

**RED HAT  
SUMMIT**

## Beyond MVP: Prometheus installation to monitor OpenShift Cluster at Sabre

Lukasz Stanczak  
Norbert Wnuk



# Agenda:

- Sabre & Next Generation Platform
- NGP Monitoring - Metrics
  - CMO vs NGP Metrics Stack
  - Self-Serviced Grafana
  - Self-Serviced Prometheus
  - Prometheus/Thanos HA configuration
  - Prometheus/Thanos performance
  - Future work

# Sabre.... Who are we?

# Global scale and reach



Every minute of every day,  
Sabre technology powers travel.



62.5M  
API Calls  
via Dev Studio



132K  
Flights searched



5.2K  
Hotel rooms shopped



Headquarters in Dallas/Fort Worth



Global Development Centers

○ Dallas/Fort Worth, Texas

○ Krakow, Poland

○ Bangalore, India



Sabre Labs

○ Boston, Massachusetts



Sabre Red 360

SabreSonic Mobile  
Concierge

GetThere Mobile  
Booking

SynXis Property Hub

# OpenShift at Sabre

# Sabre's technology evolution

## Next Generation Platform / System services

- Logging
  - Elasticsearch, Kibana
- Monitoring
  - Prometheus, Thanos, Grafana
- Tracing
  - Jaeger
- Service Mesh
  - Istio



# OpenShift Monitoring at Sabre

# Where we come from

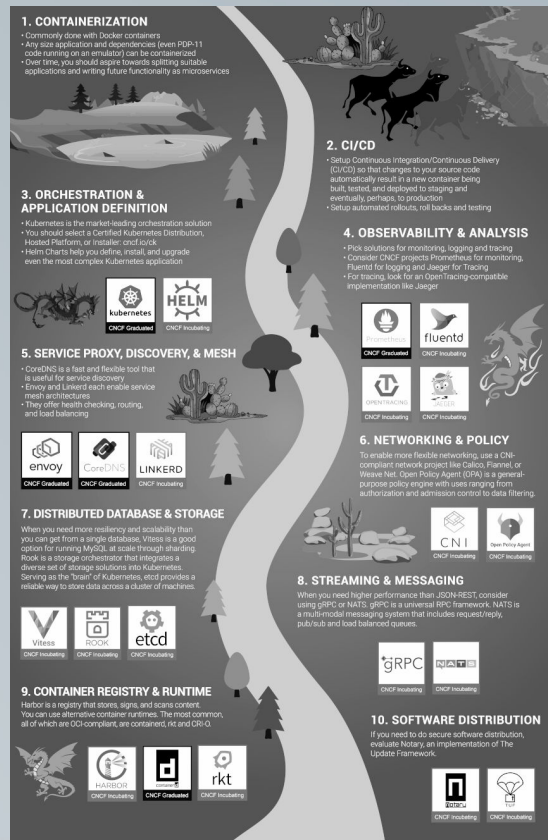
source: <https://github.com/cncf/landscape>

## NGP SaaS metrics vision

- SaaS with multi tenancy (including replacement for existing tools)
- Adhere to Open API standards (e.g. Open Metrics)
- Open Source stack & commoditization
- Aligned with K8S & CNCF & RH vision
- Fully self-serviced (K8S CRDs)
- Participate in the community (GitHub SabreOSS)

## In-house on-premise metrics

- Multiple systems, various technologies, developed over years
- Hundreds of heterogeneous applications with custom requirements
  - 100 mln incoming (sparse) data points per minute
  - 60 mln aggregated (sparse) data points per minute
  - 140 mln unique aggregated time series daily





# Cluster Monitoring Operator (CMO)

## Shipped with the Platform

- provides monitoring of the cluster components
  - ships with set of Alerts
  - ships with set of Dashboards

