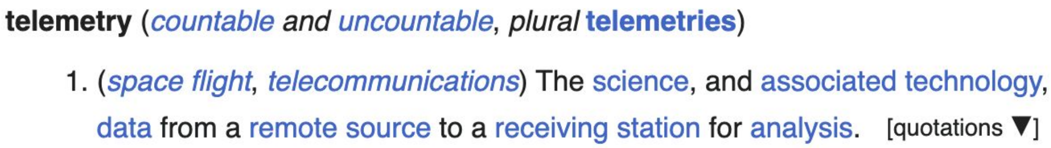
Ceph Telemetry

The Why, What, and How

Ceph Day, Silicon Valley, 2025-03-25
Yaarit Hatuka, IBM
yhatuka@ibm.com

Telemetry



telemetry (*countable* *and* *uncountable*, *plural* **telemetries**)

1. (*space flight*, *telecommunications*) The science, and associated technology, of the automatic recording and transmission of data from a remote source to a receiving station for analysis. [quotations ▼]

Wiktionary
The free dictionary

Why

- Ceph is an open source project
  - Observability into deployments in the wild
- How many clusters are out there
- Which versions they are running
- Their storage capacity distribution
- What drive models they use
- What crashes they experience
- ... and more

Why

- Developers can get feedback on what features are in use
- Version adoption rate
- See what CPU, HDD, and SSD models users are deploying
- Learn about new bugs and issues as soon as possible
- Prioritize issues
  - Focus on most common bugs
- Discover crash trends through versions
- Verify that solutions work
  - Identify regressions

Why

- Users can validate their installations by looking at “what is common”
- **No need to actively report issues or open tickets for each crash**
- Use an open data set of crashes to better understand an issue
  - Via our bug tracking system
    - tracker.ceph.com
  - See what version it is fixed in

What

- First introduced in Mimic (V13, 2019)
- Community engagement
- Steady opt-in increase trend
  - ~ 3.7K clusters
  - ~ 1.7 EB total storage capacity
  - ~ 205K OSD

What

- Clusters “phone home” to report anonymized, non-identifying data about their installation, configuration, etc.

- Data is aggregated and presented in public dashboards
  - telemetry-public.ceph.com

What

- The telemetry report is compiled daily from several channels, each with a different type of information
- Once the user is opted-in, telemetry channels can be turned on/off

What

- **Basic channel**
  - Ceph/kernel versions, cluster size, number of daemons, storage utilization, etc.
  - **On** by default

What

- **Crash channel**
  - Backtrace, where in the Ceph code the crash occurred, Ceph/kernel versions
  - **On** by default

What

- **Device channel**
  - Drive health metrics (SMART)
  - **On** by default

What

- **Ident channel**
  - Contact name, email, organization
  - **Off** by default

What

- **Perf channel**
  - Various performance counters, histograms
  - **Off** by default

How

- Telemetry / phone home capabilities are built into Ceph
  - A built-in ceph-mgr module, no extra tools are required
  - By default **home** is upstream at telemetry.ceph.com
  - If needed, a proxy can be set:
  - `ceph config set mgr mgr/telemetry/proxy http[s]://<address>:<port>`

- Clusters can be configured to phone elsewhere

- IPv6 connectivity

How

- By default telemetry reporting is **off**
- Users need to explicitly opt-in, by agreeing to CDLA-Sharing-1.0 license
  - CLI command
    - `ceph telemetry on`
  - Ceph Dashboard wizard

- Preview report with:
  - `ceph telemetry show-all`
  - `ceph telemetry preview-all`
  - `ceph telemetry show <channel>`
  - `ceph telemetry preview <channel>`

How

- Reports do not contain sensitive or identifying data like:
  - Pool name
  - Host name
  - Object name
  - Object content

How

- Anonymization
  - Cluster is assigned a 128 bit UUID
    - For telemetry
  - FSID is **not** reported
  - Disk serial ID is redacted (Device channel)
  - IPs are not stored

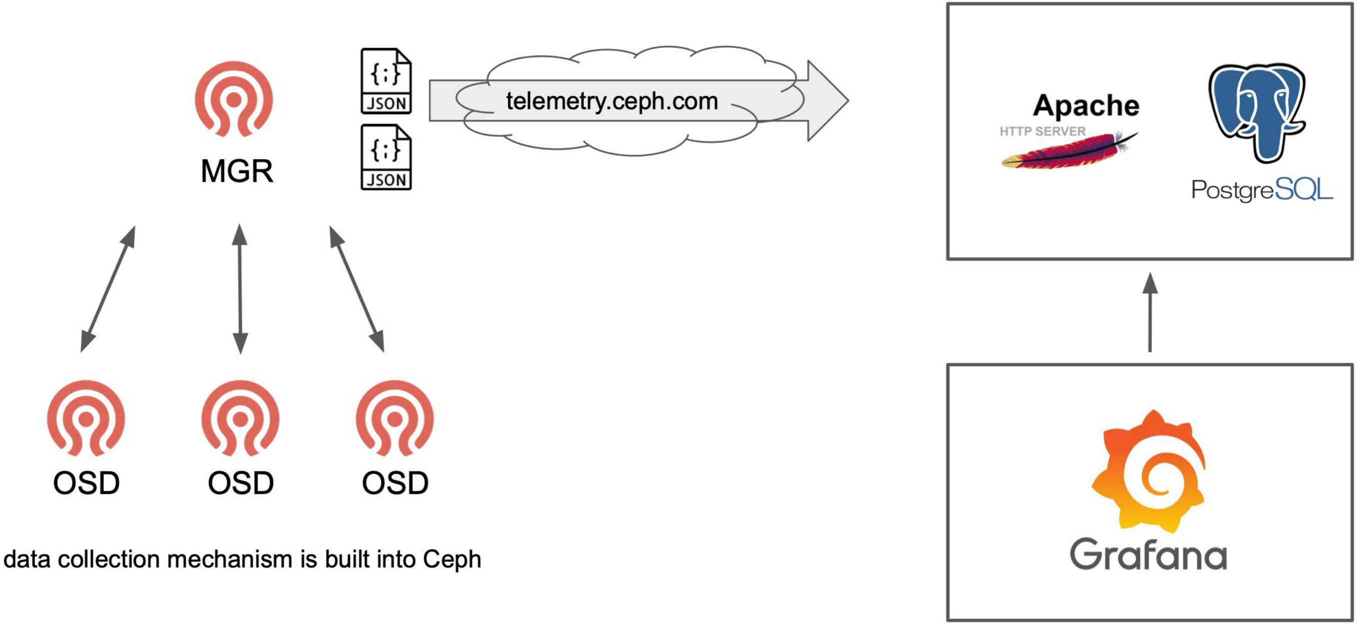
How

- Re-opt-in on Ceph upgrade
  - On any new data model version
    - `ceph telemetry diff`
  - Otherwise send current opted-in version
    - `ceph telemetry collection ls`

How

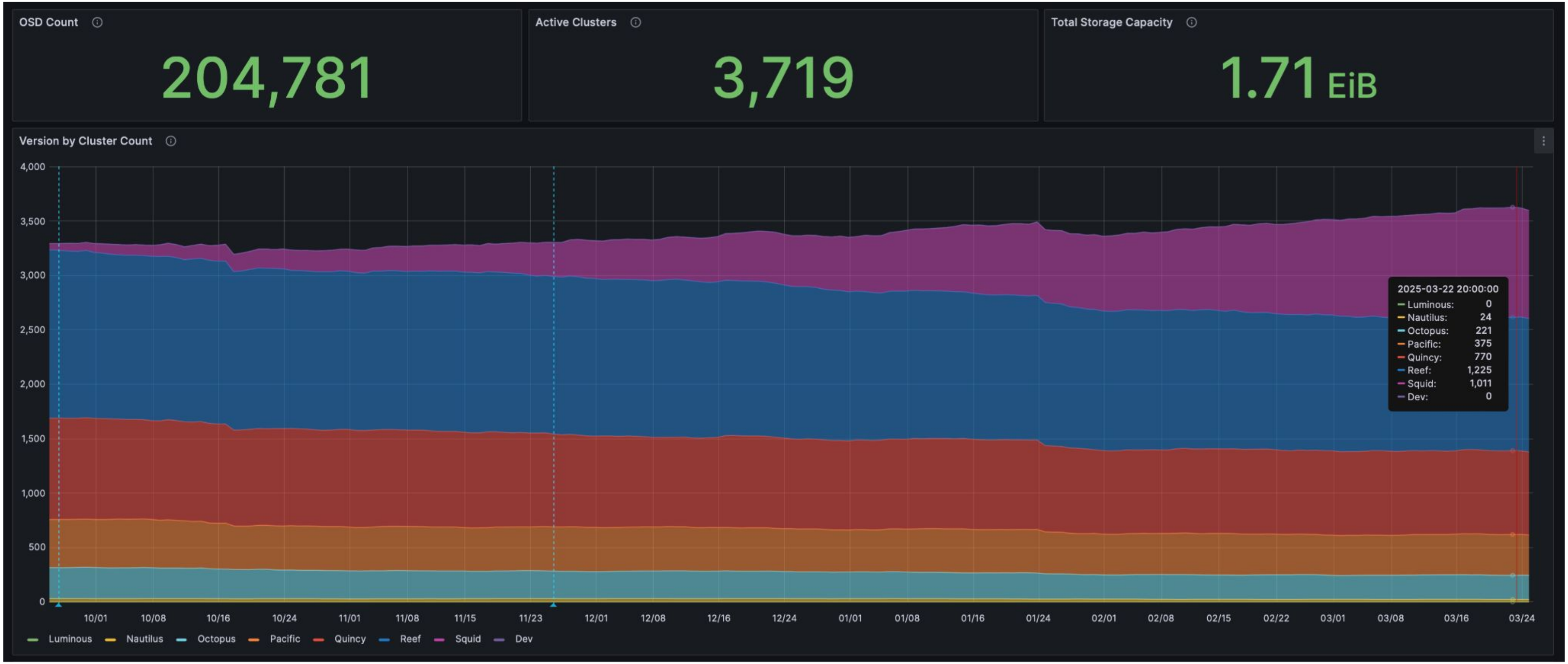
- Separate telemetry reports
  - Anonymized cluster data
  - Anonymized device health metrics
  - Sent to different endpoints

