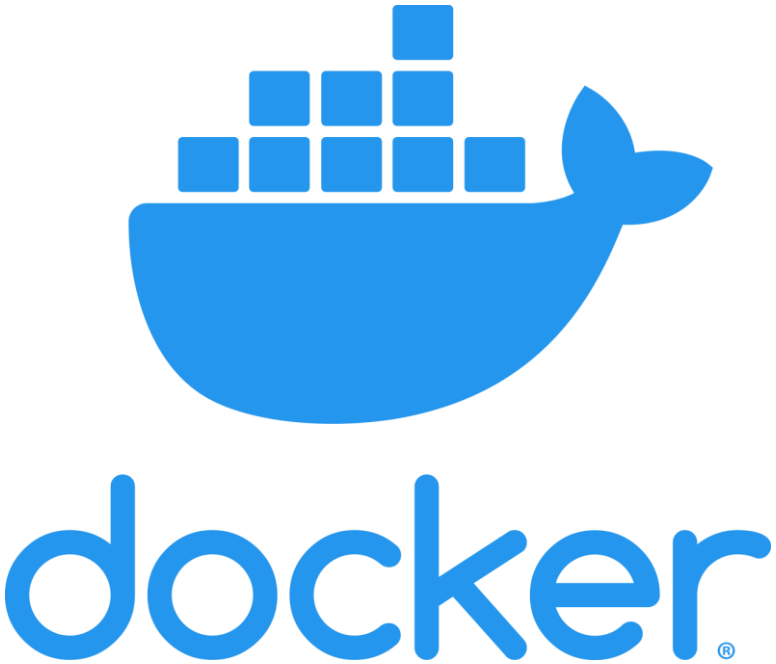
HTCondor in K8s

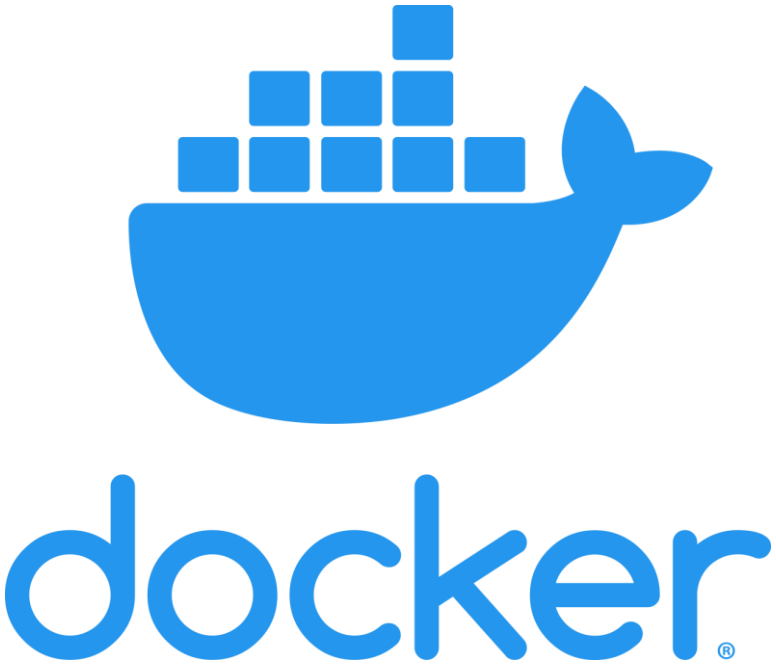
Center for High Throughput Computing

In the Beginning…

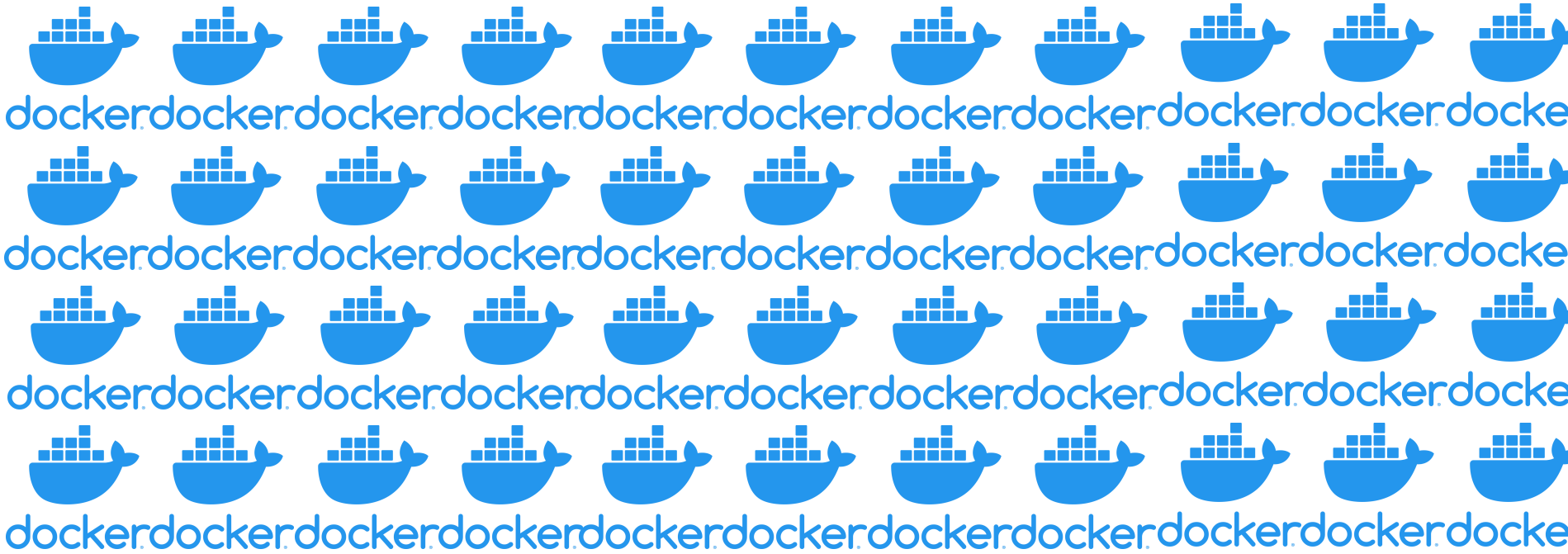
In the Beginning…



…and it was good



Until it was an unmanaged mess



Enter Kubernetes (k8s)

Tames the mess of containers on clusters

And their networks

And the storage

Kubernetes

Many users are here or working to getting here

K8s as distributed operating system
in an abstract way – like the schedd is db

Can k8s do some management things that
admins have to roll by hand today?

Summary of Kubernetes

Container Orchestrator

Sets of containers as pods

Sets of pods as deployments, etc.

Manages network and storage

Kubernetes architecture

