Docker SWARM: Light-weight orchestrator

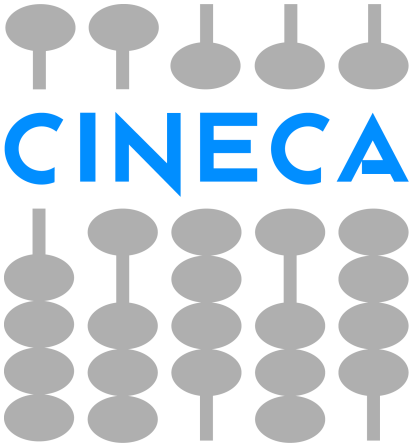
HPCW @ ISC2019

Abdulrahman Azab

Dept. of Research Computing

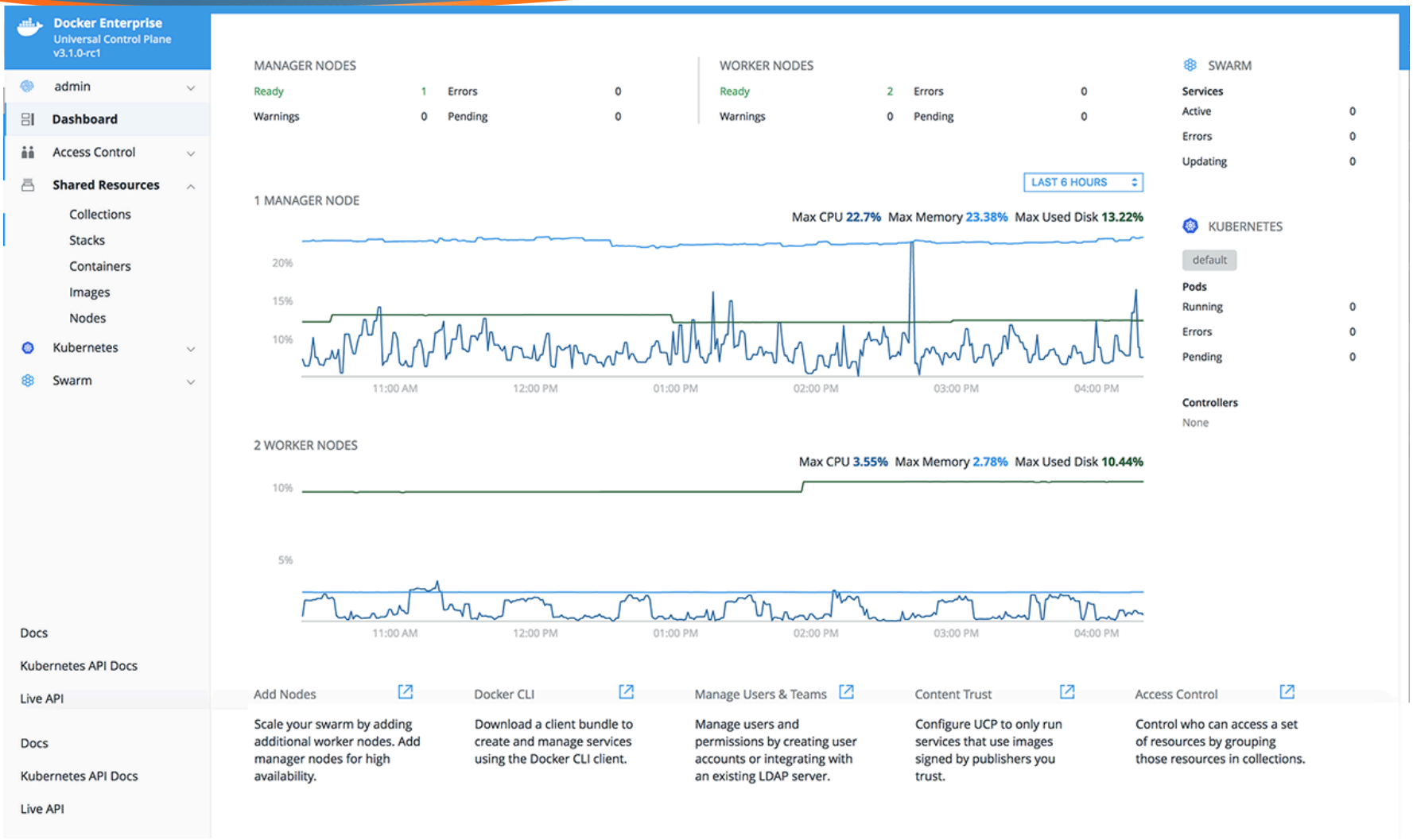
University of Oslo, Norway

PRACE 6IP WP6.2.3: The deployment of containers and full virtualised tools into HPC infrastructures

