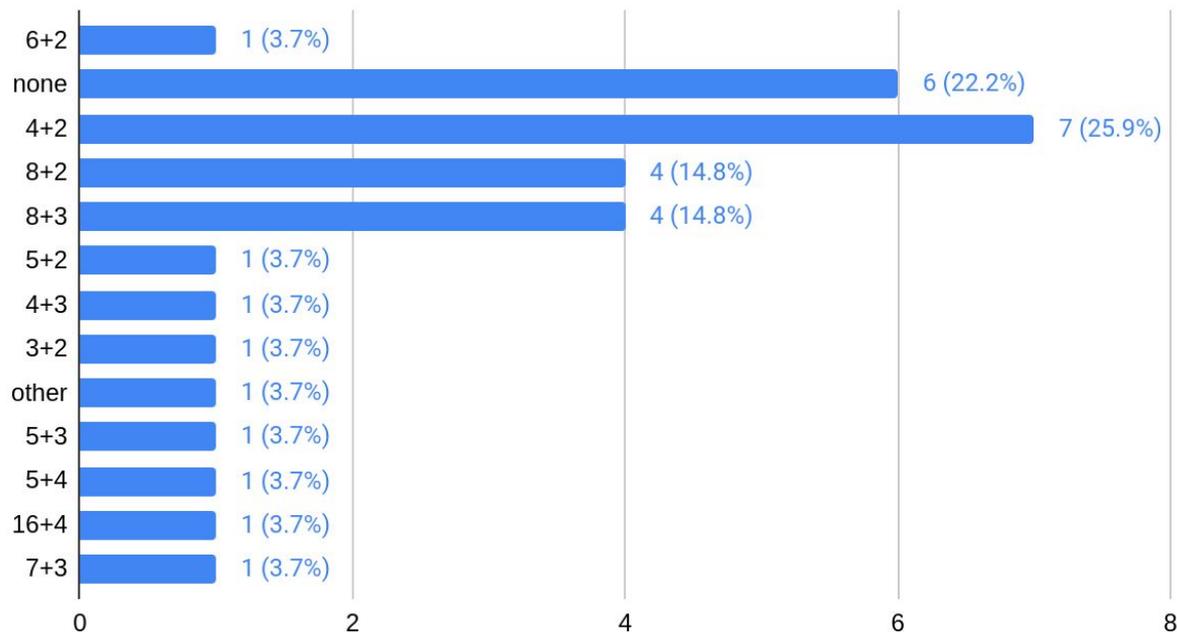
Number of Replicas

(participants could indicate more than one choice)

Performance Highlights

- Most participants who use erasure coding have **4+2** EC profiles.
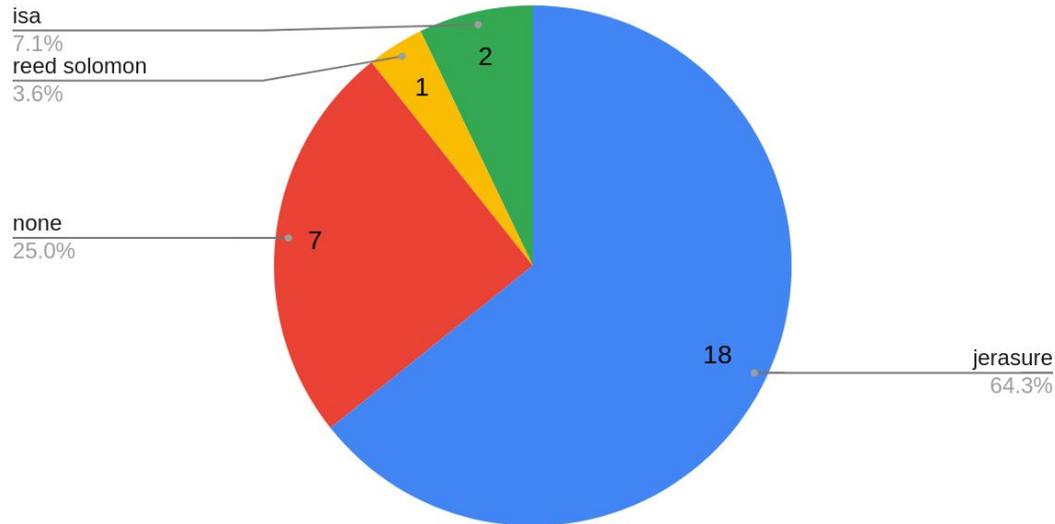- **8+2** and **8+3** placed second.