## Monitoring the health of the Cluster

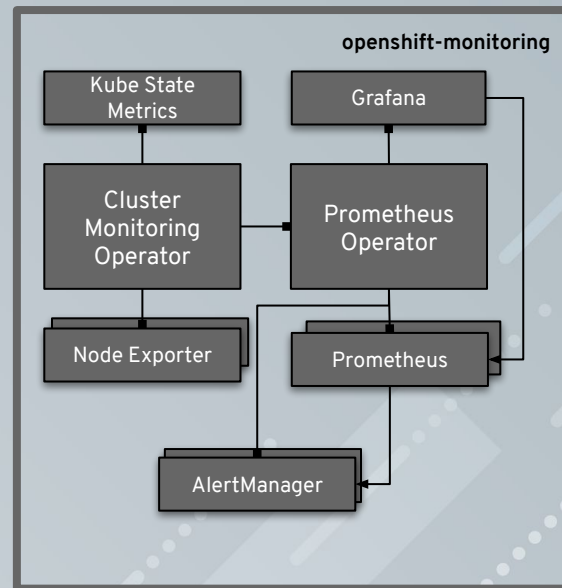
- K8S Nodes
- K8S Api server
- Kubelets

## Exporters

- node exporter
- kube-state-metrics

## What it does not provide

- application metrics
- application alerts
- application dashboards
- integration with ServiceNow



# CMO vs NGP Metrics stack

CMO installation is effectively immutable

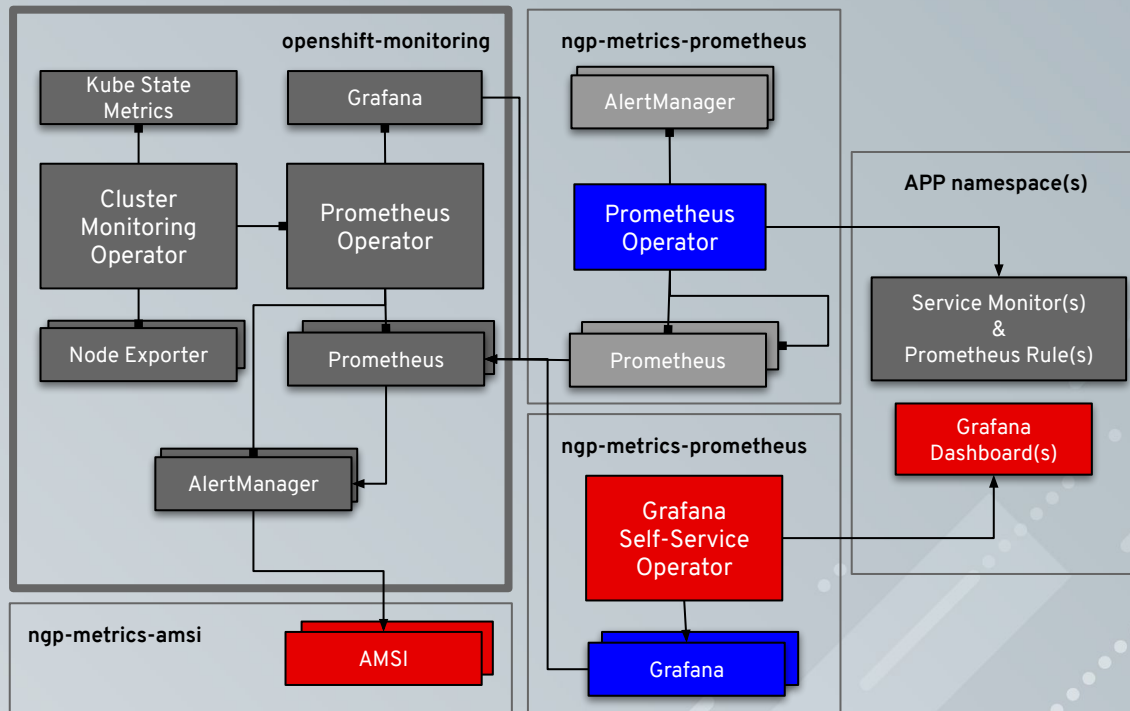
- except integration with SNOW

Custom NGP Prometheus stack for APPs:

- self - service (CRDs)
- multi tenancy (RBAC)

Sabre specific tweaks:

