Scalable JupyterHub deployment on bare-metal Kubernetes

Rohini Joshi <u>r.joshi@skatelescope.org</u>





JupyterHub Prototype

- Prototype running on <u>http://130.246.212.44/escape</u>
- Openstack running bare-metal Kubernetes cluster
- Helm based JupyterHub deployment
- Dynamic persistent storage with NFS
- ESCAPE IAM auth integration example <u>here</u>
- Custom user environments

 Instance Name

kubernetes-worker-3

kubernetes-worker-2

kubernetes-worker-1

kubernetes-head-01

• DEMO

		$\leftarrow \rightarrow$	G	(i) Not Se	cure ′	130.	246.212.44/es	scape/	· 🕸	0	<u>ا</u>	👼 Incognito		
<u>escape</u>		💭 Jupyter <mark>hub</mark>												
re-metal Kubernetes														
ub deployment														
orage with NFS														
egration example			Sign in with ESCAPE IAM											
0	1													
nents														
Image Name	IP Address	Flavor	Key Pair		Status		Availability Zone	Task	Power State	Age		Actions		
Ubuntu-Bionic-NoGui	192.168.50.78	c2.large	RohiniAE	NEASUIMachine	Active		ceph	None	Running	3 weeks	, 3 days	CREATE SNAPSHO	т - т	
Ubuntu-Bionic-NoGui	192.168.50.29	c3.medium	RohiniAE	NEASUIMachine	Active		ceph	None	Running	1 month	ı, 1 week	CREATE SNAPSHO	T	
Ubuntu-Bionic-NoGui	192.168.50.61	c3.medium	RohiniAE	NEASUIMachine	Active		ceph	None	Running	1 month	ı, 1 week	CREATE SNAPSHO	т 🔹	
Ubuntu-Bionic-NoGui	192.168.50.44, 130.246.212.44	c3.large	RohiniLap	ptop	Active		ceph	None	Running	1 month	ı, 1 week	CREATE SNAPSHC	T -	

× +

JupyterHub





Lessons learnt

Start small, adding a variable at a time

Persistent volumes are hard!

Magnum clusters or another cluster orchestration service will simplify things

Kubernetes

- Double check yaml file for misconfigs
- Lots of things can change between versions
- K8s documentation, blog posts are your friend

More useful monitoring strategy probably after testing the prototype

Custom user environments: Be careful of permissions of the default jovyan user

ssh tunnels used initially. Exposed port 80 needed for IAM auth to work



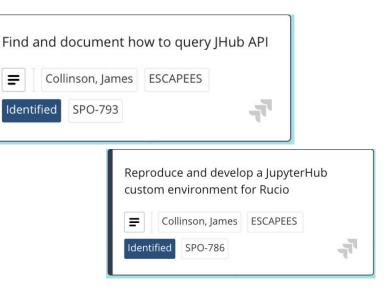


Next steps

Documenting how to query the JupyterHub REST API. <u>Gitlab issue</u>

Custom environment for Rucio to interact with the data lake from a notebook

Using auth tokens and propagation of auth tokens







Useful Links

- Small self-contained deployment <u>https://medium.com/@georgepaw/jupyterhub-with-kubernetes-on-single-bare-metal-instance-tutorial</u> <u>-67cbd5ec0b00</u>
- Zero to JupyterHub <u>https://zero-to-jupyterhub.readthedocs.io/en/latest/</u>
- Customisation Guide

https://zero-to-jupyterhub.readthedocs.io/en/latest/customizing/index.html

- Current deployment documentation https://github.com/rohinijoshi06/jupyterhub-on-k8s
- Stelios' previous tech talk <u>https://github.com/stvoutsin/escape-wp5-dev/blob/master/doc/20200608-ESCAPEWP5.pdf</u>