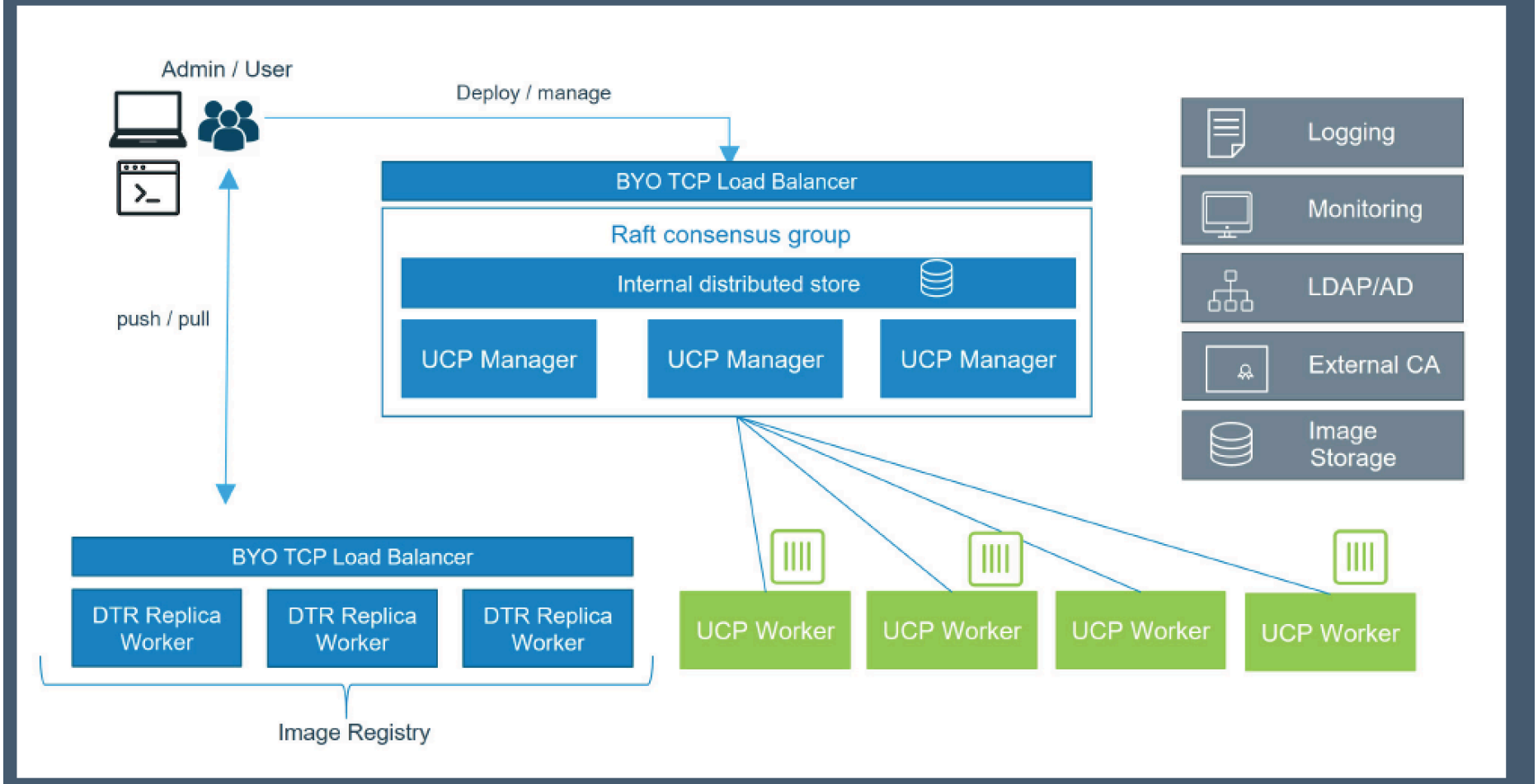


ISC High Performance
The HPC Event.