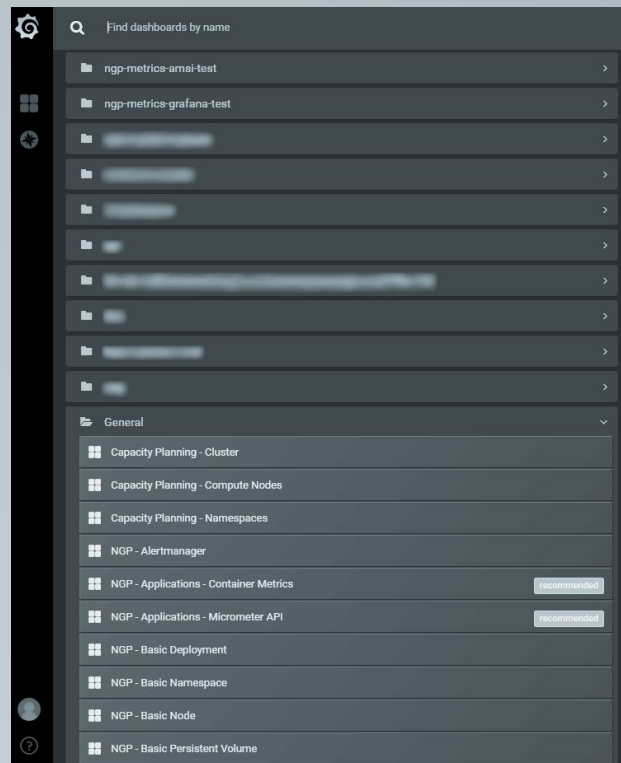
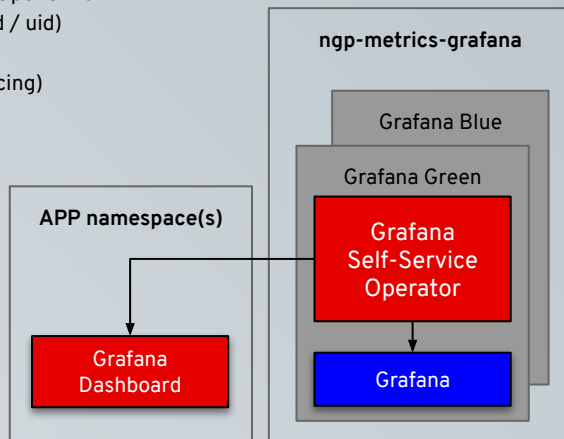
- Prometheus 2.8.1 (performance improvements)
- Customized Prometheus Operator (0.29)
- Customized Grafana image (6.1.1)



# Self-serviced Grafana

## Sabre specific tweaks:

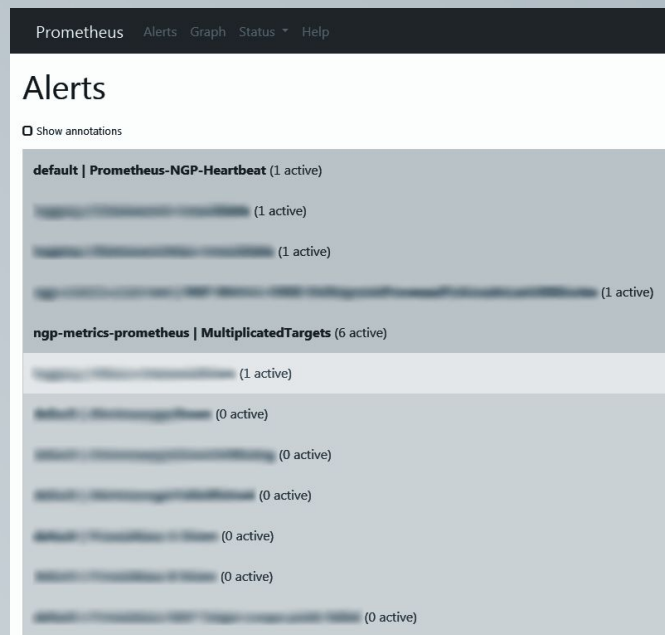
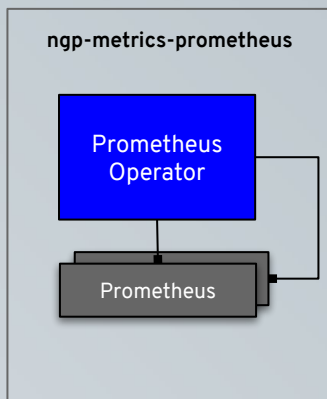
- Grafana deployment managed manually
  - stateful data (snapshots, sandbox organization)
  - shared db file in HA configuration
  - OAuth (htpasswd-file)
- Multi-tenancy based on Folders & OpenShift RBAC
  - [#12948](#), [#10339](#)
- Java based Operator
  - sync users and teams between Grafana and OpenShift
  - upload dashboards accordingly (e.g. tweak id / uid)
- Additional data sources
  - AlertManager, Elastic Search (Logging / Tracing)



