

WLCG k8s WG

A. Forti

K8s

30 January 2020




Motivation

- k8s not part of WLCG infrastructure officially yet
- Interest is growing
- Several activities are ongoing in different groups
- Similar efforts but no direction
- Not a community effort yet
 - It is becoming in US because of IRIS-HEP
 - Some EU sites have k8s infrastructure, but none for the LHC experiments or public
 - CERN well into it for all its core and experiment services



BoF & pre-GDB





pre-GDB - kubernetes many faces 🔍


📅 Tuesday 10 Dec 2019, 10:30 → 19:00 Europe/Zurich
📍 31/S-028 (CERN)


Description Monthly meeting of the WLCG Grid Deployment Board See also Twiki [GDB area](#) for actions and summaries


If you plan to attend in person and require a visitor pass, please contact [lcg.office AT cern.ch](mailto:lcg.office@cern.ch) in advance of travel (please don't arrive at CERN without having arranged your pass in advance).



 Minutes


Videoconference Rooms
 pre-GDB



10:30 → 10:45 Introduction
Speaker: Alessandra Forti (University of Manchester (GB))


10:45 → 11:00 k8s at PIC
Speaker: Vanessa Acin Portella (Institut de Física d'Altes Energies)


11:00 → 11:15 CERN deployment with HELM
Speaker: Ricardo Brito Da Rocha (CERN)


11:15 → 11:30 Managing the CERN Batch System with Kubernetes
Speaker: Luis Fernandez Alvarez (CERN)
 

11:30 → 11:45 CVMFS snapshotter I
Speaker: Theodoros Tsioutsias (CERN)


11:45 → 12:00 CVMFS snapshotter II
Speaker: Simone Mosciatti (CERN)
 

- Starting point
 - BoF @CHEP
 - pre-GDB

- 1) Analysis facilities
 - * as a batch system including federated clusters and schedulers
 - * [jupiter hub/binder hub/reana/kubeflow](#)
 - * submission from the experiments
 - * easier access to alternative architectures
- 2) Service deployment
 - * large scale service deployment: [rucio](#), [cms webservices](#), [htcondor](#), other services
 - * centralised remote deployment (slate)
- 3) k8s deployment and operations
 - * on [openstack](#)
 - * on [baremetal](#)
 - * on a commercial cloud
- 4) Security
 - * k8s security in general
 - * k8s security for centralised remote deployment
 - * [AAI](#) (x509, tokens, [VO SSO](#)...)
- 5) Storage integration
- 6) Image distribution (not strictly k8s but still)
 - * registries/registries caches
 - * [cvmfs snapshotter](#)
 - * [cvmfs plain cvmfs](#)
- 7) [CNCF](#) landscape (if we want to interact with the community we need to know the landscape <https://landscape.cncf.io>)
- 8) k8s [CRI](#) and use of different [runtimes](#)



Reorganisation

- Topics that came out of the pre-GDB and BoF have been re-organised in topics that have someone working on them in some form and those that not
 - K8s as a batch system
 - Installation and deployment
 - Common infrastructure/ documentation
 - Service deployment
 - Large scale at CERN
 - Remote deployment SLATE/dodas/simple
 - Image distribution
- Couple of other topics
 - Analysis facilities
 - Security



k8s as a batch system

- k8s can do resource management very well
- Still needs a lot of development to have some of the features that we take for granted
 - Multi-tenancy and fair shares
 - Traceability techniques might have to be relearned
- Reasons to do this
 - Potential to simplify a lot some of the experiment infrastructure by using some of the native functionality
 - Spill over cloud resources seamlessly without using custom made tools, but using native functionality
 - Integrating analysis infrastructure resources



K8s as a batch system

- Making possibly smoother to submit to cloud and grid
- ATLAS + CERN-IT led effort
- Sites where ATLAS can submit CERN + UniVic
- Effort will continue to make this more robust
 - Possibly add other features like token authorization
- UniVic setup in SSL



Installation and Deployment

- Sys admins interest needs a good environment to grow
- Don't have a reference point need to create the sys admin community
 - Leverage existing presentations and expertise from people already working on it
 - CERN-IT, UniVic, Chicago, pic, DESY.....
- Ongoing discussion on how to organise the group
 - Mailing list, where to place the docs
 - Too CERN centric tools might exclude people
 - It depends on the access



Common infrastructure

- Registries and configuration tools common repos
 - Nothing new we do it also for puppet at least for non confidential code
- Still images and Helm charts are slightly more than puppet modules and might need sanitising
 - There are tools to do automated scanning and secrets can be isolated
- But the infrastructure would need to be agreed and built
 - Or we need to agree on a space on public repositories and how to maintain them and protect them
- CERN-IT is working on sharing their resources
 - Need to organise community sharing around this

Not only k8s

Not only k8s

Image distribution & CVMFS

- 2 containerd solutions to use CVMFS and avoid download everything from a registry
 - Particularly for users aim is to be able to use CVMFS for common layers and get only the user layer from the registry
 - Need to converge and cooperate on a common solution
- This is a long standing problem also for other types of container runtimes
 - Singularity also has different solutions being implemented either to use cvmfs or squids in front of a registry
- Benefits from common work.



Not only k8s
Not only k8s

Centralised installation of services

- Depending on the model
 - Hardware owned by project managed remotely
 - Hardware owned locally needs access from external project
- Simplifies installation and maintenance of complicated services
 - Local people might need knowledge of k8s but little else
- Raises a lot of questions about security and trust model
 - There is already a WLCG WG about this started by SLATE
 - SLATE not the only one all projects that do install services at sites should participate
 - WG also dominated by US sites European sites need to participate too
 - WG charter



Analysis Facilities

- Two types of services mostly required
 - Local batch system
 - Jupiter hubs
- Jupiter hubs handled by k8s can be also seen as an alternative to more classical batch system k8s still queues jobs even if the hub is interactive
- A more futuristic vision is to have federated jupiter hubs accessible using a federated identity
 - Components to do this are already there
- Hope is to cooperate also with HSF people deciding how an analysis facility should look like



Security

- Means....
 - k8s security in general
 - AAI (x509, tokens, VO SSO..., federated identity providers)
 - Registries and helm charts sanitisation
 -
- All encompassing might not need a group in itself but each group need to highlight the security features related to their theme



WIP document

- <https://docs.google.com/document/d/1zBKagc0RIa8xM1la7QiLQOq9-xZVcaW7OXfqPeTpJs0/edit#>

