



Exploring Kubernetes at FNAL

Anthony Tiradani

K8s-HEP Meetup ('Chicagoland')


30 January 2020

Kubernetes Motivation

- Increasing demands for hardware
 - However hardware not fully utilized
- Evolving user demands
 - Elastic Analysis Facility
- Increasing number of services demanded
 - Batch system
 - Supporting services
- External project adoption (SLATE, NERSC, LCFs, etc)
- Industry adoption

Exploring Kubernetes at FNAL – Big Data

- Apache Spark powered up with some COFFEA ☕ [*] adapted to do xrootd remote reads through custom HDFS-XrootD plugin
- Storage systems: Hadoop, MiNIO
- MongoDB, Dask
- JupyterHub with multiple Notebook flavors to choose from (Pyspark, TensorFlow, SciPy, etc.)

 **Spark Master at spark://10.131.8.4:7077**

URL: spark://10.131.8.4:7077
 REST URL: spark://10.131.8.4:6066 (cluster mode)
 Alive Workers: 15
 Cores in use: 15 Total, 0 Used
 Memory in use: 335.8 GB Total, 0.0 B Used
 Applications: 0 Running, 0 Completed
 Drivers: 0 Running, 0 Completed
 Status: ALIVE

Workers (15)

Worker Id	Address	Status	Local Resources	Used Resources
worker-20190522164819-10.131.8.8-36773	10.131.8.8:36773	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)
worker-20190522164821-10.131.8.2-41401	10.131.8.2:41401	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)
worker-20190522164821-10.131.8.3-40063	10.131.8.3:40063	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)
worker-20190522164821-10.131.8.5-34938	10.131.8.5:34938	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)
worker-20190522164821-10.131.8.6-41661	10.131.8.6:41661	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)

Name	OpSys	Arch	State	Activity	LoadAv	Mem	ActvtyTime
slot1@k8s-wn-68-69fh8.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:13:29
slot1@k8s-wn-68-mshv6.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:14:24
slot1@k8s-wn-68-n96j9.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:14:45
slot1@k8s-wn-68-ncpwb.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:13:51
slot1@k8s-wn-68-sd8kv.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:14:14
slot1@k8s-wn-test-7-cvf9v.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+06:09:39

Machines	Owner	Claimed	Unclaimed	Matched	Preempting	Drain
X86_64/LINUX	6	0	0	6	0	0
Total	6	0	0	6	0	0

[*] <https://news.fnal.gov/2019/04/cms-explores-new-technologies-for-big-data-analysis/>

Exploring Kubernetes at FNAL – Big Data



Spark Master at spark://10.131.8.4:7077

URL: spark://10.131.8.4:7077
 REST URL: spark://10.131.8.4:6066 (cluster mode)
 Alive Workers: 15
 Cores in use: 15 Total, 0 Used
 Memory in use: 335.8 GB Total, 0.0 B Used
 Applications: 0 Running, 0 Completed
 Drivers: 0 Running, 0 Completed
 Status: ALIVE

Workers (15)

Worker Id	Address	State	Cores	Memory
worker-20190522164819-10.131.8.8-36773	10.131.8.8:36773	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)
worker-20190522164821-10.131.8.2-41401	10.131.8.2:41401	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)
worker-20190522164821-10.131.8.3-40063	10.131.8.3:40063	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)
worker-20190522164821-10.131.8.5-34938	10.131.8.5:34938	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)
worker-20190522164821-10.131.8.6-41661	10.131.8.6:41661	ALIVE	1 (0 Used)	22.4 GB (0.0 B Used)

Name	OpSys	Arch	State	Activity	LoadAv	Mem	ActvtyTime
slot1@k8s-wn-68-69fh8.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:13:29
slot1@k8s-wn-68-mshv6.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:14:24
slot1@k8s-wn-68-n96j9.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:14:45
slot1@k8s-wn-68-ncpwb.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:13:51
slot1@k8s-wn-68-sd8kv.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+07:14:14
slot1@k8s-wn-test-7-cvf9v.fnal.gov	LINUX	X86_64	Unclaimed	Idle	0.000	22923	5+06:09:39

Machines	Owner	Claimed	Unclaimed	Matched	Preempting	Drain
X86_64/LINUX	6	0	0	6	0	0
Total	6	0	0	6	0	0