How



data collection mechanism is built into Ceph

Public Dashboards



OSD Count

204,781

Active Clusters

3,719

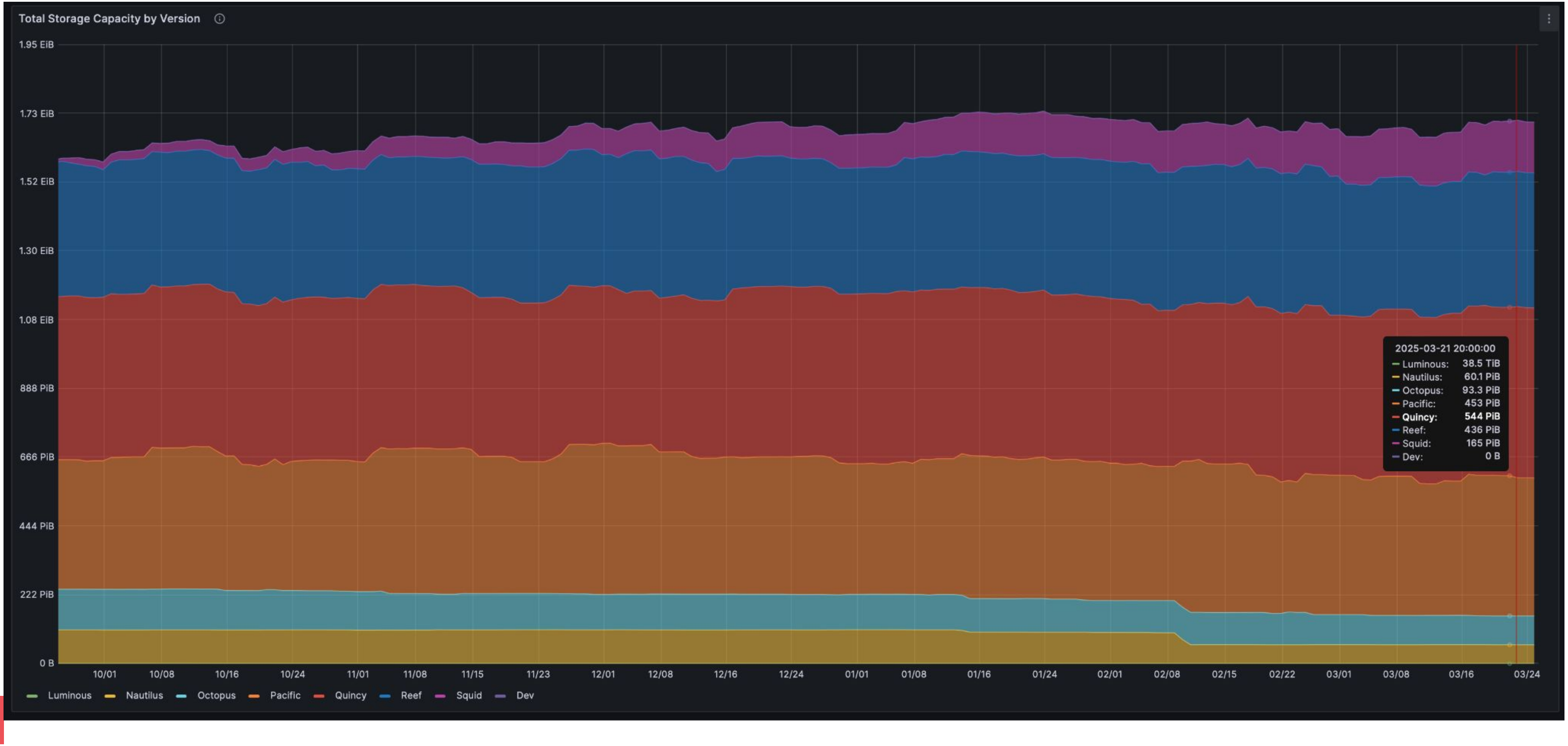
Total Storage Capacity

1.71 EiB

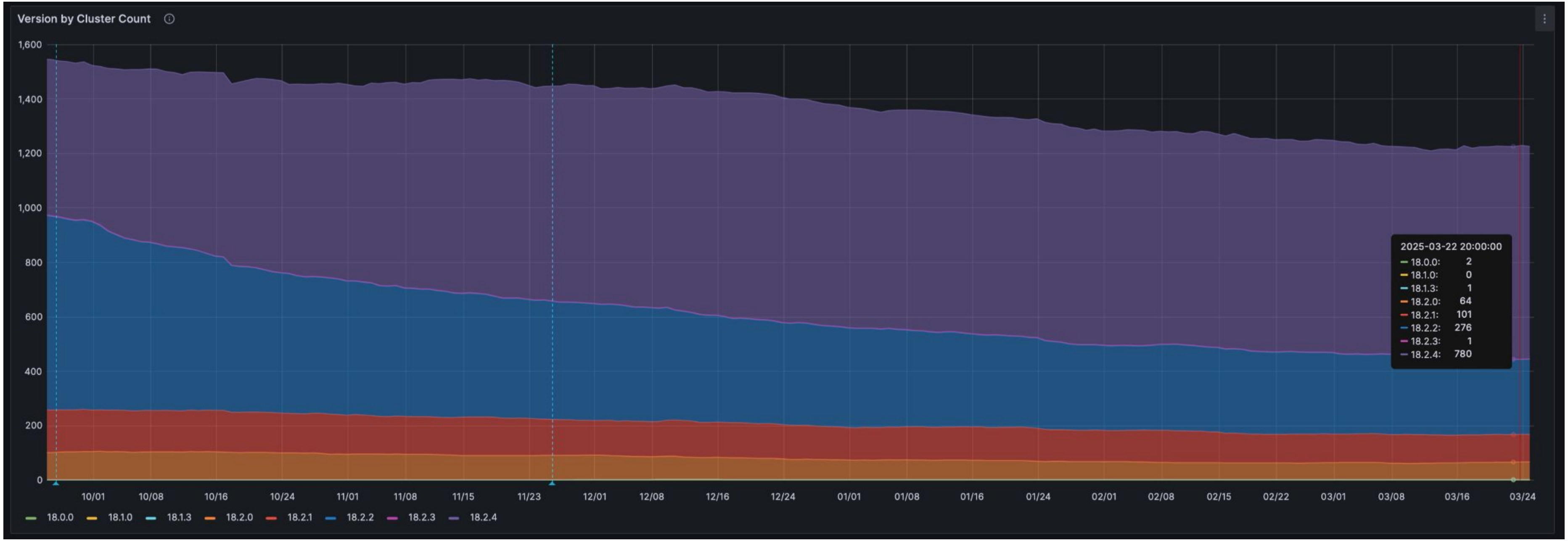
Version by Cluster Count

```
2025-03-22 20:00:00
Luminous:      0
Nautilus:     24
Octopus:     221
Pacific:     375
Quincy:      770
Reef:      1,225
Squid:     1,011
Dev:           0
```

