MONITORING RED HAT CEPH STORAGE
"THE EASY WAY™"

Paul Cuzner
Storage Architect
May 2018

SESSION GOALS

1. The Monitoring Challenge
2. Solution Overview
3. Demo
4. What's Next?

THE MONITORING CHALLENGE

“We” need monitoring today - RHCS 2.x and RHCS 3.x

So many ceph counters, not enough time...confusion

How to better enable Operations;

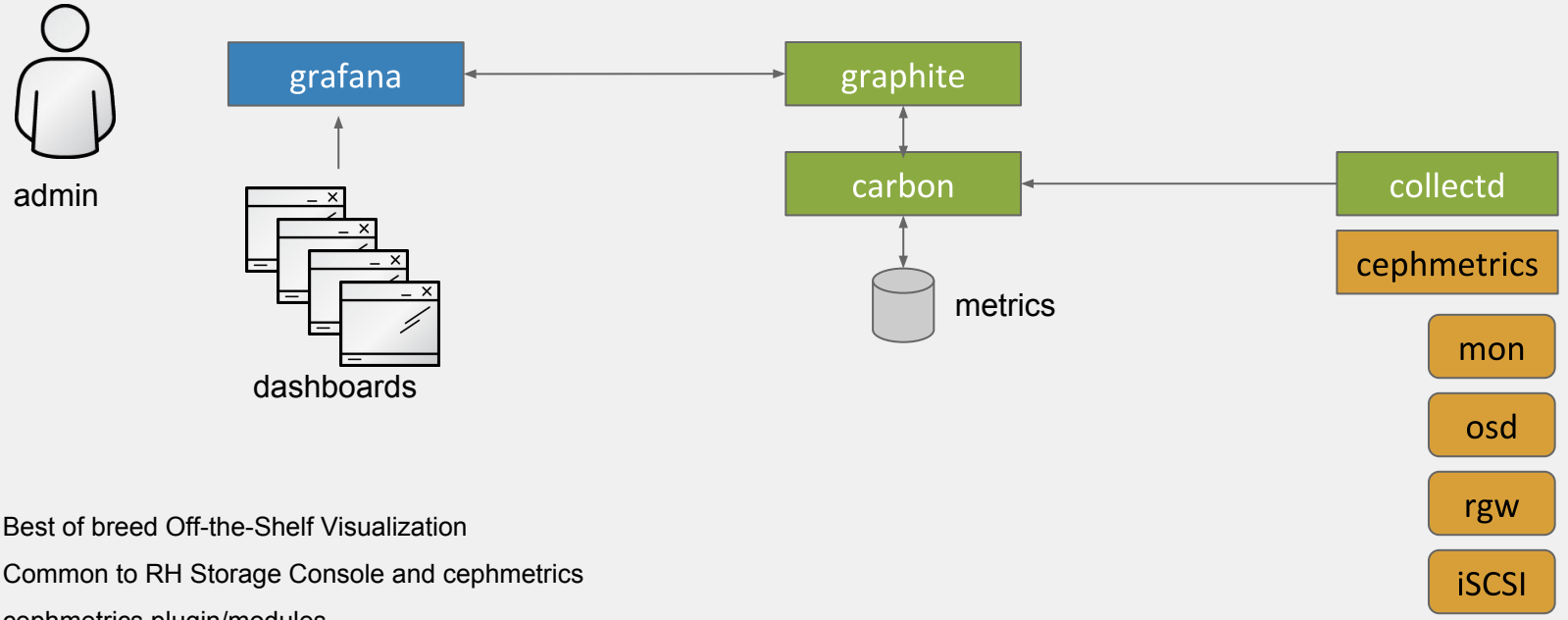
- Which metrics make the most sense to monitor
- Which metrics benefit from graphs/charts
- What are the common operation workflows

MONITORING STRATEGY

- Build on ‘best of breed’
- Use meaningful visualizations
- Combine Ceph and OS level metrics
- Support near real time analytics
- Identify key metrics
- Deliver a simple deployment with Ansible
- Flexibility to extend!