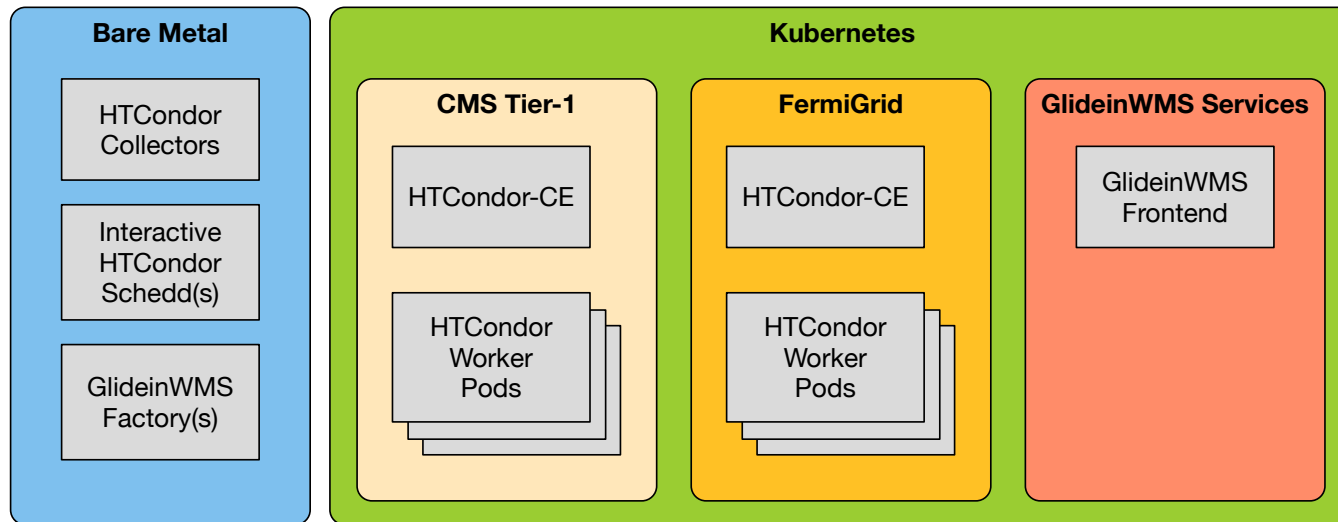


Work in progress:

- Authorization?
- User Management?
- Multi-Tenancy?
- Resource Allocation Policies?

Exploring Kubernetes at FNAL – Grid Technologies

- Collaborating with CHTC on HTCondor Docker images and Kubernetes deployment charts
- Deployment of HTCondor-CE for FermiGrid and CMS Tier-1
- Deployment of Grid Worker Nodes for FermiGrid and CMS Tier-1
- Deployment of GlideinWMS Frontend
- Hybrid Hardware for specialized workloads (GPUs, FPGAs)