Public Dashboards



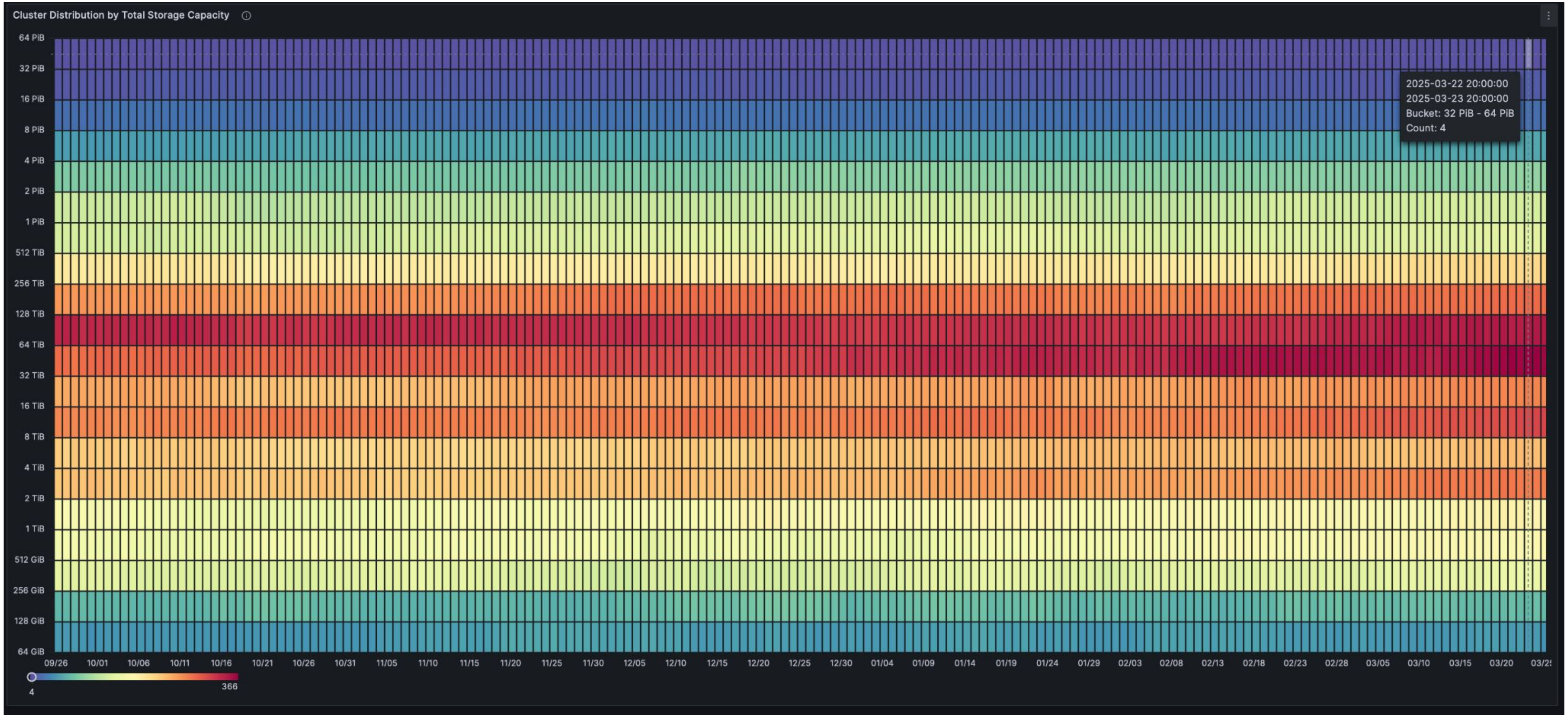
Public Dashboards



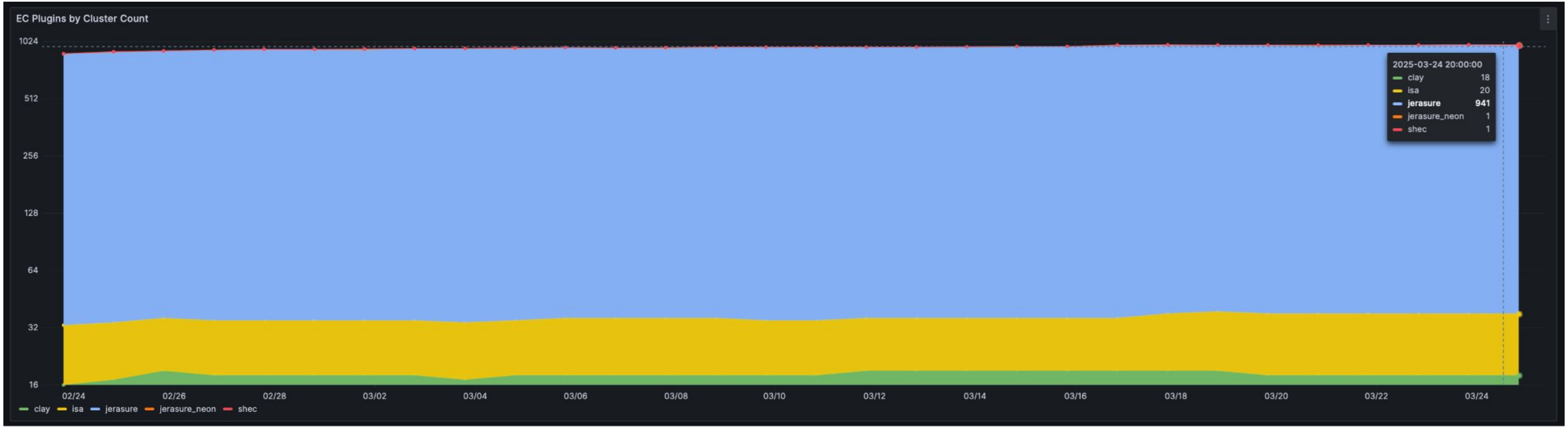
2025-03-22 20:00:00

18.0.0:	2
18.1.0:	0
18.1.3:	1
18.2.0:	64
18.2.1:	101
18.2.2:	276
18.2.3:	1
18.2.4:	780

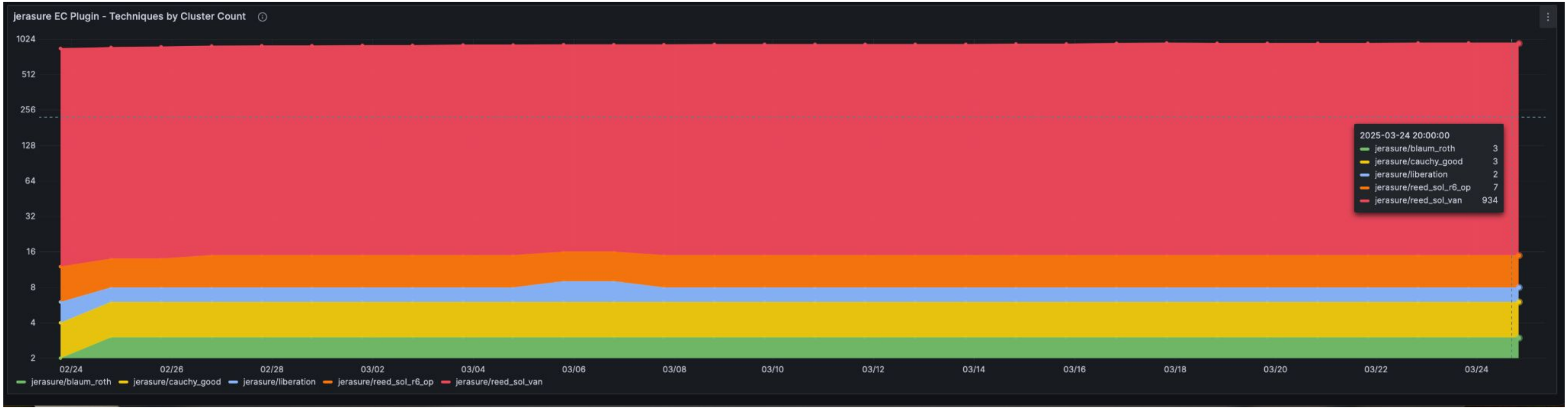
Public Dashboards



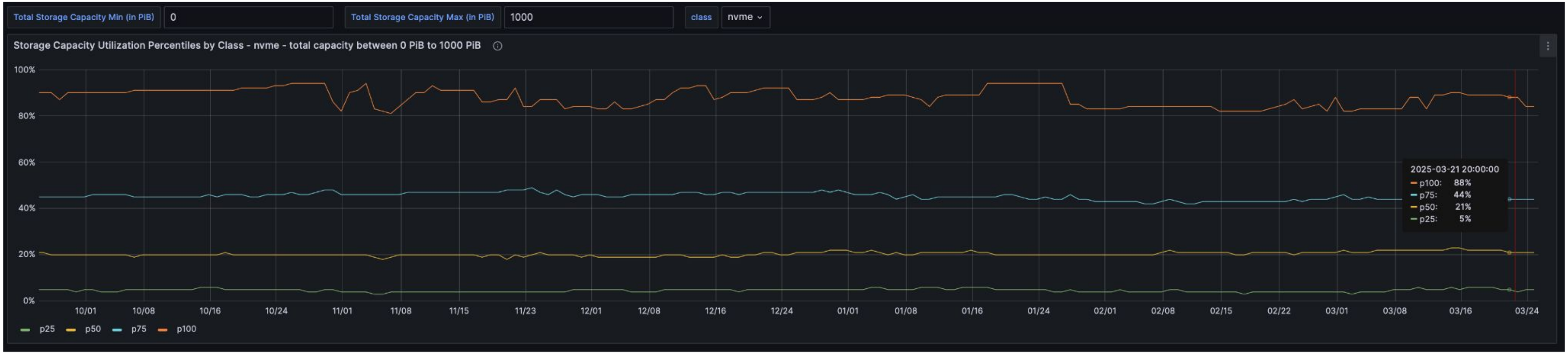
Public Dashboards



Public Dashboards



Public Dashboards



Total Storage Capacity Min (in PiB) 0

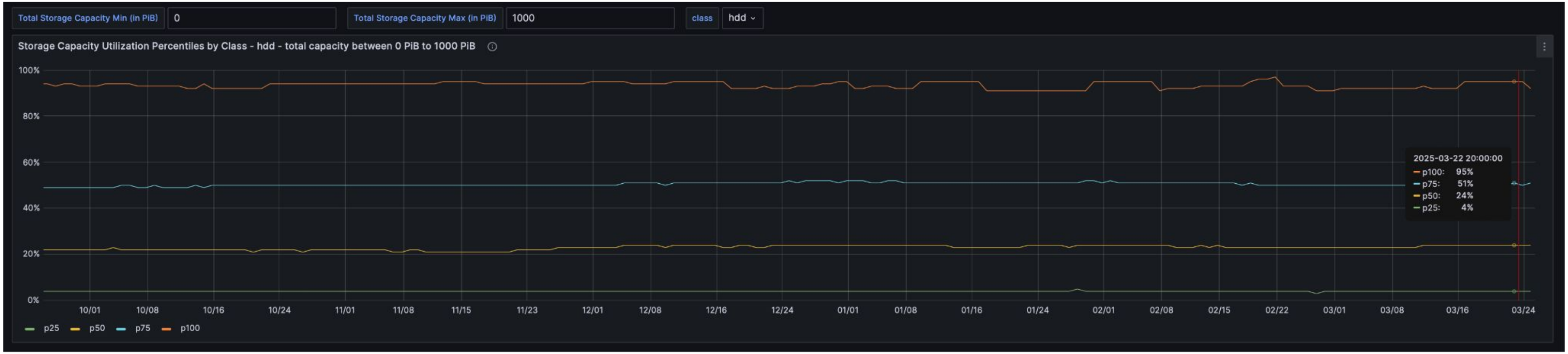
Total Storage Capacity Max (in PiB) 1000

class nvme

Storage Capacity Utilization Percentiles by Class - nvme - total capacity between 0 PiB to 1000 PiB