...and above all, don’t reinvent the wheel!

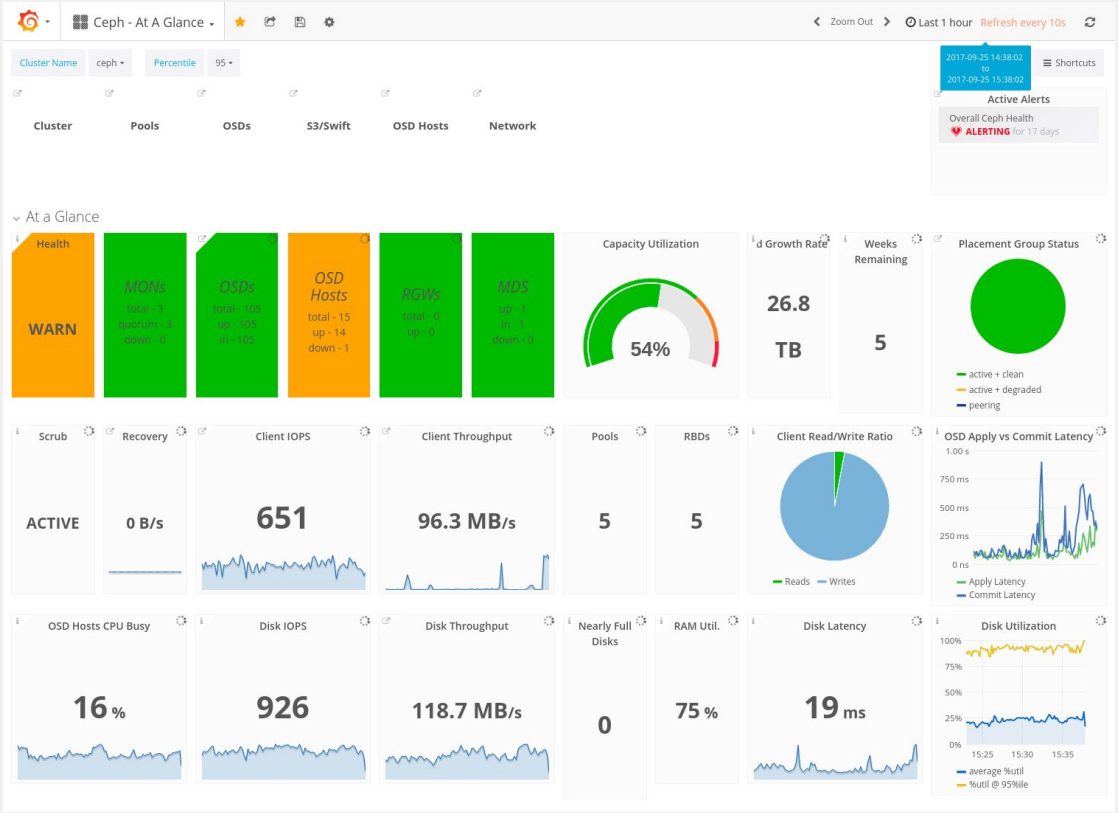
SOLUTION OVERVIEW



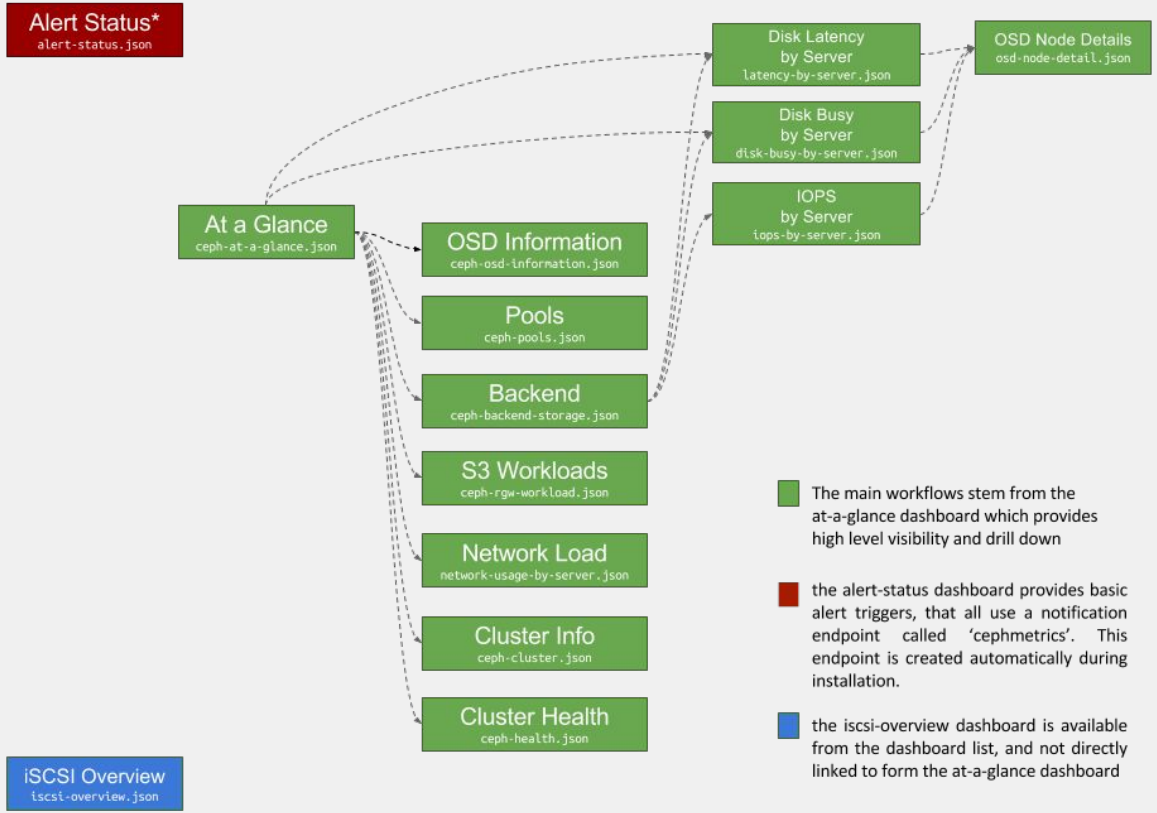
COLLECTORS

Collector Type	Data Source	Operational Visibility
mon	admin_socket perf dump ceph health ceph df ceph osd pool stats	Monitor stats and cluster health and client side performance metrics Health events with triggers
osd	admin_socket perf dump /proc/diskstats	Per OSD latencies for backing device and journals OS level latencies and performance per device with cluster level aggregation
rgw	admin_socket perf dump	RGW object operations and latency overview, by host Aggregated across all hosts
iSCSI	LIO via python rtslib_fb	Overview metrics (IOPS/Throughpout) and configuration (client count, LUNs/capacity exported) Per client IOPS/throughput, with drill-down to each LUN

THE AT-A-GLANCE DASHBOARD



DASHBOARD RELATIONSHIPS



DEMO

FEATURE SUMMARY

General Features

RHCS 2.x and RHCS 3.x
SELINUX support
Filestore and Bluestore OSD support
Encrypted and non-encrypted OSD support
MON/OSD/RGW and iSCSI roles supported
Initial support for MDS (*state ONLY*)
Drill down / dashboard links
10 second granularity
Configurable alerts
HDD/SSD/NVMe/IntelCAS support

Host Metrics support

CPU and RAM Usage
Network load

Alerts

Alerts and triggers “OOB”
Notification target automatically defined
Ceph Health Summary

Cluster Summary

OSD Configuration breakdowns
- OSD filestore vs bluestore
- OSD encrypted vs non-encrypted

Ceph version breakdown by node type
Disk Size Summary
Host Size by Capacity and disk count
PG Status breakdown
RBD and Pool counts