- Central database holds all objects
  - Pods, services, storage, nodes
  - Note difference from condor – tightly coupled
- All objects described in yaml
- One command kubectl interacts with k8s

Deploying Kubernetes

- Local or in the cloud
- Most deployments are cloud based
  - could be universal cloud interface
  - ephemeral condor pools
- Local deployment is a lot of work
  - many distros

K8s and HTCondor worldview:

k8s assumes you have…

A **finite** number of services*

each of which run for an **unlimited** time

*mostly stateless

HTCondor assumes you have

An **unlimited** number of jobs

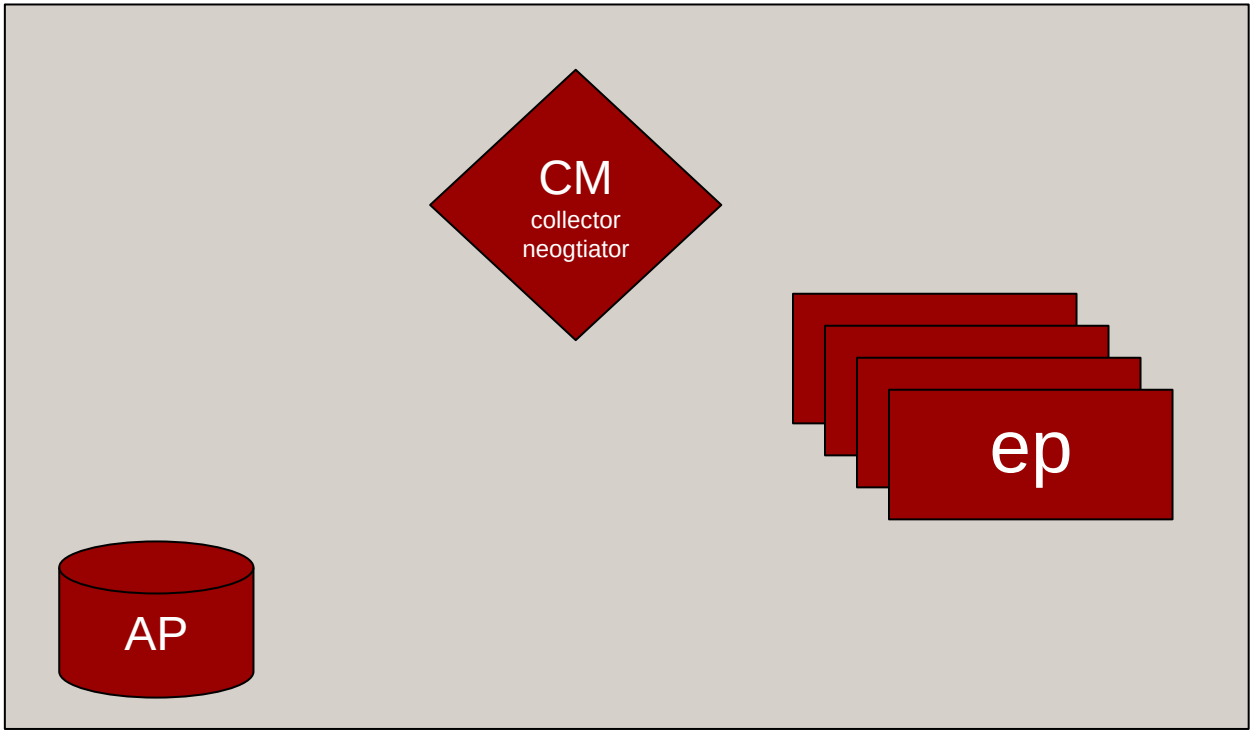
each of which run for a **finite** time

NOT k8s VS HTcondor…

K8s manages services
HTCondor is composed of services ..