# Self-serviced Prometheus

## Sabre specific tweaks:

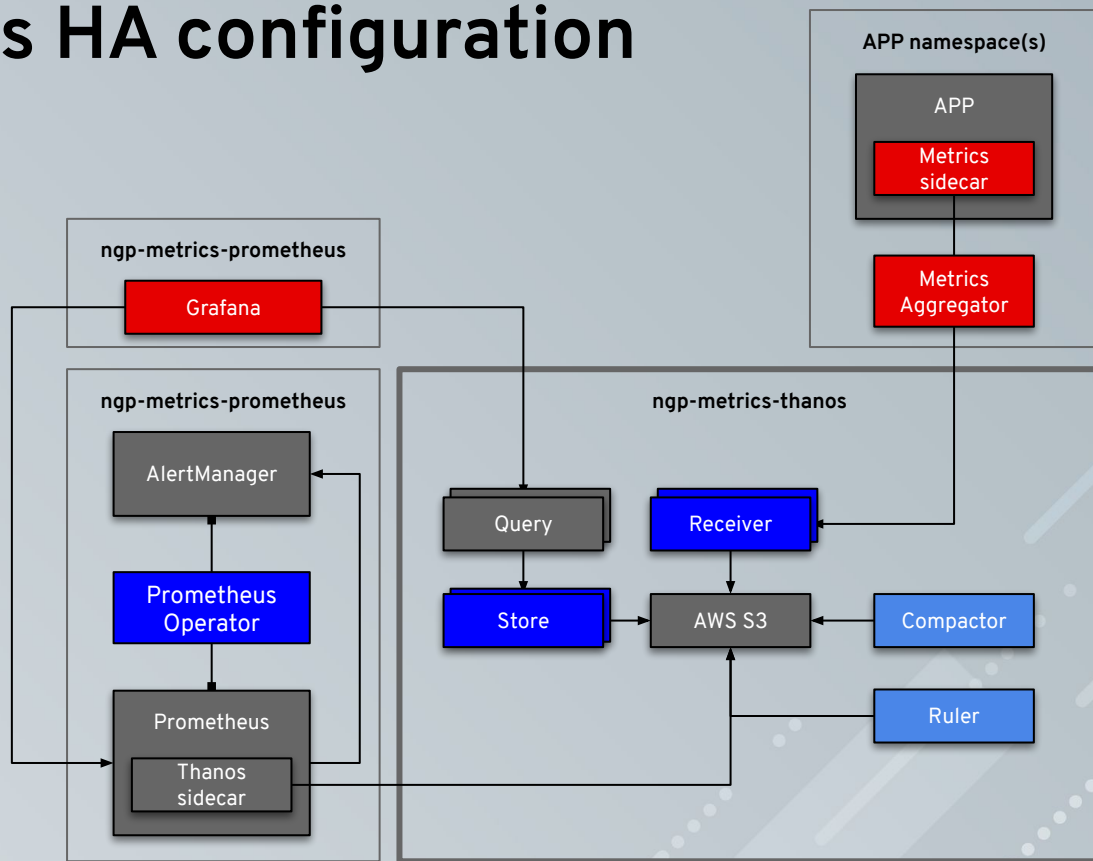
- Prometheus 2.8.1 includes all critical tweaks
  - Merge postings (tsdb): [#480](#), [#486](#), [#531](#), ...
  - Slowdown in Web UI: [#5139](#)
  - Size based storage retention
- Upgrade procedure
  - Single vs multiple operator instances
- Restore from snapshot procedure
  - PVC snapshot not vs manual procedure
- CRDs
  - Reuse CMO cluster level setup
- Custom operator image
  - CRD error handling: [#2273](#), [#380](#), ...
- Prometheus Web UI is not multi-tenant
  - Delegate alerts view to Grafana