2025-03-21 20:00:00
p100: 88%
p75: 44%
p50: 21%
p25: 5%

Public Dashboards



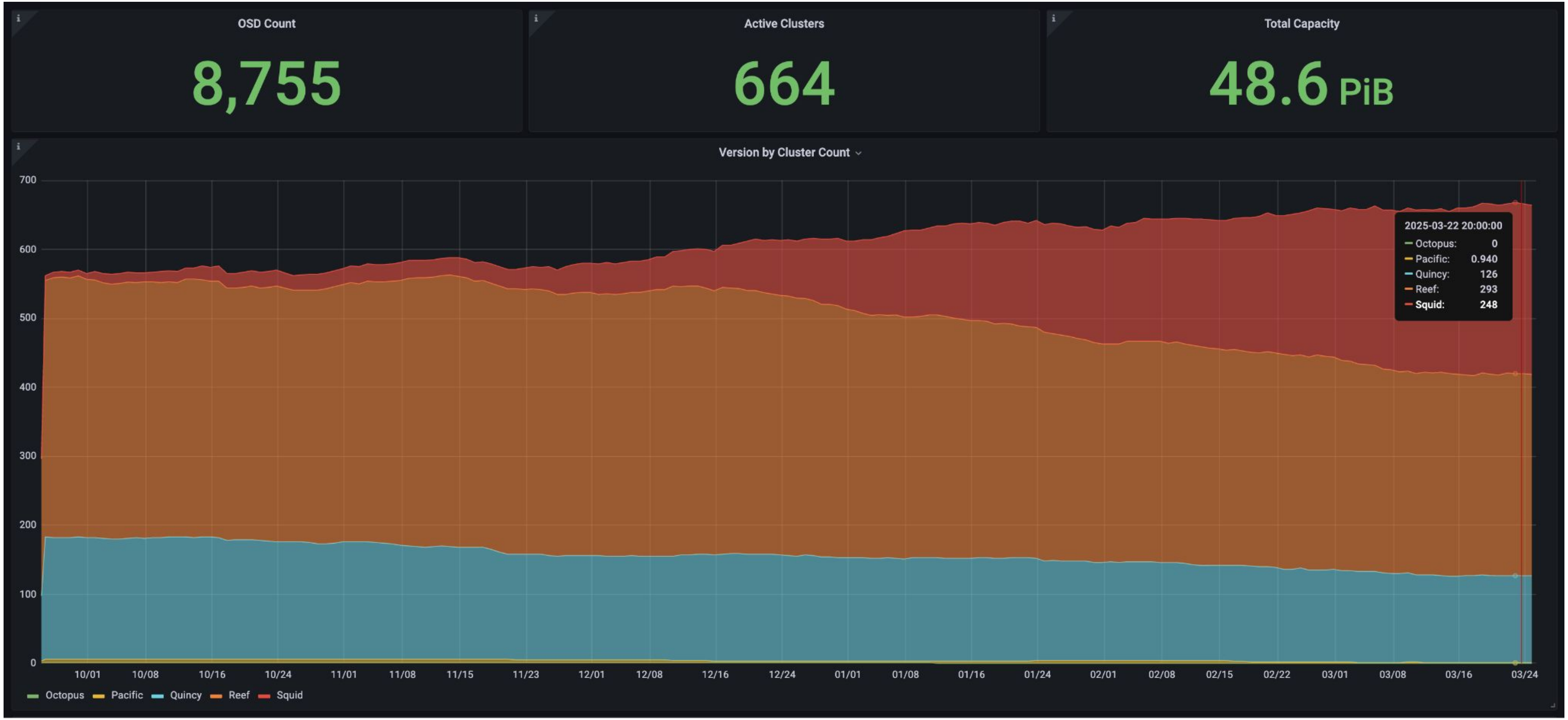
Total Storage Capacity Min (in PiB) 0

Total Storage Capacity Max (in PiB) 1000

class hdd

Storage Capacity Utilization Percentiles by Class - hdd - total capacity between 0 PiB to 1000 PiB

