



# Evaluating Ceph Deployments with Rook

Presented by: Rubab Zahra Sarfraz

Supervised by: Dan van der Ster, Julien Collet (IT-ST-FDO)

16-08-18

# What is Ceph?

Ceph is an open source **distributed storage system** designed for excellent reliability, performance and scalability.

- OpenStack Images and Volumes (RBD)
- HPC scratch spaces (CephFS)
- Private NFS-like file shares (CephFS)
- S3.CERN.CH: object storage compatible with Amazon S3 (RGW)
- lhc@home, mattermost...



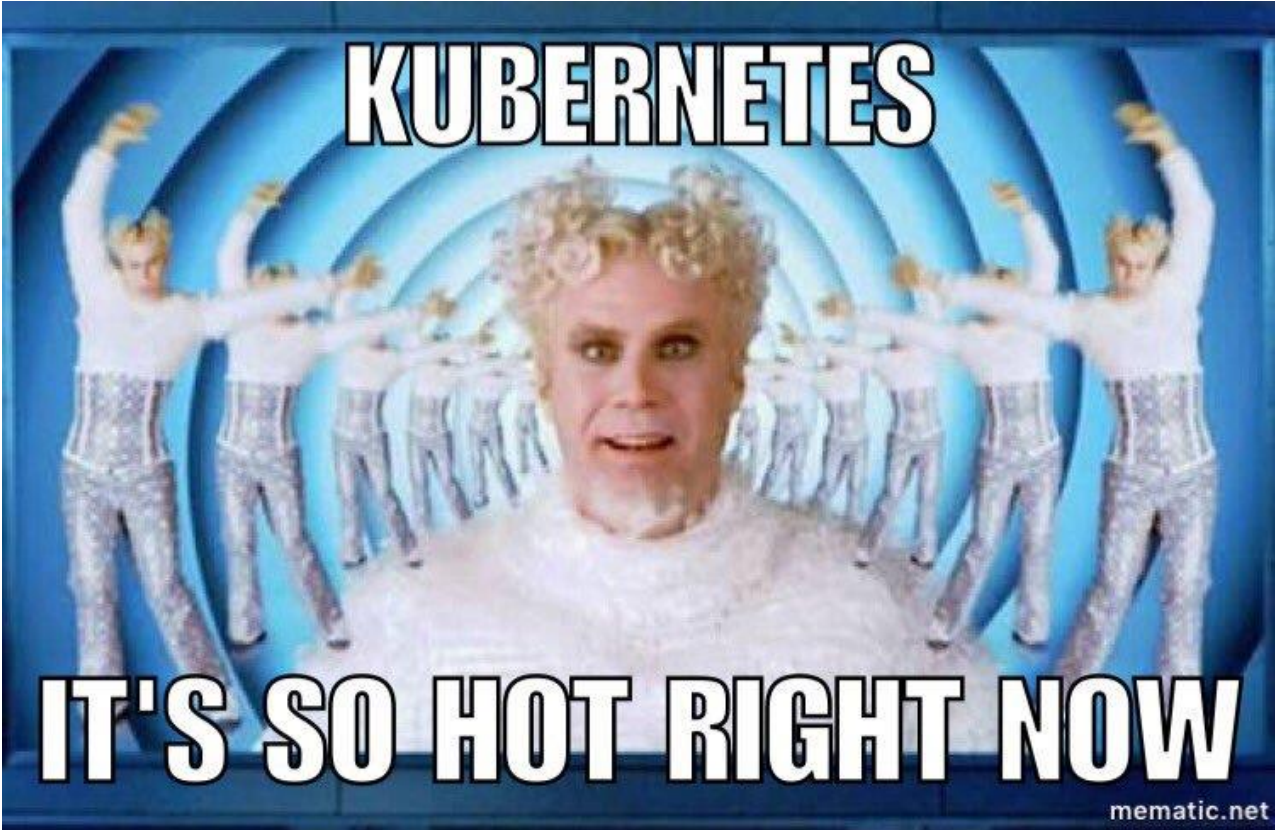


# It's not always rainbows and butterflies....

Ceph operations can be improved:

- Time consuming to deploy
- Non-trivial upgrade process
- Slow to adapt number of servers to client load
- Failure tolerance requires several standby servers (wasting resources)
- Requires expert operators

# Solution:



# Solution: ROK ❤️ Kubernetes

Rook: An **orchestrator for Ceph** that integrates with cloud native environment



# My Project

Evaluate operating Ceph clusters in Kubernetes using Rook according to CERN's storage use cases

# My Project

Evaluate operating Ceph clusters in Kubernetes using Rook according to CERN's storage use cases

Deployment



deploy Ceph clusters with  
code that is still in development stage:  
Orchestrator CLI



# My Project

Evaluate operating Ceph clusters in Kubernetes using Rook according to CERN's storage use cases

Deployment + Development

enhance code base (contribute to open source) to fit to our use-cases at CERN

deploy Ceph clusters with code that is still in development stage:  
Orchestrator CLI

# Evaluation Criteria

- Time to deploy whole cluster
- Time to add new OSDs in a running cluster
- Autoscaling: adding/removing S3/CephFS daemons
- Ceph upgrades: how much is it automated?

# Evaluation Criteria

- Time to deploy whole cluster
- Time to add new OSDs in a running cluster
- Autoscaling: adding/removing S3/CephFS daemons
- Ceph upgrades: how much is it automated?

# Evaluation Environment

## Deployed clusters for S3 workloads:

- **Virtual Environment:** 3 OpenStack VMs with 10+ volume attachments
- **Physical Environment:** 3 Physical Servers with 5.5 TB 50+ disks each

# Results and Conclusion

Metric	Puppet	Rook w/ K8s
Time to deploy whole cluster	> 3 hours	
Time to add new OSDs in a cluster	> 1 hour	
Autoscaling S3/CephFS daemons	> 1 hour	
CEPH upgrades	Manual	

# Results and Conclusion

Metric	Puppet	Rook w/ K8s
Time to deploy whole cluster	> 3 hours	< 15 minutes
Time to add new OSDs in a cluster	> 1 hour	< 2 minutes
Autoscaling S3/CephFS daemons	> 1 hour	< 2 minutes
Ceph upgrades	Manual	Manual (WIP)

# Results and Conclusion

- Elastic: can load-adapt number of S3 and CephFS servers
- Failure tolerant: replace daemons automatically without hot-standbys
- Point and click: orchestration requires much less expertise
- S3 and Automated Upgrades: still work-in-progress (*Future work!*)
- Rook + Kubernetes: fast growing community

# Results and Conclusion

- Elastic: can load-adapt number of S3 and CephFS servers
- Failure tolerant: replace daemons automatically without hot-standbys
- Point and click: orchestration requires much less expertise
- S3 and Automated Upgrades: still work-in-progress (*Future work!*)
- Rook + Kubernetes: fast growing community

**Verdict:** Spawn your next cluster with Rook already!

# Acknowledgements

- Dan van der Ster
- Julien Collet
- Theofilos Mouratidis
- Herve Rousseau
- Spyridon Trigazis
- Ricardo Brito Da Rocha
- Arne Wiebalck

# Thank you!

*rubab.zahra.sarfraz@cern.ch*