# Prometheus/Thanos HA configuration

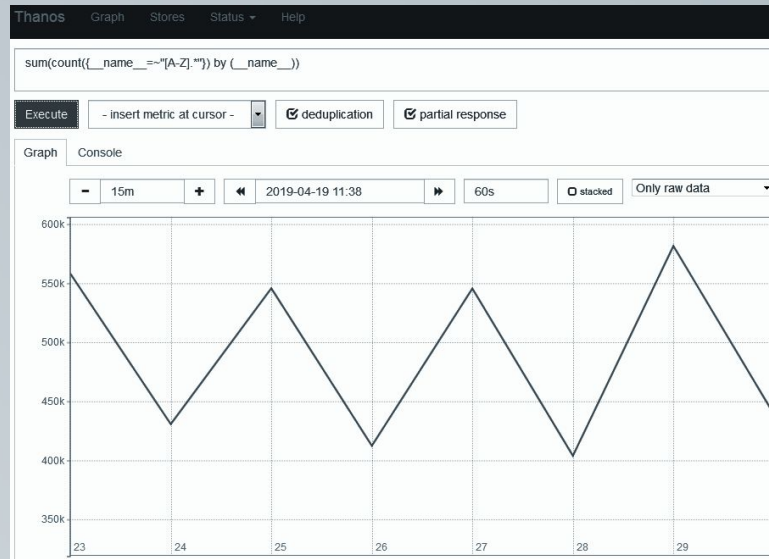
## Sabre specific tweaks:

- Based on Thanos 0.4.0 / master
- Long term storage (multiple years)
  - [#957](#)
- Multi-cluster aggregation
  - [#454](#)
- Thanos receiver (remote write) by RedHat - [#659](#)
  - Short blocks (3-5 min) for sparse metrics
  - Long duration queries (up to 10 minutes)



# Prometheus/Thanos performance

- Standard metrics
  - Prometheus (tsdb 0.6.1): 10-20M/minute with 16 CPUs / 64 GB RAM
- Sparse metrics
  - Prometheus (tsdb 0.6.1): 3M/minute with 16 CPUs / 64 GB RAM
  - Thanos receiver (tsdb 0.6.1): 3M/minute with 16 CPUs / 64 GB RAM
- 5 mln per minute metrics test
  - Queries over 24h period:
    - Combining 200 time series ~2 sec
    - Combining 1500 time series ~3sec
    - Combining 12k time series ~12sec
  - Queries over “very sparse” dataset
    - Combining 30k time series (15 min period) ~8 sec
    - Combining 2mln time series (5 min period) ~10 min
- Known limitations:
  - Thanos: [#814](#)
  - TSDB: [#561](#)



# Future work

Q3&Q4:

- Mainline Sabre specific tweaks to Operators and Tools
  - Open source Grafana operator
- Further improvements in Prometheus / Prometheus Operator
  - #1871 - Skip malformed rules when processing PrometheusRule CRDs
  - #2273 - Prometheus Operator should validate Prometheus Rule expression syntax
- Remove need to use Prometheus / Thanos Web UIs
  - Report Service Monitor status via Events in K8S
- Metrics delivery via asynchronous channels (Kafka)
- Strong(er) multi tenancy separation on TSDB level
  - Prom-label-proxy
- Data science / ML / AI for alerting and forecasting
  - AIOps: Anomaly detection with Prometheus
- Re-evaluate Cortex & M3DB

Our previous mindset was build it,  
now we're contributing to it.  
Will you join in, too?



Questions?

RED HAT  
**SUMMIT**

THANK YOU



[linkedin.com/company/Red-Hat](https://www.linkedin.com/company/Red-Hat)



[youtube.com/user/RedHatVideos](https://www.youtube.com/user/RedHatVideos)



[facebook.com/RedHatinc](https://www.facebook.com/RedHatinc)



[twitter.com/RedHat](https://twitter.com/RedHat)