Public Dashboards



OSD Count

8,755

Active Clusters

664

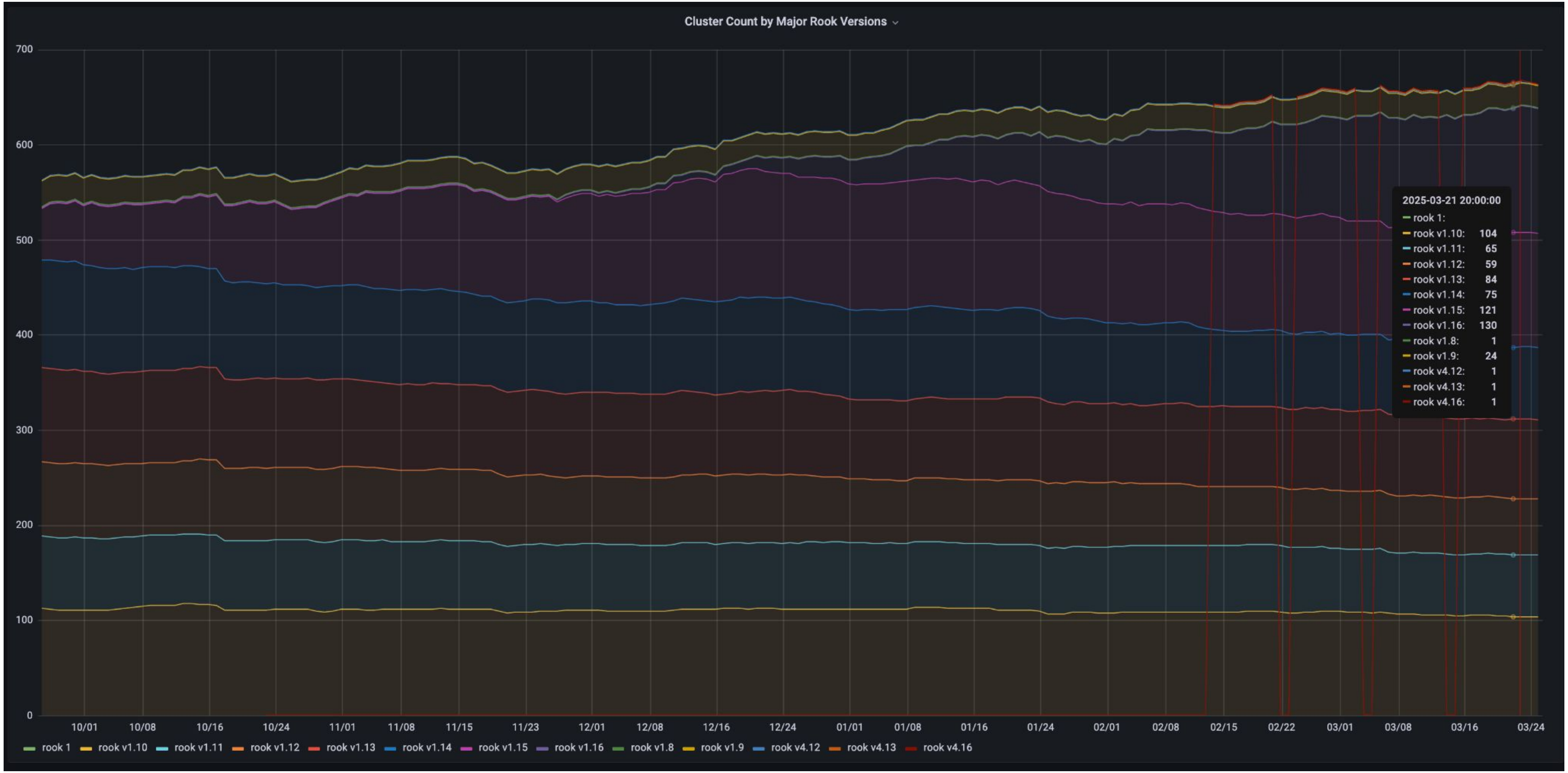
Total Capacity

48.6 PiB

Version by Cluster Count

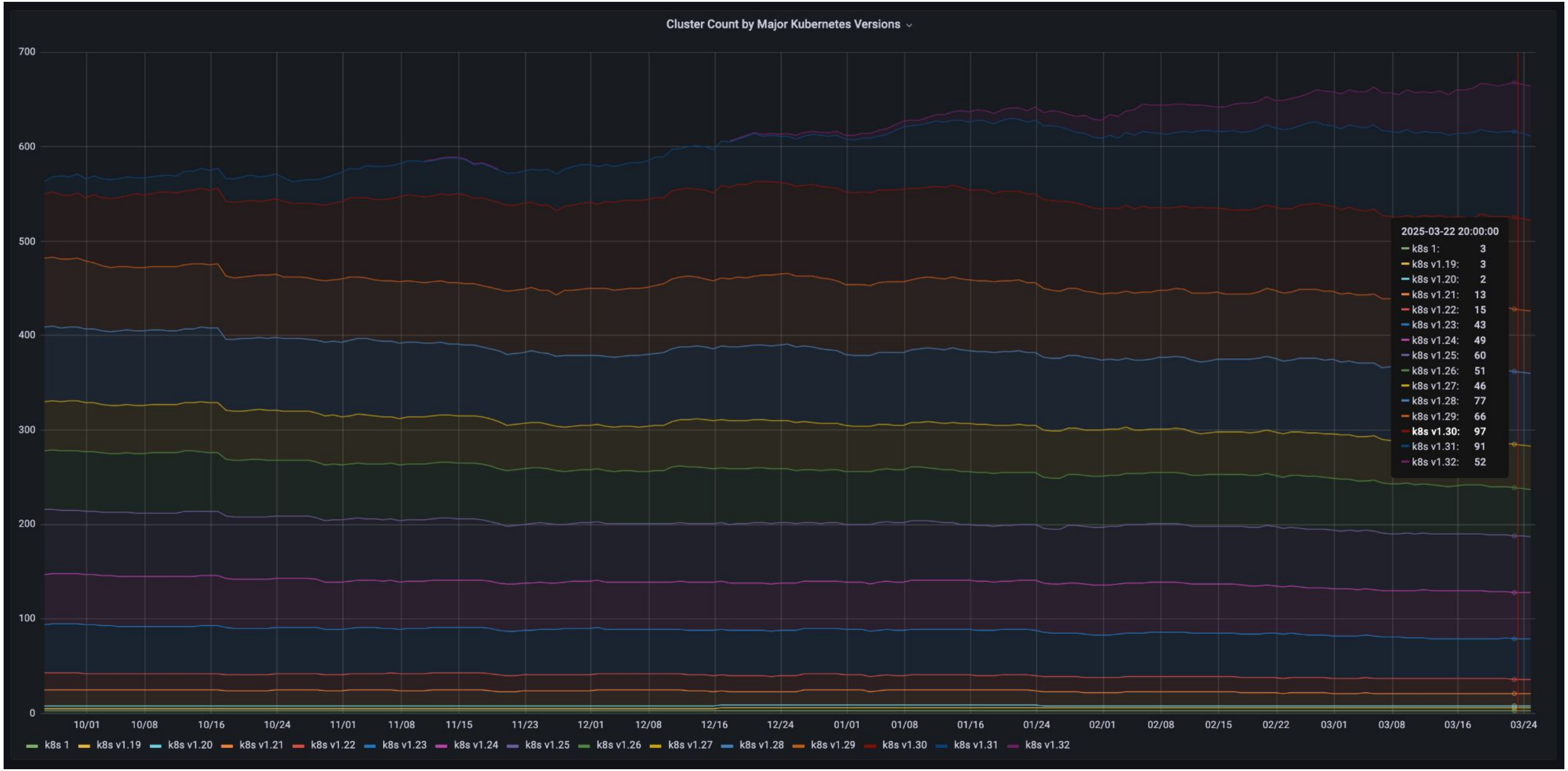
2025-03-22 20:00:00
- Octopus: 0
- Pacific: 0.940
- Quincy: 126
- Reef: 293
- Squid: 248

Public Dashboards



```
2025-03-21 20:00:00
rook 1:
rook v1.10:   104
rook v1.11:    65
rook v1.12:    59
rook v1.13:    84
rook v1.14:    75
rook v1.15:   121
rook v1.16:   130
rook v1.8:      1
rook v1.9:     24
rook v4.12:     1
rook v4.13:     1
rook v4.16:     1
```

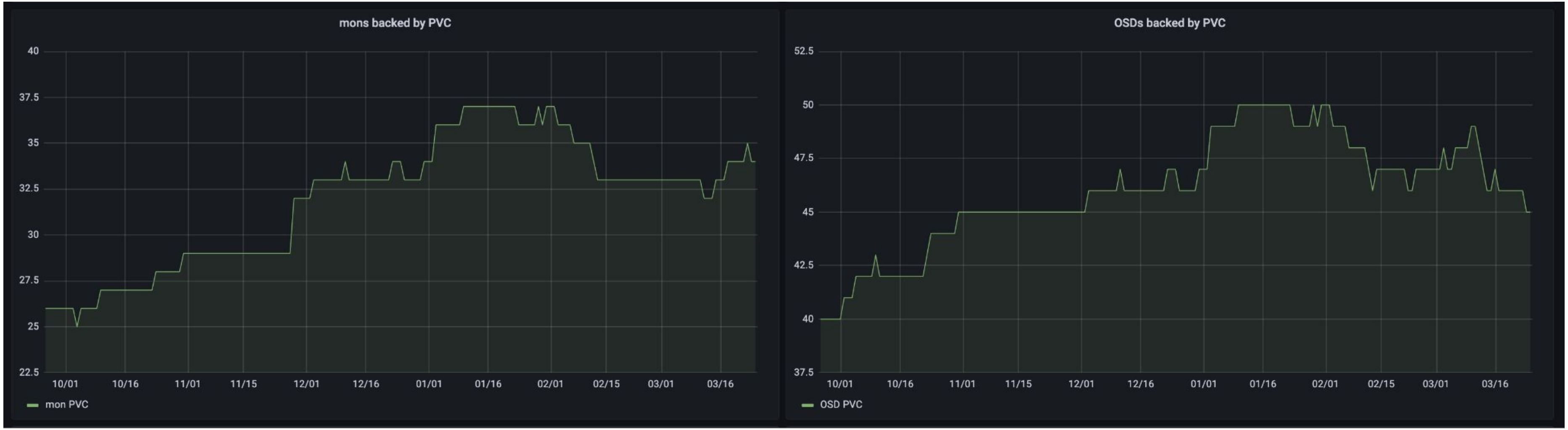
Public Dashboards



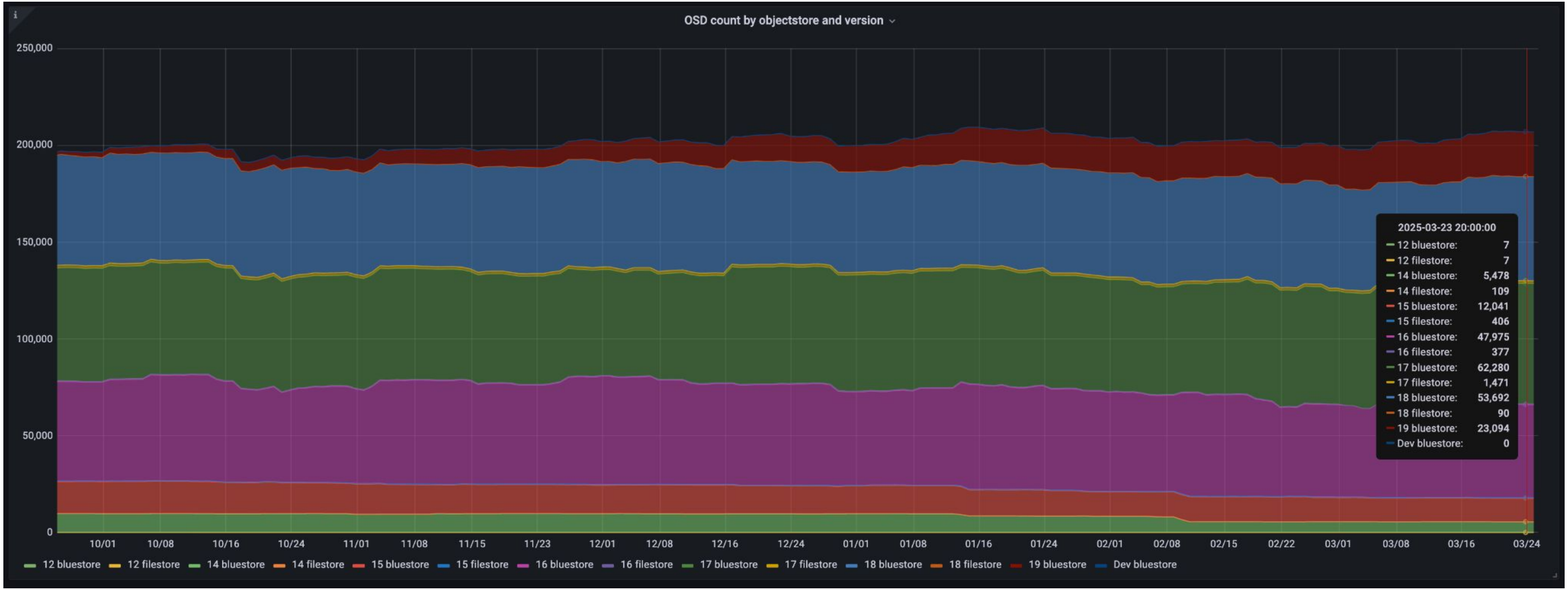
Cluster Count by Major Kubernetes Versions

2025-03-22 20:00:00	
k8s 1:	3
k8s v1.19:	3
k8s v1.20:	2
k8s v1.21:	13
k8s v1.22:	15
k8s v1.23:	43
k8s v1.24:	49
k8s v1.25:	60
k8s v1.26:	51
k8s v1.27:	46
k8s v1.28:	77
k8s v1.29:	66
k8s v1.30:	**97**
k8s v1.31:	91
k8s v1.32:	52