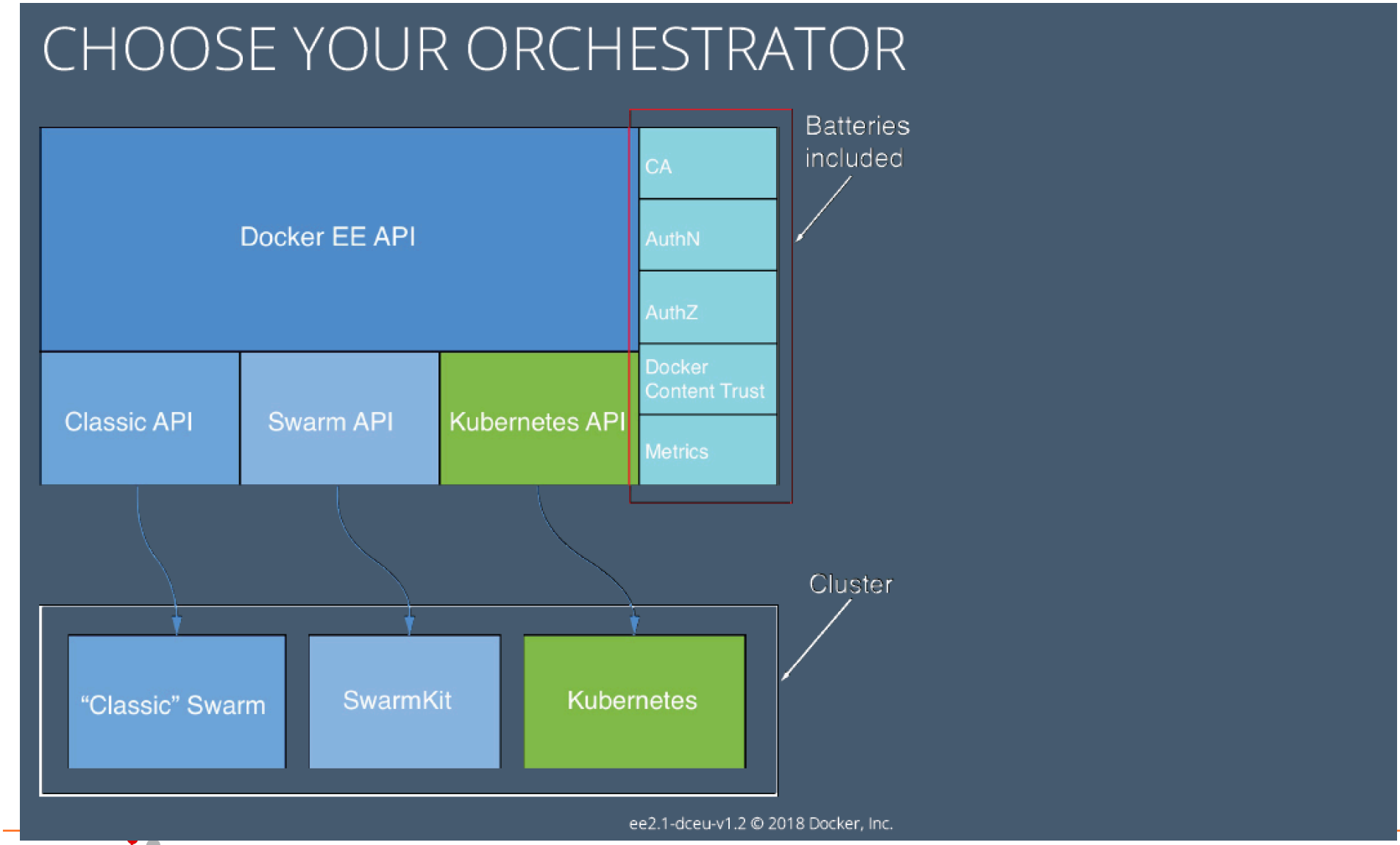
Docker UCP


Docker Enterprise
Universal Control Plane
v3.1.0-rc1
admin
Dashboard
Access Control
Shared Resources
Collections
Stacks
Containers
Images
Nodes
Kubernetes
Swarm
Docs
Kubernetes API Docs
Live API
Docs
Kubernetes API Docs
Live API
MANAGER NODES
Ready 1 Errors 0
Warnings 0 Pending 0
WORKER NODES
Ready 2 Errors 0
Warnings 0 Pending 0
SWARM
Services
Active 0
Errors 0
Updating 0
LAST 6 HOURS
1 MANAGER NODE
Max CPU 22.7% Max Memory 23.38% Max Used Disk 13.22%
20%
15%
10%
11:00 AM
12:00 PM
01:00 PM
02:00 PM
03:00 PM
04:00 PM
KUBERNETES
default
Pods
Running 0
Errors 0
Pending 0
Controllers
None
2 WORKER NODES
Max CPU 3.55% Max Memory 2.78% Max Used Disk 10.44%
10%
5%
11:00 AM
12:00 PM
01:00 PM
02:00 PM
03:00 PM
04:00 PM
Add Nodes
Scale your swarm by adding additional worker nodes. Add manager nodes for high availability.
Docker CLI
Download a client bundle to create and manage services using the Docker CLI client.
Manage Users & Teams
Manage users and permissions by creating user accounts or integrating with an existing LDAP server.
Content Trust
Configure UCP to only run services that use images signed by publishers you trust.
Access Control
Control who can access a set of resources by grouping those resources in collections.