- Most participants who use erasure coding use the **jerasure** plugin (default setting).

EC Plugin
(participants could indicate more than one choice)

isa
7.1%

reed solomon
3.6%

none
25.0%

jerasure
64.3%

18

7

1

2

Performance Highlights

- Varied network configurations:
  - 10G, 25G, 40G, and 100G ethernet connections
  - A mix of routing protocols: (LACP, BGP, FRR)
  - Some clusters with dedicated replication networks; some without
  - Some reported unique use cases, including 2 x 10Gb for clients access and 2 x 40G for backend

- Varied latency and IOPS requirements depending on configuration:
  - Some indicated they had no specific requirements
  - "Lower is always better"
  - Some had more specific requirements, such as 150k IOPS below 10ms latency

Performance Highlights

- Overall satisfaction with performance in the non-error scenario:
    - Most expressed satisfaction with Ceph's performance
    - Some noted occasional delays, for example in database use cases and certain workloads
    - Some said they have consistent performance with exceptions, such as slow RBD mirroring and slowness for heavy random read workloads on HPC storage

- Mixed reviews on performance during error scenarios:
    - mClock has helped rebalance performance in newer releases, but users would like more control over prioritizing client I/O over other operations
    - Dissatisfaction with performance during longer periods of backfill
    - Peering degrades performance (large latency impact when shutting down/starting OSDs, especially in clusters with large OMAPs)

Performance Analysis Feedback

- Various telemetry systems are used to debug Ceph performance issues:
  - Prometheus, grafana, and mimir for metrics collection and visualization
  - Linux tools like iostat, iotop, atop
  - Ceph-specific tools like the Ceph Dashboard, Ceph CLI, ceph-exporter

- Constructive feedback about performance "best practices" methods:
  - Several reported reading the ceph-users mailing list for performance advice
  - Others reported trying out configurations on test clusters before moving them to production
  - Another report mentioned keeping things as "default" as possible
  - Some pointed out that we could use more "best practices" in the Ceph documentation

Orchestration Systems Distribution

- Participants evenly deploy Ceph on **containers** and **bare metal**.

Orchestration Deployment Targets
(participants could indicate more than one choice)