Public Dashboards



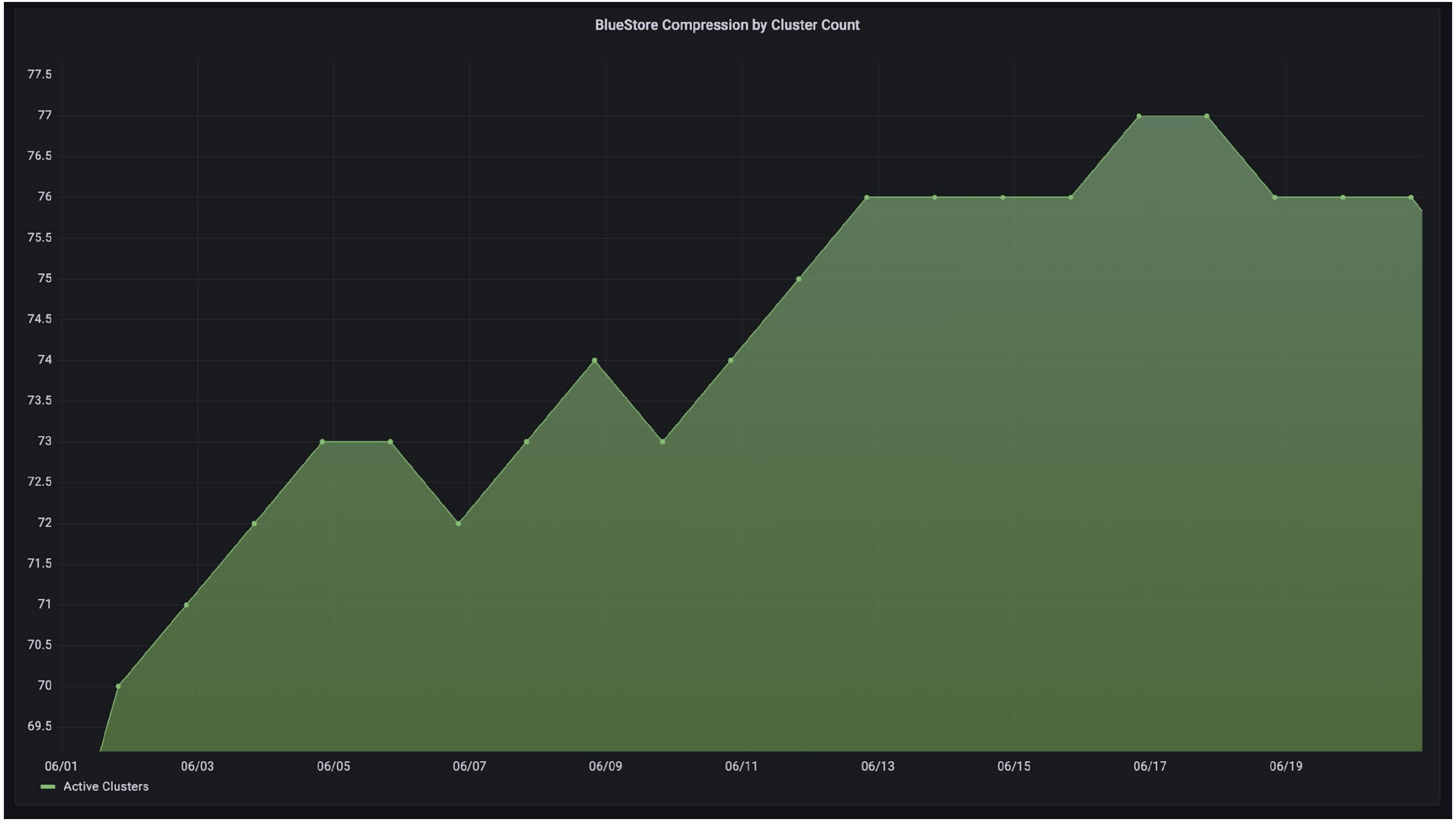
Public Dashboards



2025-03-23 20:00:00

12 bluestore:	7
12 filestore:	7
14 bluestore:	5,478
14 filestore:	109
15 bluestore:	12,041
15 filestore:	406
16 bluestore:	47,975
16 filestore:	377
17 bluestore:	62,280
17 filestore:	1,471
18 bluestore:	53,692
18 filestore:	90
19 bluestore:	23,094
Dev bluestore:	0

Public Dashboards



Crash Reports

- Crash report content
  - Crash id (timestamp + random UUID)
  - Daemon type and name
  - Ceph version
  - Backtrace
  - OS information (distribution, kernel version)
  - Assert message, condition, function, file (if applicable)

- Along with cluster information (basic channel, if enabled)
  - Available / used capacity
  - Number of daemons, their metadata, etc.

- Logs are **not** collected

Finding Similarities

- At its backend, telemetry has a **crash processor** that identifies similarities among different raw crash reports
  - Groups raw crashes into signatures (fingerprints)
  - Represent the same issue
    - Across all reporting clusters
- Supports multiple generations / recipes of signatures
  - Allows for backward compatibility
- Populates database

Tracking Issues

- A **redmine bot** queries the database for the most recent crash signatures
- Maps crash signatures to Redmine issues
  - Searches Redmine for these signatures
  - Updates an existing issue
  - Otherwise creates a new issue
    - Picks the right project
- Updates issues with essential information from crash telemetry database
  - Affected versions
  - Sanitized & raw backtrace
  - Link to a dynamic dashboard

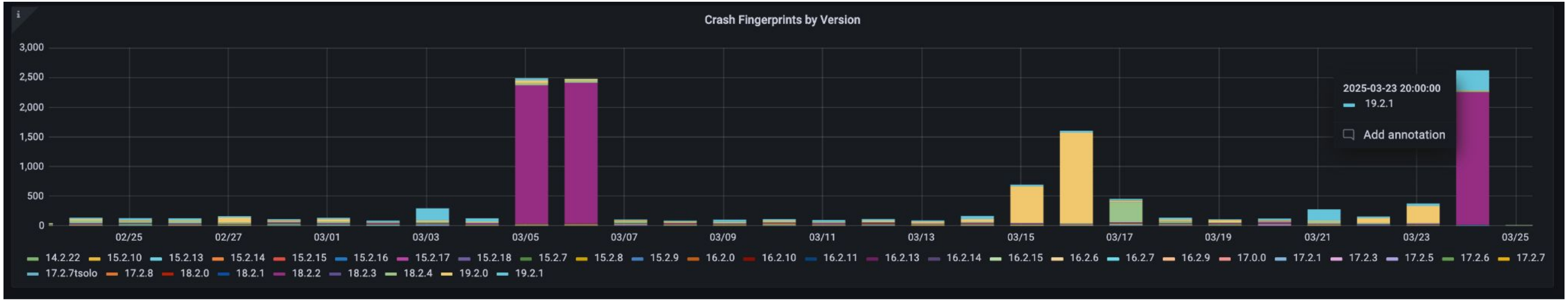
Fix Validation

- Identifies regressions
  - Opens new issues and links to the original, fixed issue
  - Emails if crash is seen in a newer point release of an older major
    - Maybe there is no need to open a new issue

Crash Insights

- A **powerful dashboard** allows searching by:
  - Backtrace frames
  - Versions (major / minor)
  - Signatures, all revisions
  - Assert function / condition
  - Number of affected clusters
  - Crash status
- Drilldowns to cluster information
  - Capacity (available / used), number of daemons, etc.
  - Current cluster version
- Shows trends in crashes through versions

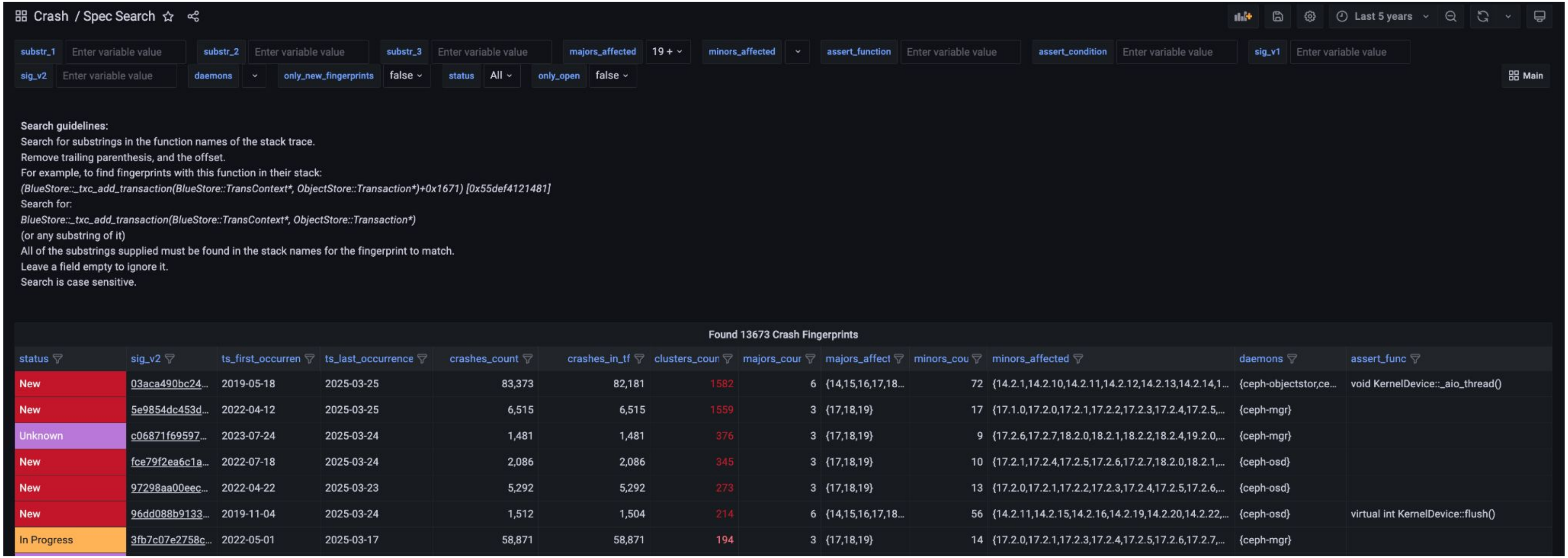
Breakdown by version



Crash Insights

- Not all crashes mean Ceph bugs!
  - Hardware, IO errors
  - Environment / resource limitation / misconfiguration
  - Issues of other dependencies