DOCKER EE ARCHITECTURE

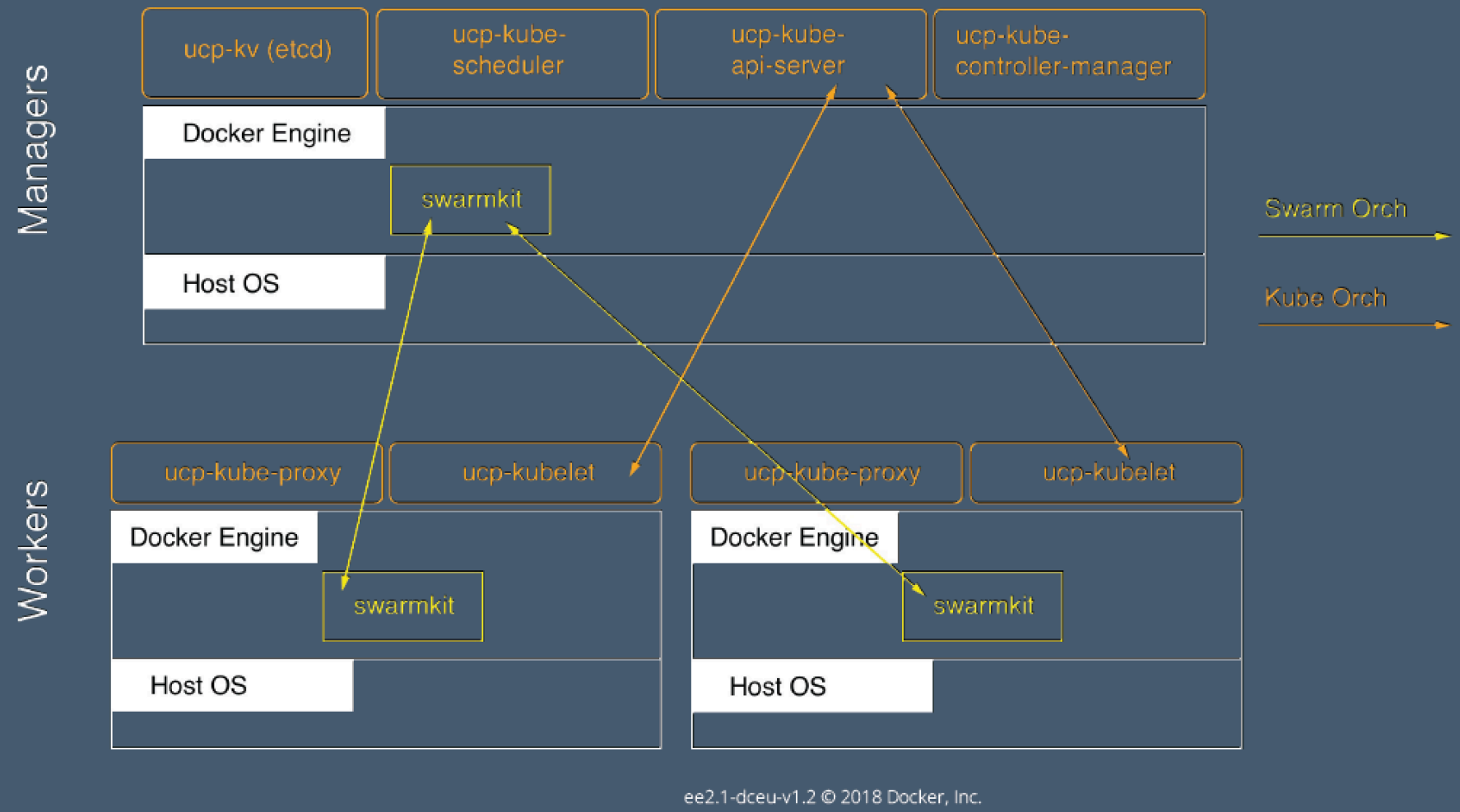


```
[centos@ucp-manager-0 ~]$ UCP_IP=<ucp-manager-0 IP>
[centos@ucp-manager-0 ~]$ UCP_FQDN=<ucp-manager-0 FQDN>
[centos@ucp-manager-0 ~]$ docker container run --rm -it --name
ucp \
-v /var/run/docker.sock:/var/run/docker.sock \
docker/ucp:3.1.0 install \
--admin-username admin \
--admin-password adminadmin \
--san ${UCP_IP} \
--san ${UCP_FQDN}
```