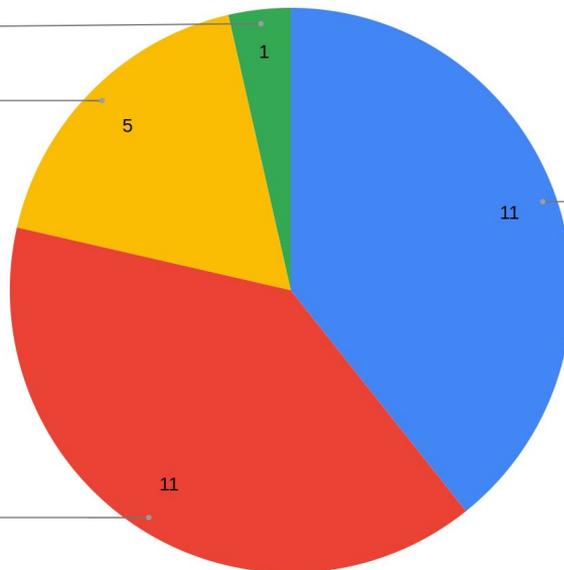In-house orchestration
3.6%

Both Containers and Bare Metal
17.9%

Containers (cephadm/Rook)
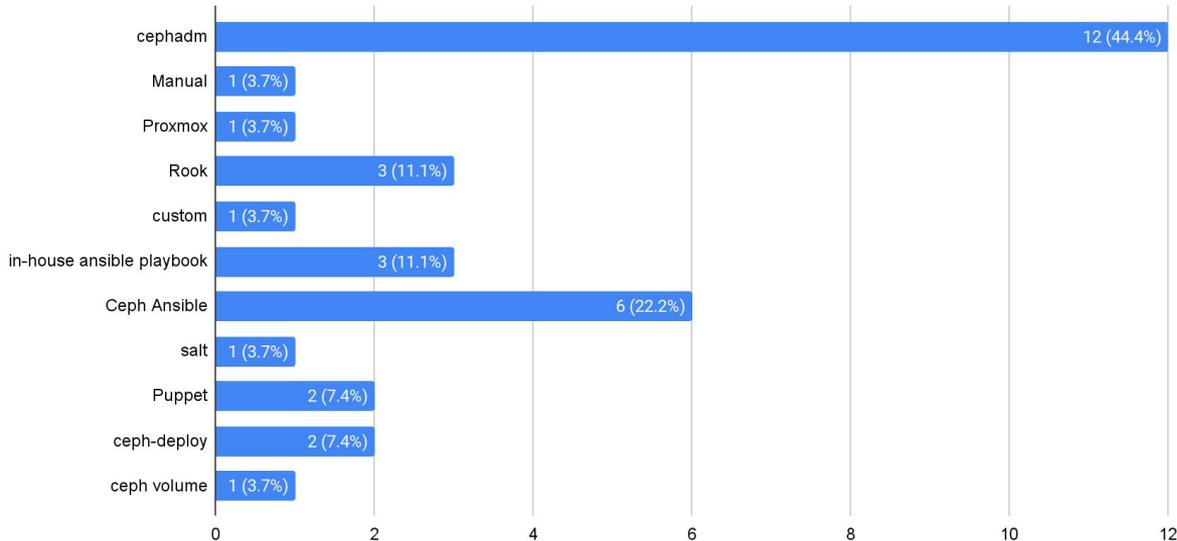39.3%

Bare metal (RPMs/Binary)
39.3%

Orchestration Systems Distribution

- The majority of participants use **cephadm** as their orchestrator
- **Ceph Ansible** was the second most popular choice

Count of Orchestration System

(participants could indicate more than one choice)

System	Count
cephadm	12 (44.4%)
Manual	1 (3.7%)
Proxmox	1 (3.7%)
Rook	3 (11.1%)
custom	1 (3.7%)
in-house ansible playbook	3 (11.1%)
Ceph Ansible	6 (22.2%)
salt	1 (3.7%)
Puppet	2 (7.4%)
ceph-deploy	2 (7.4%)
ceph volume	1 (3.7%)

Orchestration and Deployment Highlights

- Positive feedback:
  - Satisfaction with automation and ease of use
  - "Makes updates / upgrades fast and easy"

- Areas to improve:
  - Call for better logging and handling of silent failures to improve debugging
  - Lack of documentation for certain tasks, such as replacing a drive
  - Certain commands are too verbose and/or not intuitive
  - Hard to debug cases when containers enter error states
  - Rook is simplistic and lacks some of cephadm's features (easy disk replacement, allowance for more flexibility)

Next Steps

- Orchestration and Deployment feedback will be shared with Ceph Leadership Team and appropriate stakeholders
- Increase outreach of future surveys:
  - Social media (LinkedIn, X)
  - Mailing lists
  - Slack
  - Word of mouth!
- Involve future champions to cover additional pillars such as Quality and Scalability

QUESTIONS?

Ceph User + Dev Meeting Details

- When: every third Thursday at 14:00 UTC
- Next meeting: May 16th, 2024
- Where: https://meet.jit.si/ceph-user-dev-monthly
- Agenda: https://pad.ceph.com/p/ceph-user-dev-monthly-minutes
- Interested in becoming a champion in the user council? Email meeting coordinators:
  - Neha Ojha nojha@ibm.com
  - Laura Flores lflores@ibm.com