Extensive Search



Crash / Spec Search

Last 5 years

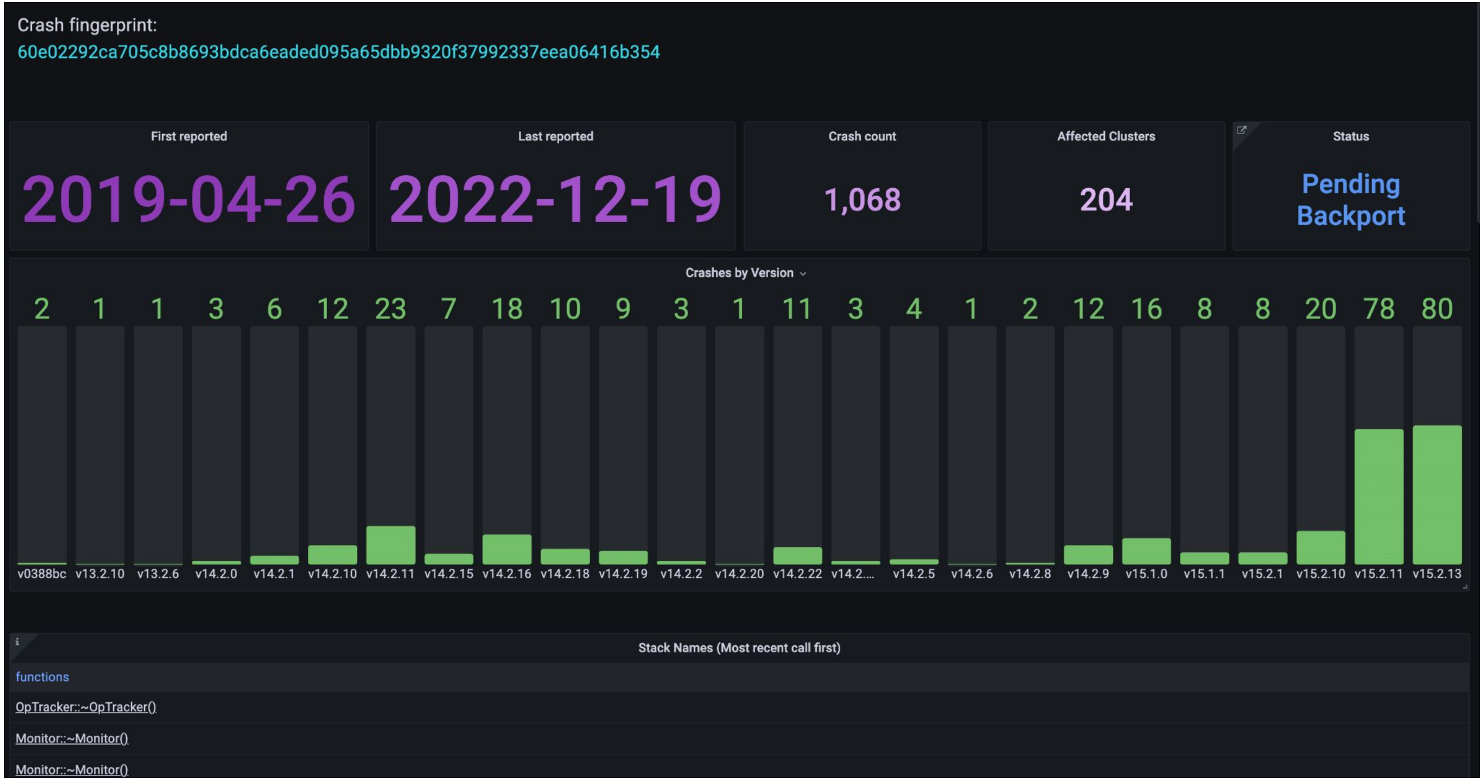
substr_1: Enter variable value
substr_2: Enter variable value
substr_3: Enter variable value
majors_affected: 19 +
minors_affected:
assert_function: Enter variable value
assert_condition: Enter variable value
sig_v1: Enter variable value
sig_v2: Enter variable value
daemons:
only_new_fingerprints: false
status: All
only_open: false

Search guidelines:
Search for substrings in the function names of the stack trace.
Remove trailing parenthesis, and the offset.
For example, to find fingerprints with this function in their stack:
(BlueStore::_txc_add_transaction(BlueStore::TransContext*, ObjectStore::Transaction*)+0x1671) [0x55def4121481]
Search for:
BlueStore::_txc_add_transaction(BlueStore::TransContext*, ObjectStore::Transaction*)
(or any substring of it)
All of the substrings supplied must be found in the stack names for the fingerprint to match.
Leave a field empty to ignore it.
Search is case sensitive.

Found 13673 Crash Fingerprints

status	sig_v2	ts_first_occurren	ts_last_occurrence	crashes_count	crashes_in_tf	clusters_coun	majors_cour	majors_affect	minors_cou	minors_affected	daemons	assert_func
New	03aca490bc24...	2019-05-18	2025-03-25	83,373	82,181	1582	6	{14,15,16,17,18...	72	{14.2.1,14.2.10,14.2.11,14.2.12,14.2.13,14.2.14,1...	{ceph-objectstor,ce...	void KernelDevice::_aio_thread()
New	5e9854dc453d...	2022-04-12	2025-03-25	6,515	6,515	1559	3	{17,18,19}	17	{17.1.0,17.2.0,17.2.1,17.2.2,17.2.3,17.2.4,17.2.5,...	{ceph-mgr}	
Unknown	c06871f69597...	2023-07-24	2025-03-24	1,481	1,481	376	3	{17,18,19}	9	{17.2.6,17.2.7,18.2.0,18.2.1,18.2.2,18.2.4,19.2.0,...	{ceph-mgr}	
New	fce79f2ea6c1a...	2022-07-18	2025-03-24	2,086	2,086	345	3	{17,18,19}	10	{17.2.1,17.2.4,17.2.5,17.2.6,17.2.7,18.2.0,18.2.1,...	{ceph-osd}	
New	97298aa00eec...	2022-04-22	2025-03-23	5,292	5,292	273	3	{17,18,19}	13	{17.2.0,17.2.1,17.2.2,17.2.3,17.2.4,17.2.5,17.2.6,...	{ceph-osd}	
New	96dd088b9133...	2019-11-04	2025-03-24	1,512	1,504	214	6	{14,15,16,17,18...	56	{14.2.11,14.2.15,14.2.16,14.2.19,14.2.20,14.2.22,...	{ceph-osd}	virtual int KernelDevice::flush()
In Progress	3fb7c07e2758c...	2022-05-01	2025-03-17	58,871	58,871	194	3	{17,18,19}	14	{17.2.0,17.2.1,17.2.3,17.2.4,17.2.5,17.2.6,17.2.7,...	{ceph-mgr}	

Crash Fingerprint



Crash fingerprint:
60e02292ca705c8b8693bdca6eaded095a65dbb9320f37992337eea06416b354

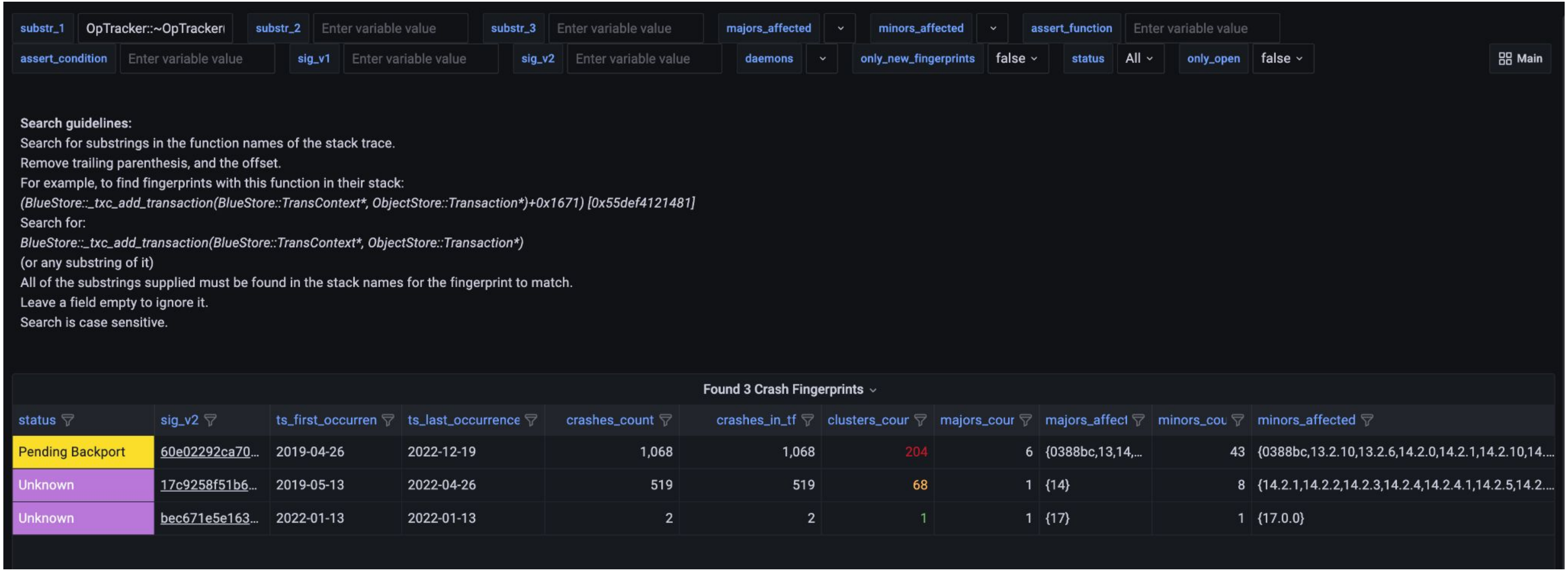
First reported	Last reported	Crash count	Affected Clusters	Status
2019-04-26	2022-12-19	1,068	204	Pending Backport

Crashes by Version

Stack Names (Most recent call first)

functions
OpTracker::~OpTracker()
Monitor::~Monitor()
Monitor::~Monitor()

Search by Frame



substr_1 OpTracker::~OpTracker substr_2 Enter variable value substr_3 Enter variable value majors_affected minors_affected assert_function Enter variable value