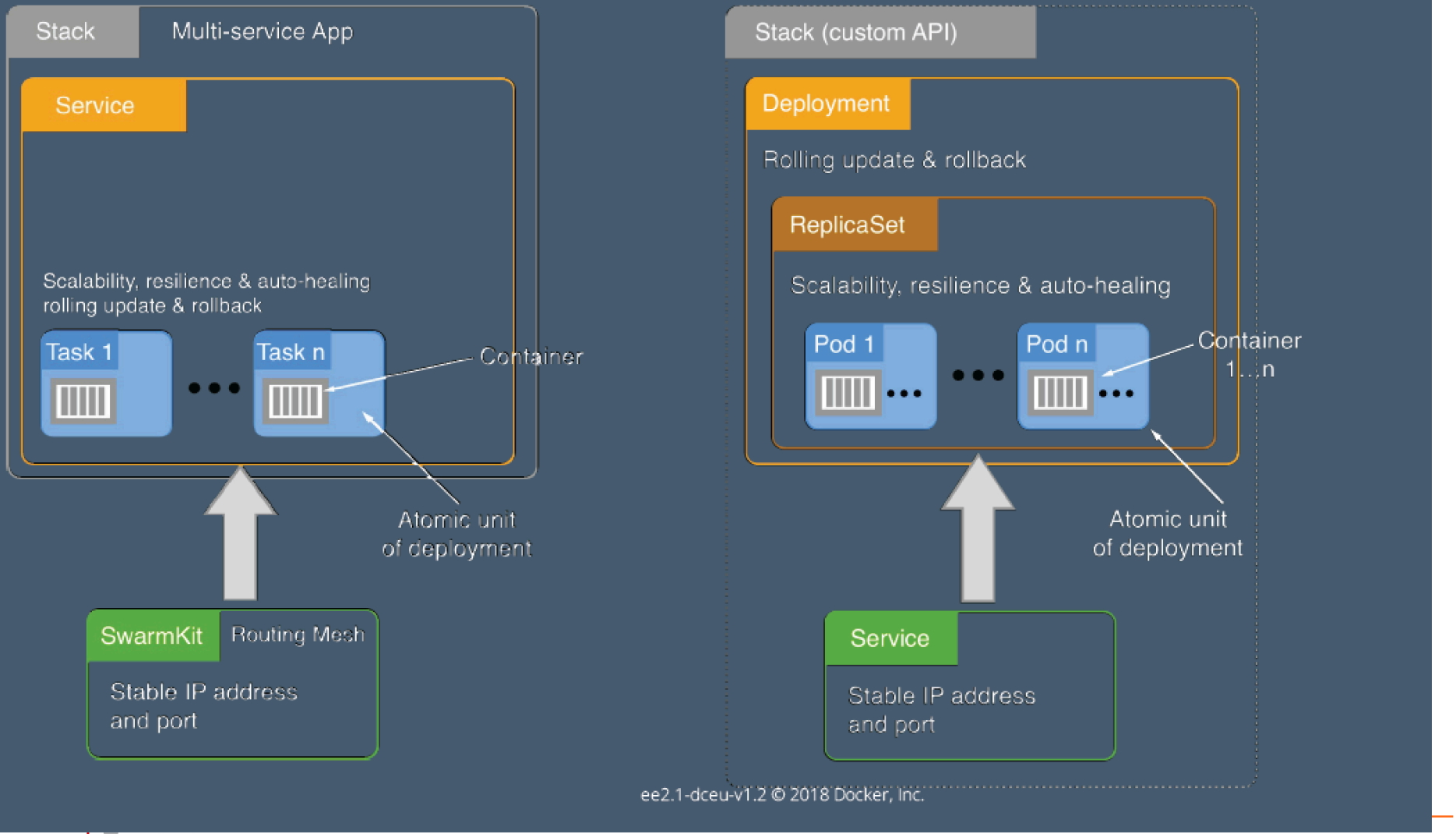

CHOOSE YOUR ORCHESTRATOR
Batteries included
CA
AuthN
AuthZ
Docker Content Trust
Metrics
Docker EE API
Classic API
Swarm API
Kubernetes API
Cluster
"Classic" Swarm
SwarmKit
Kubernetes
ee2.1-dceu-v1.2 © 2018 Docker, Inc.

Orchestration

ORCHESTRATORS IN UCP



ORCHESTRATION COMPONENTS



Container Network Operations

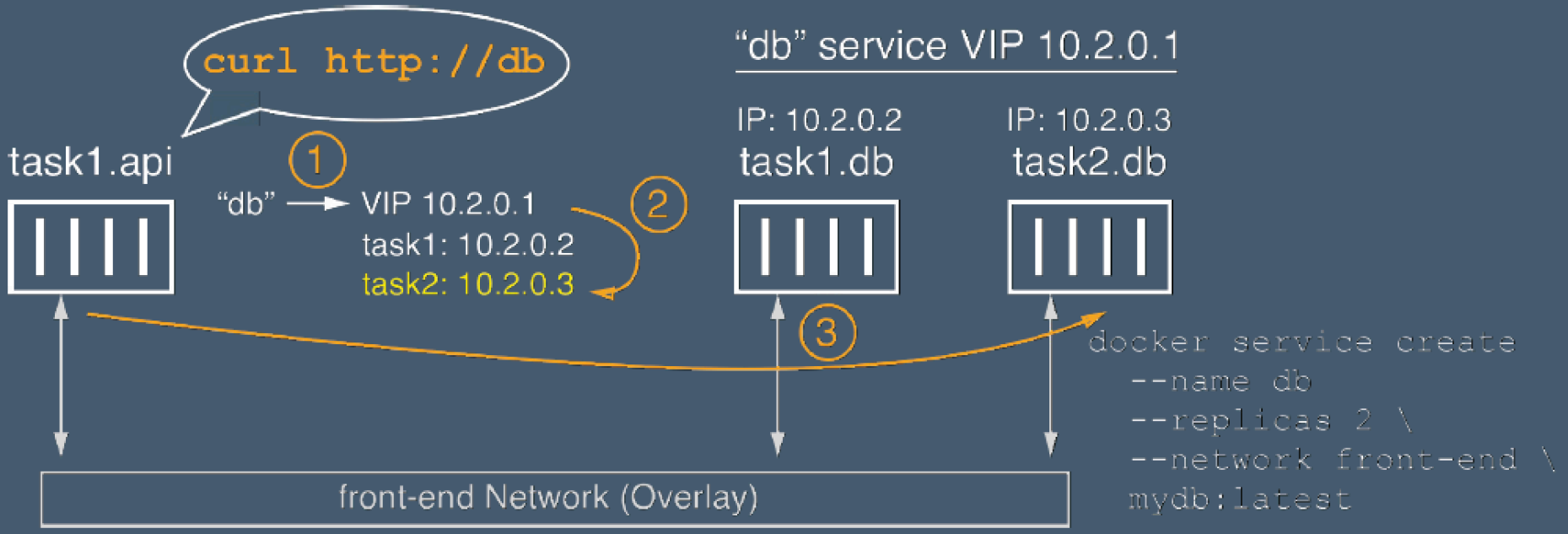
CHOOSING THE RIGHT NETWORKING TOOL

Three questions:

- Is the request originator internal or external to your cluster?
- Are the destination containers stateless or stateful?
- Are you using Swarm or Kubernetes?