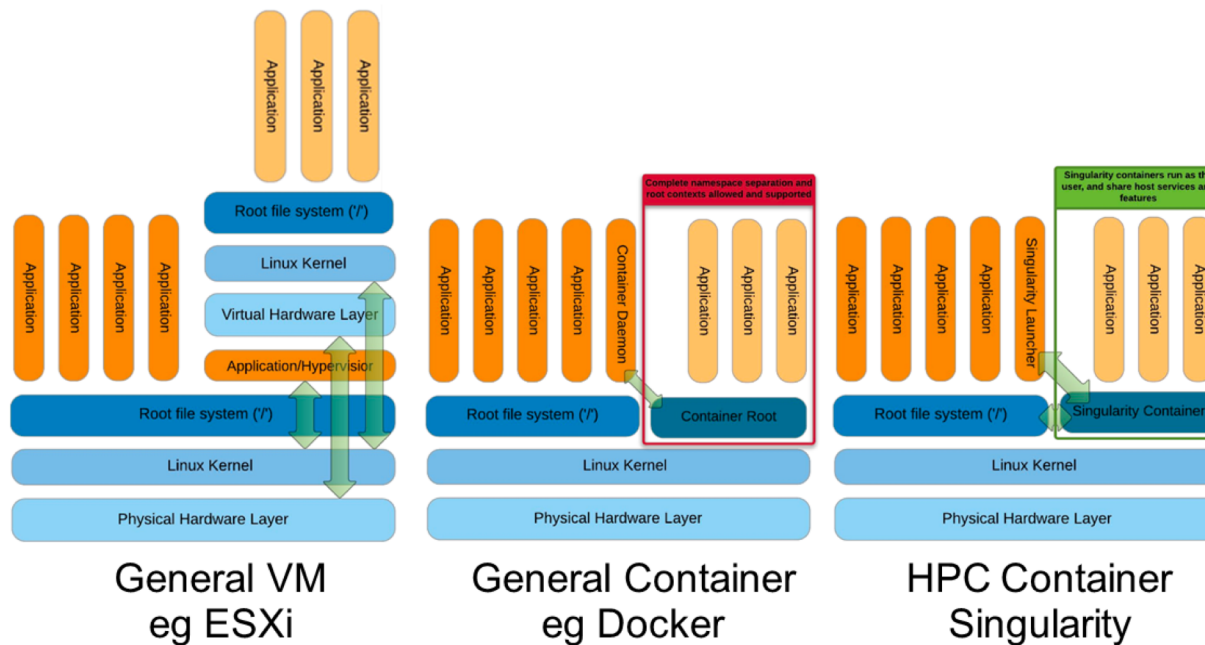


Exploring Kubernetes at FNAL – Data Management/Monitoring

- Rucio
 - Used by ATLAS and CMS
 - FNAL adding meta-data capabilities
 - Deployment planned for Kubernetes
- FTS
 - FTS3 image loosely based on SLATE image
- Grafana, Kibana and all things Landscape

Kubernetes Topics of Interest (low priority)

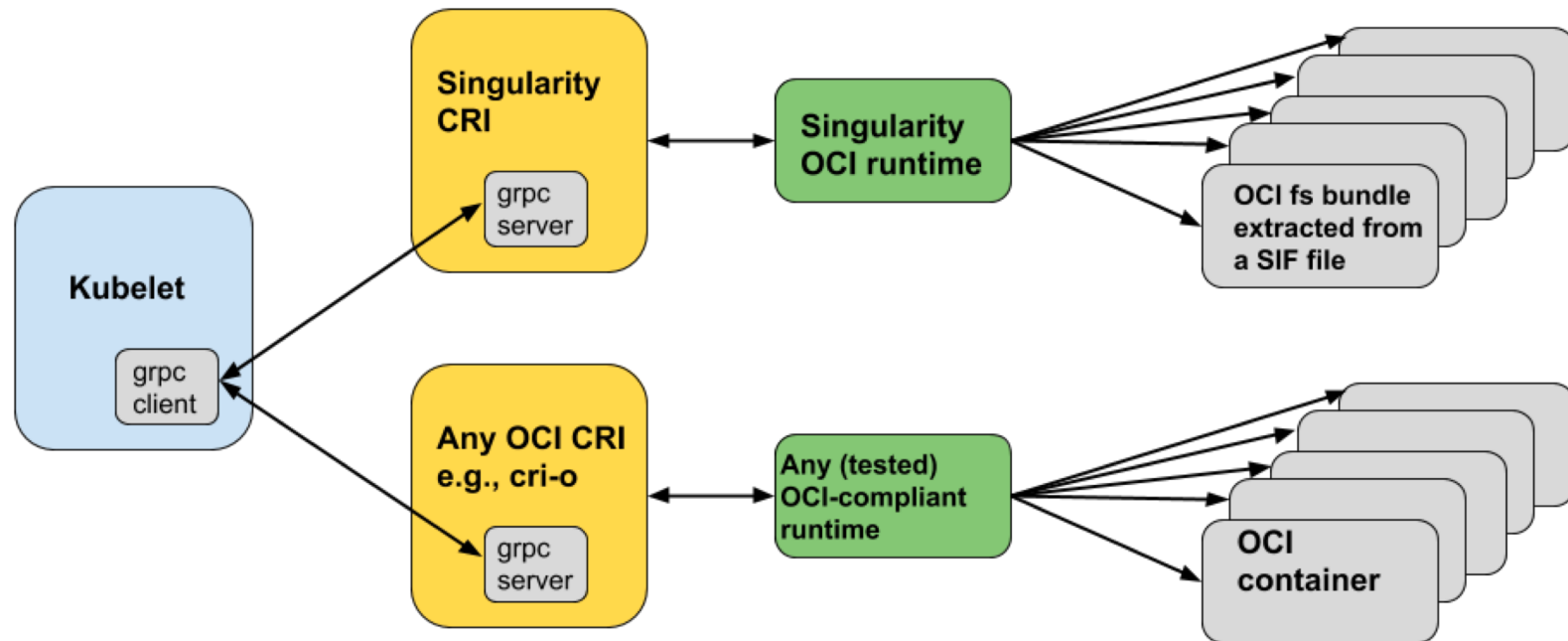
Docker Vs. Singularity



http://www.hpcadvisorycouncil.com/events/2017/stanford-workshop/pdf/GMKurtzer_Singularity_Keynote_Tuesday_02072017.pdf#43

Kubernetes Topics of Interest (low priority)

Singularity + k8s



<https://github.com/sylabs/singularity-cri>