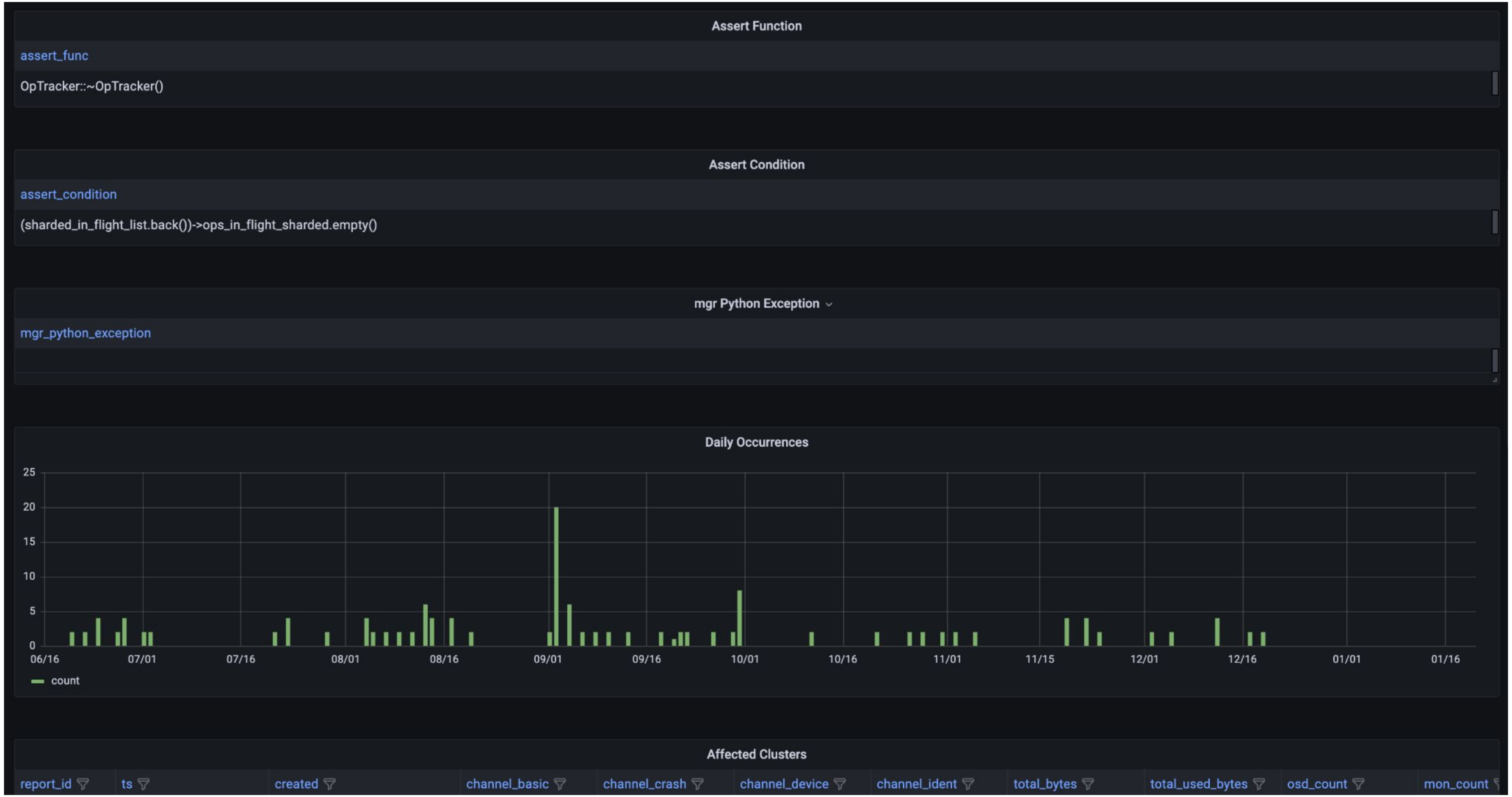
assert_condition Enter variable value sig_v1 Enter variable value sig_v2 Enter variable value daemons only_new_fingerprints false status All only_open false

Search guidelines:
Search for substrings in the function names of the stack trace.
Remove trailing parenthesis, and the offset.
For example, to find fingerprints with this function in their stack:
(BlueStore::_txc_add_transaction(BlueStore::TransContext, ObjectStore::Transaction*)+0x1671) [0x55def4121481]*
Search for:
BlueStore::_txc_add_transaction(BlueStore::TransContext, ObjectStore::Transaction*)*
(or any substring of it)
All of the substrings supplied must be found in the stack names for the fingerprint to match.
Leave a field empty to ignore it.
Search is case sensitive.

Found 3 Crash Fingerprints

status	sig_v2	ts_first_occurren	ts_last_occurrence	crashes_count	crashes_in_tf	clusters_cour	majors_cour	majors_affect	minors_cou	minors_affected
Pending Backport	60e02292ca70...	2019-04-26	2022-12-19	1,068	1,068	204	6	{0388bc,13,14,...	43	{0388bc,13.2.10,13.2.6,14.2.0,14.2.1,14.2.10,14....
Unknown	17c9258f51b6...	2019-05-13	2022-04-26	519	519	68	1	{14}	8	{14.2.1,14.2.2,14.2.3,14.2.4,14.2.4.1,14.2.5,14.2....
Unknown	bec671e5e163...	2022-01-13	2022-01-13	2	2	1	1	{17}	1	{17.0.0}

Daily Occurrences



Assert Function

assert_func

OpTracker::~OpTracker()

Assert Condition

assert_condition

(sharded_in_flight_list.back())->ops_in_flight_sharded.empty()

mgr Python Exception ⌄

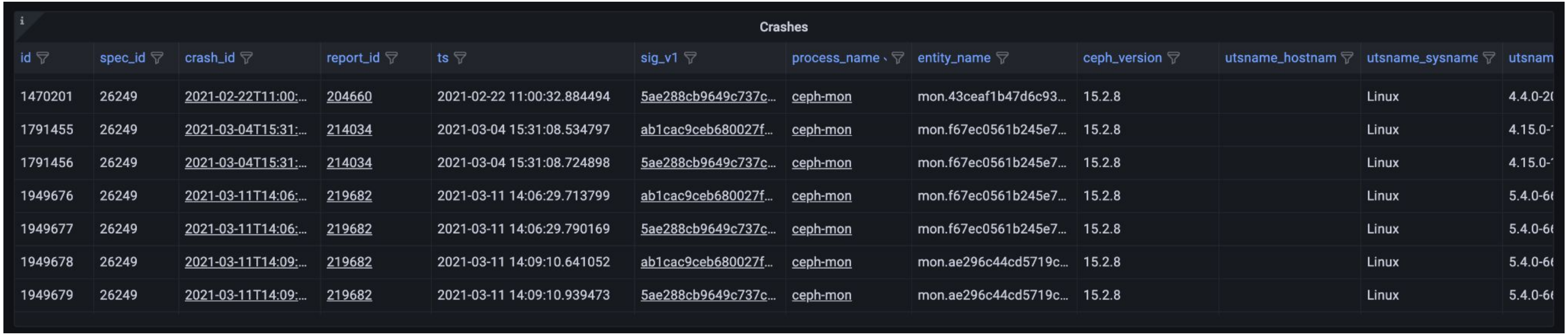
mgr_python_exception

Daily Occurrences

Affected Clusters

report_id	ts	created	channel_basic	channel_crash	channel_device	channel_ident	total_bytes	total_used_bytes	osd_count	mon_count

Crash Details



Crashes

id	spec_id	crash_id	report_id	ts	sig_v1	process_name	entity_name	ceph_version	utsname_hostnam	utsname_sysname	utsnam
1470201	26249	2021-02-22T11:00:...	204660	2021-02-22 11:00:32.884494	5ae288cb9649c737c...	ceph-mon	mon.43ceaf1b47d6c93...	15.2.8		Linux	4.4.0-2
1791455	26249	2021-03-04T15:31:...	214034	2021-03-04 15:31:08.534797	ab1cac9ceb680027f...	ceph-mon	mon.f67ec0561b245e7...	15.2.8		Linux	4.15.0-
1791456	26249	2021-03-04T15:31:...	214034	2021-03-04 15:31:08.724898	5ae288cb9649c737c...	ceph-mon	mon.f67ec0561b245e7...	15.2.8		Linux	4.15.0-
1949676	26249	2021-03-11T14:06:...	219682	2021-03-11 14:06:29.713799	ab1cac9ceb680027f...	ceph-mon	mon.f67ec0561b245e7...	15.2.8		Linux	5.4.0-6
1949677	26249	2021-03-11T14:06:...	219682	2021-03-11 14:06:29.790169	5ae288cb9649c737c...	ceph-mon	mon.f67ec0561b245e7...	15.2.8		Linux	5.4.0-6
1949678	26249	2021-03-11T14:09:...	219682	2021-03-11 14:09:10.641052	ab1cac9ceb680027f...	ceph-mon	mon.ae296c44cd5719c...	15.2.8		Linux	5.4.0-6
1949679	26249	2021-03-11T14:09:...	219682	2021-03-11 14:09:10.939473	5ae288cb9649c737c...	ceph-mon	mon.ae296c44cd5719c...	15.2.8		Linux	5.4.0-6

Crash Insights

- Monitoring of crash reports of new releases

- Bug fixes to crashes in the wild
  - https://tracker.ceph.com/issues/54653
  - https://tracker.ceph.com/issues/54742
    - Including backports!

Crash Insights

- Reported first by a developer:
  - https://tracker.ceph.com/issues/52535
  - Crash dashboard helps to estimate when it was first introduced

- Users respond to telemetry crash reports:
  - https://tracker.ceph.com/issues/56210
  - A fix from the community

- Some users identify themselves
  - Allows us to contact them for more crash related information

Public Dashboards

telemetry-public.ceph.com

```
ceph telemetry on
```