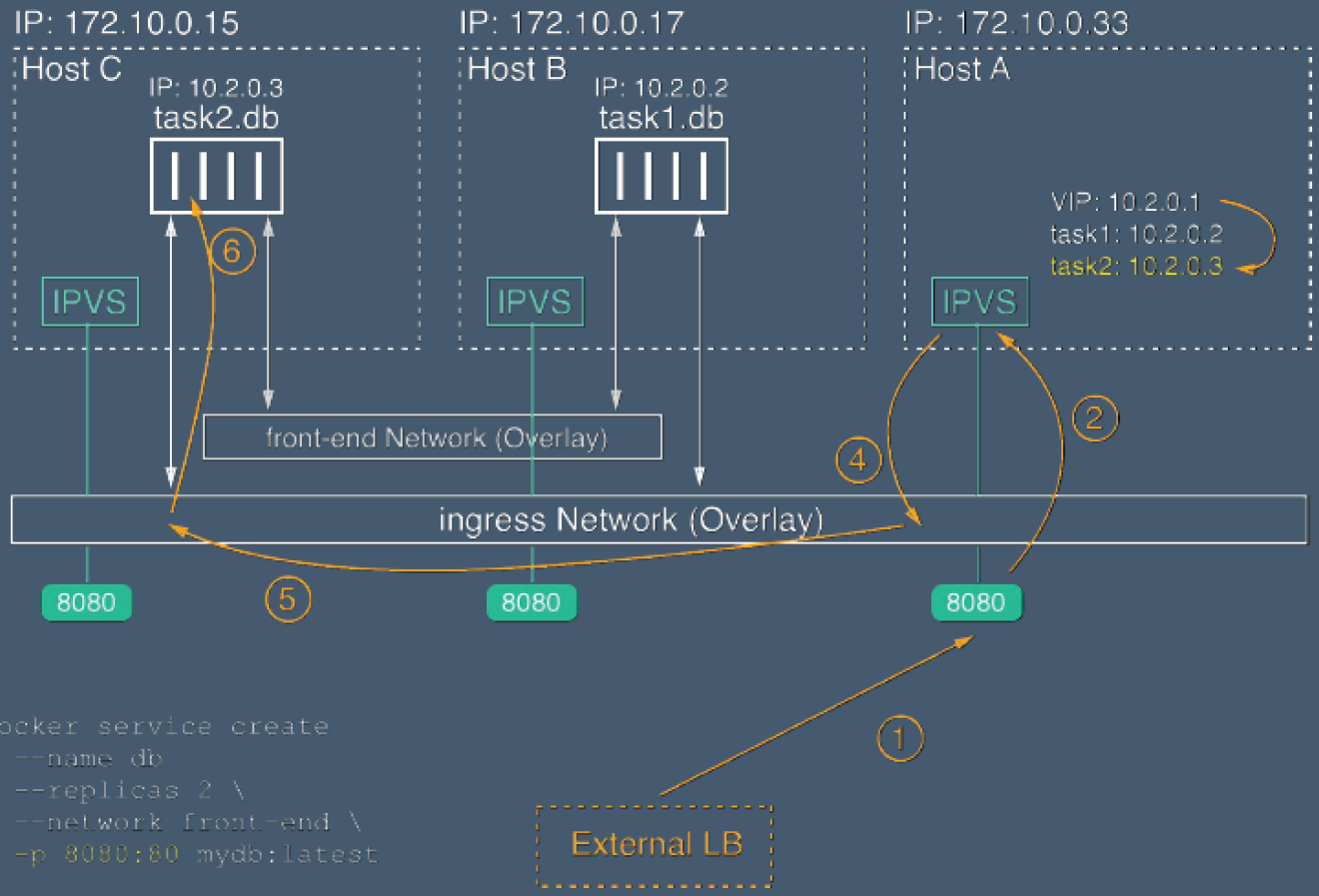
INTERNAL / STATELESS / SWARM

Solution: Swarm VIPs



EXTERNAL / STATELESS / SWARM

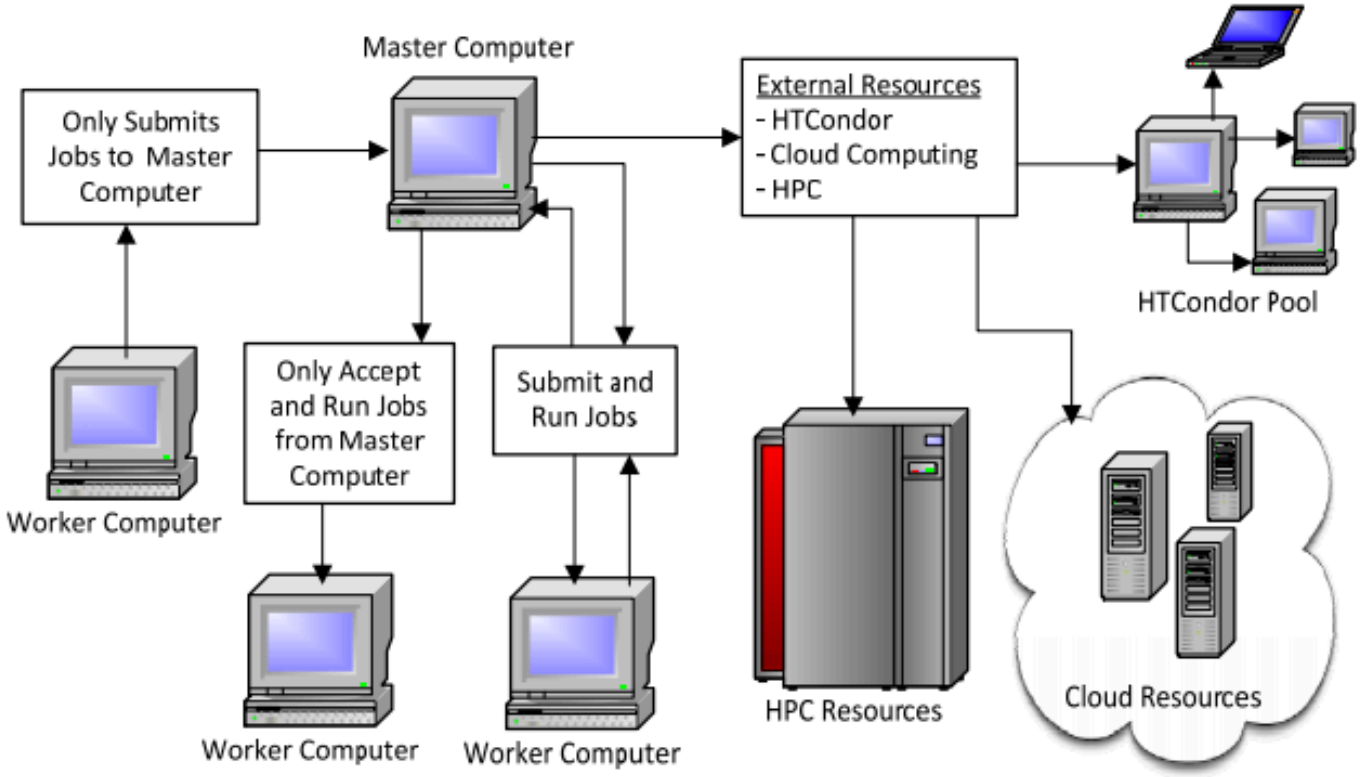
Solution: Swarm L4 Routing Mesh




```
docker service create
  --name db
  --replicas 2 \
  --network front-end \
  -p 8080:80 mydb:latest
```

Use case:
Containerised HPC
With HTConcor and swarm

HTCondor


Master Computer
Only Submits Jobs to Master Computer