Cluster Details

Cluster flags status (*noout, nodown* etc)
OSD/RGW Host up/down
Per Pool capacity usage
Raw capacity utilisation
Scrub/Recovery active indicators
Growth tracking and forecast (raw)
OSDs down, near full information (host and disk)

OSD Performance

IOPS/throughput by pool
OSD performance indicators
Disk stats (per OSD)
Cluster wide disk throughput
Read/Write ratio (client IOPS)
Disk utilisation heatmap
Network load by Ceph Role

RGW

Aggregated Load View
Per Host Latencies and throughput

iSCSI

Aggregated views
- Configuration
- Performance

Per Gateway Resource Utilisation
Per client Load and configuration
Per rbd image performance

REQUIREMENTS

- A Ceph cluster running RHEL 7
- An Ansible inventory file in the same format as ceph-ansible
- A dedicated RHEL 7 host (physical or vm) to run the dashboard
- Primary requirement for the monitoring host...fast storage!
- All hosts must share a DNS domain

ROADMAP

WHAT'S NEXT?

Planned Changes

- Ceph metrics from ceph-mgr daemon
- iSCSI metrics from gateway daemons
- Data stored in prometheus
- Operational focus
- Adoption of new Grafana features
- Prometheus node exporter
- New dashboard feature embeds Grafana

Benefits

- ✓ Ceph & iSCSI metrics without 'agents'
- ✓ Reduction in monitor host requirements
- ✓ Support for containerized Ceph
- ✓ Grafana v5 simplifies deployment
- ✓ No host configuration for the node exporter
- ✓ Node exporter delivers visibility of OS metrics
- ✓ Time series graphs inside the Ceph element manager (dashboard v2)

QUESTIONS

THANK YOU


plus.google.com/+RedHat


facebook.com/redhatinc


linkedin.com/company/red-hat


twitter.com/RedHat


youtube.com/user/RedHatVideos

DEPLOYMENT OVERVIEW

- Ansible
- Uses a ‘new’ host group in ansible’s inventory
- Installs and configures collectd, grafana and graphite
- admin user enabled by default
- Dashboards loaded
- Default alerts configured
- Auto-login to the “Ceph At A Glance” dashboard

```
[ceph-grafana]
cephmetrics.front.sepia.ceph.com
```

ADDITIONAL MATERIAL

COMPARING COLLECTD PLUGINS

ceph plugin

```
<Plugin ceph>
   LongRunAvgLatency false
   ConvertSpecialMetricTypes true
   <Daemon "osd.0">
     SocketPath "/var/run/ceph/ceph-osd.0.asok"
   </Daemon>
   <Daemon "osd.1">
     SocketPath "/var/run/ceph/ceph-osd.1.asok"
   </Daemon>
   <Daemon "mon.a">
     SocketPath "/var/run/ceph/ceph-mon.ceph1.asok"
   </Daemon>
   <Daemon "mds.a">
     SocketPath "/var/run/ceph/ceph-mds.ceph1.asok"
   </Daemon>
 </Plugin>
```

cephmetrics plugin

```
<Plugin python>
   ModulePath "/usr/lib64/collectd/python-plugins"

   LogTraces true
   Import "cephmetrics"
   <Module cephmetrics>
       ClusterName "ceph"
       EventURL "http://graphite.test.lab/events/"
   </Module>
</Plugin>
```

SUPPORTED ALERT TARGETS

Grafana supports a number of notification channels (alert targets)

- Email
- Slack
- PagerDuty
- HipChat
- Telegram
- Sensu
- Generic webhooks

For more detail, look at http://docs.grafana.org/alerting/notifications/

CUSTOMIZATION INSTRUCTIONS

To make a copy of this deck for your use, go to "File" > "Make a copy" > and save to your own Google Drive.

PRESENTATION RESOURCES

For help getting started with presentations, check out the official Red Hat Presentation Guide.

NEED HELP?

If you have any questions about your session content or using the speaker portal, contact presenters@redhat.com.

CLICK TO ADD TITLE

Click to add subtitle

Insert paragraph of copy here. Do not exceed 40 words.

- Bullet
- Bullet
- Bullet