Kubernetes and HPC

Daniel Gruber
The UberCloud

Kubernetes Intro: Native Batch Job Submission

```
kubectl apply -f job.yaml
```

```yaml
 1    apiVersion: batch/v1
 2    kind: Job
 3    metadata:
 4      name: sleep
 5    spec:
 6      template:
 7        spec:
 8          containers:
 9          - name: sleep
10            image: busybox
11            command: ["sleep", "100"]
12          restartPolicy: Never
```

Kubernetes Intro: Native Batch Job Submission

```
kubectl apply -f job.yaml
```

```yaml
1   apiVersion: batch/v1
2   kind: Job
3   metadata:
4     name: sleep
5   spec:
6     template:
7       spec:
8         containers:
9         - name: sleep
10          image: busybox
11          command: ["sleep", "100"]
12        restartPolicy: Never
```
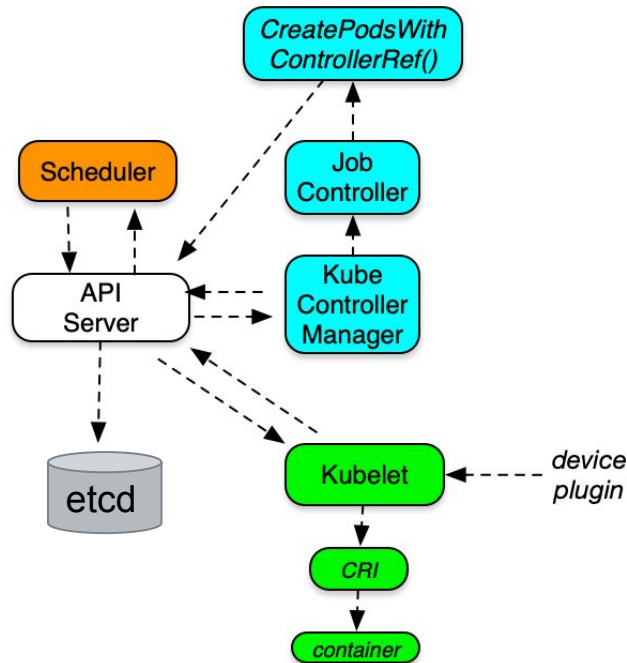
Deployment of more complex applications: *helm, Operators*

Kubernetes and HPC

HPC Requirements

➢ Fast job submission times

➢ Multi-node jobs (distributed memory)

➢ Support for GPUs, FPGAs

➢ Easily accessible by HPC applications

➢ Job queueing capabilities

➢ Job prioritization

➢ Job pre-emption

Kubernetes features

❖ *kubectl*, API

❖ No direct MPI Support

❖ GPUs supported via *device plugin*

❖ AI

❖ Pods, since 1.9

❖ 1.14

❖ 1.14

Uber Cloud

ISC High Performance
The HPC Event.

Kubernetes and HPC (continued)

HPC Requirements	Kubernetes features
➢ Advance reservation	❖ No
➢ License management capabilities	❖ No
➢ Support for Infiniband, Omnipath	❖ device plugin
➢ Full hardware utilization	❖ More daemons, typically virtualized, CPU manager
➢ Resource control	❖ Memory and CPU resources, Resource quotas
➢ Accounting and reporting	❖ Yes
➢ Storage	❖ PVs and PVCs, hostpath

References

- Kubeflow: https://github.com/kubeflow/kubeflow

- MPI Operator: https://github.com/kubeflow/mpi-operator

- Poseidon: https://github.com/kubernetes-sigs/poseidon

- CPU Manager: https://kubernetes.io/docs/tasks/administer-cluster/cpu-management-policies/

- Kube batch: https://github.com/kubernetes-sigs/kube-batch

- Slurm Operator: https://github.com/sylabs/slurm-operator

- Lustre: https://github.com/kvaps/kube-lustre

- DRMAA2: https://github.com/dgruber/drmaa2os

- RDMA: https://github.com/Mellanox/k8s-rdma-sriov-dev-plugin

- Univa's Command: http://www.univa.com/products/navops.php

- UberCloud: https://www.theubercloud.com/

Thank you

Daniel.Gruber@TheUberCloud.com

Some of my strange thoughts: www.drmaa2.org (--> https://www.gridengine.eu)

https://github.com/dgruber