External Resources
- HTCondor
- Cloud Computing
- HPC
HTCondor Pool
Worker Computer
Only Accept and Run Jobs from Master Computer
Submit and Run Jobs
HPC Resources
Cloud Resources
Worker Computer
Worker Computer

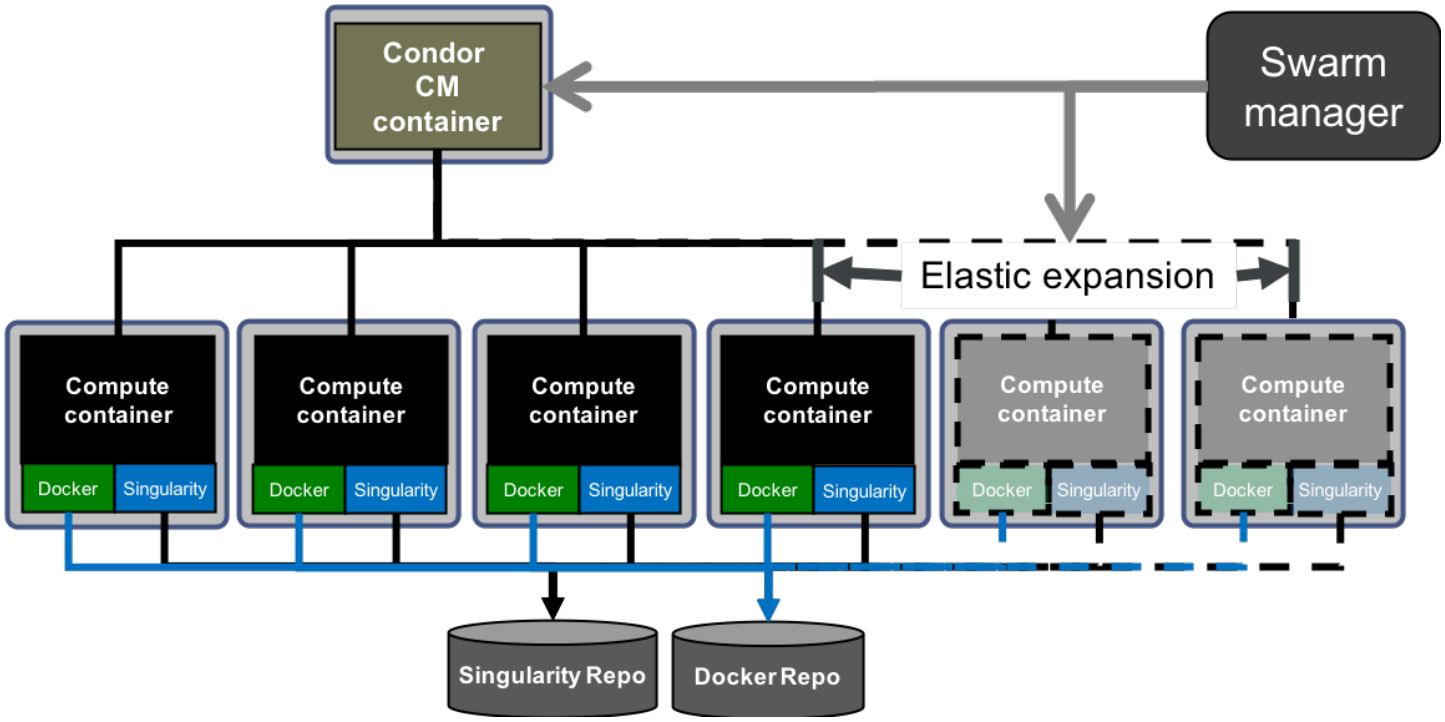
HTCondor VM universe

```
universe = vm
executable = vmware_sample_job
log = simple.vm.log.txt
vm_type = vmware
vm_memory = 64
vmware_dir = C:\condor-test
vm_checkpoint = true
queue
```