So, can k8s manage the HTCondor services?

Review of HTCondor services



This requires docker images

- So, we built some…
- $ docker run htcondor/cm
- $ docker run htcondor/execute
- $ docker run htcondor/submit

(Excuse the old naming scheme…)

This requires docker images

$ docker run htcondor/cm:23.9.6-el9

Operating System

HTCondor version

And we'll be releasing with HTCondor release

You can rebase on your own distro favs

Note that the OS is coupled with the HTCondor

htcondor/cm - Docker Image

hub.docker.com/r/htcondor/cm

dockerhub

Search Docker Hub ctrl+K

Sign In

Sign up

Explore / htcondor/cm

http://hub.docker.com/htcondor/cm

htcondor/cm

By htcondor • Updated 20 days ago

A technology preview of an HTCondor central manager image.

IMAGE

Pulls 9.4K

9.4k pulls

See sources!

Overview

This is a technology preview of an HTCondor central manager Docker container.

Sources are located at https://github.com/htcondor/htcondor/tree/master/build/docker/services.

Documentation is located at https://github.com/htcondor/htcondor/blob/master/build/docker/services/README.md.

Docker Pull Command

```
docker pull htcondor/cm
```

Copy

Image Layer Details - htcondor/ | htcondor/build/docker/services | htcondor/build/docker/services | htcondor/build/docker/services

github.com/htcondor/htcondor/blob/main/build/docker/services/build-images

htcondor / htcondor

Type / to search

Code | Pull requests 32 | Actions | Security | Insights

Files

main

Go to file

- .github
- bindings
- build
  - cmake
  - docker
    - centos7
    - hpc-annex-pilot
    - k8s
    - services
      - base
      - cm
      - execute
      - mini
      - submit

htcond ... ices / build-images

timth

4c6c5a5 · last month | History

http://github.com/htcondor/htcondor/blob/main/build/docker/services/build-images

```
10          %(prog)s
11
12      Build the latest daily build for all distros that support daily builds
13      and push to the default registry:
14          %(prog)s --distro=all --push --version=daily
15
16      Build version 9.11.0 for EL8 and push to the OSG registry:
17          %(prog)s --distro=el8 --version=9.11.0 --push --registry=osg
18
19      Build the latest version of the 9.x series for Ubuntu 20.04:
20          %(prog)s --distro=ubu20.04 --series=9.x
```

Start with the CM

Best example of stateless* service

Needs port 9618 everywhere

Configure to restart with same hostnam

clients need reconfig to pick up new IP

* Mostly stateless – AccountantNew.log file

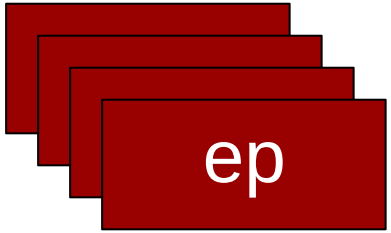
And some example yaml

```
# This is the service that names the ip address of the collector
# All pods get an environment variable with this ip in it
apiVersion: v1
kind: Service
metadata:
  name: condor
spec:
  selector:
    htcondor-role: cm
 ports:
   - protocol: TCP
     port: 9618
     targetPort: 9618
---
# This is pod yaml to describe the single htcondor central manager
apiVersion: v1
kind: Pod
metadata:
```

Debugging the CM

- Kubectl exec (or maybe ssh)
  - Is a MUST!
- /var/log/condor/*Log files very very useful
- Offline ads? Gangliad? More state

Next step: EP



EPs – where jobs run

Also mostly stateless

But – how to size? Pslots help, but…?

Like CM, logs, ssh v useful, port 9618 helpful

Consequence of an EP pod

- EP is running inside docker container
- OS of EP is OS job sees
- Usually means job can't also be a docker
  - (singularity/apptainer OK, though)
- Also means not privileged
  - No cgroups, priv switching, etc. etc. etc.

Provisioning via Rooster

- Can use offline ads to automatically provision
- Create "fake" ads, if they match, submit pod to create real capacity
- PATh Fac works this way, better support for 1st class coming with Glidein Manager

Final step: AP

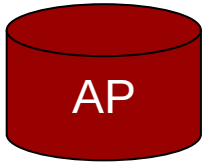
AP – 100 % state

Does it even make sense to put in k8s?

May want to keep it outside k8s

If so, home directories, /var/spool/condor

Login sessions, etc. etc.



Post final step -- CE

CEs make a lot of sense in k8s

Most OSG CEs are "managed CEs", and run in k8s

We can reuse OSG scripts to run your CE

CHTC Experience w k8s

Generally good, but …

- Everything has consequences
- Debugging inside is tricky
- k8s moves fast – prepare to upgrade a lot
- Need to live with immutable image mindset

Thank you

Questions?

Center for High Throughput Computing