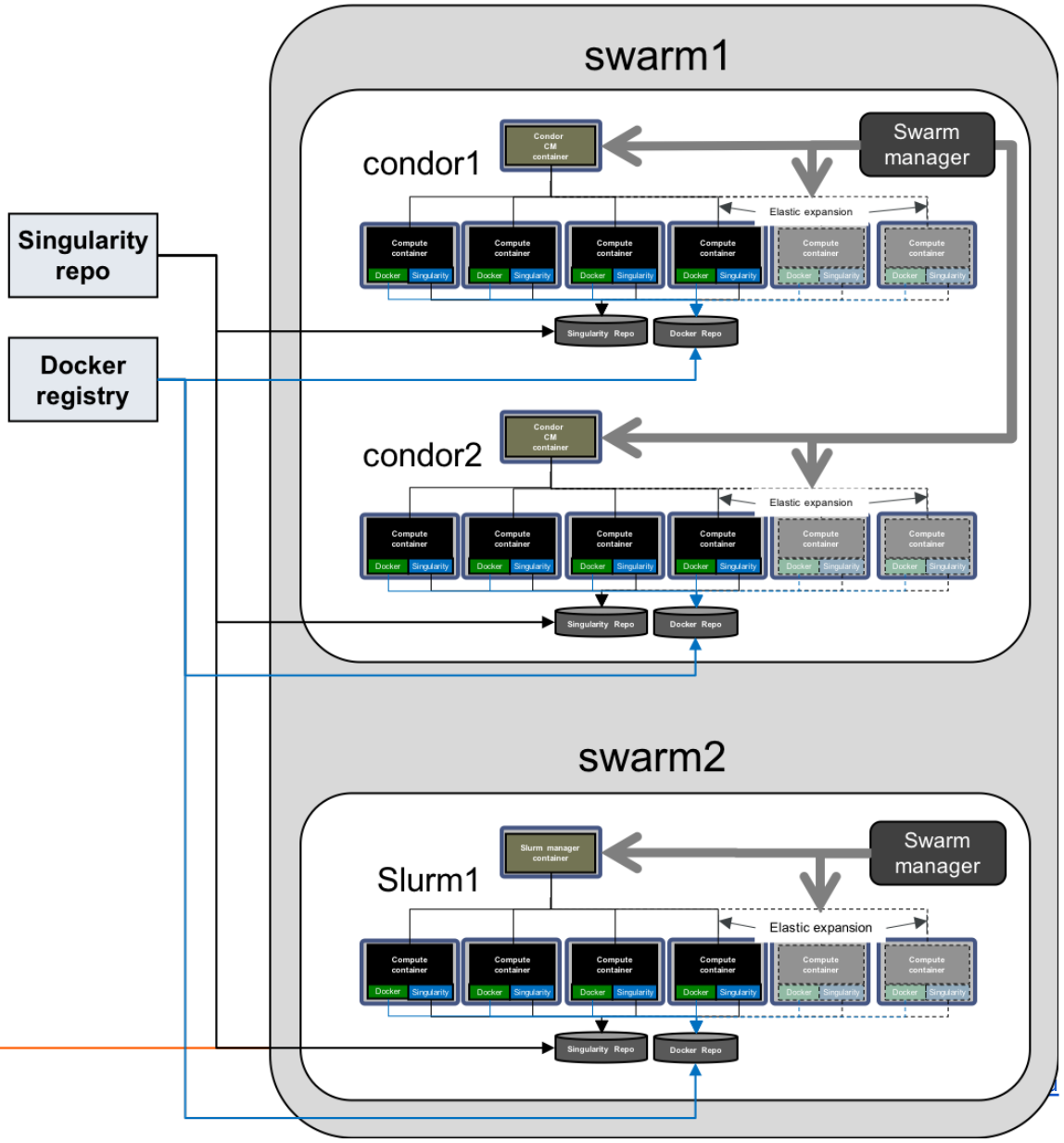
HTCondor Docker universe

```
universe = docker
docker_image = debian
executable = /bin/cat
arguments = /etc/hosts
output = out.$(Process)
error = err.$(Process)
request_memory = 100M
queue 10
```

Containerised HTCondor on Swarm



Docker only cluster



```
[cloud-user@swarm-1 ~]$ docker node ls
ID                            HOSTNAME              STATUS              AVAILABILITY        MANAGER STATUS
RSION
smzp49rwzhc6g5m3rbnwa4o41 *   swarm-1.novalocal     Ready               Active              Leader
e
2x9m0zr61qnm7gvnhwwjgn263      swarm-2.novalocal     Ready               Active
e
n99dyxddexkr0w0wq8ysz3tw3      swarm-3.novalocal     Ready               Active
```

```
[cloud-user@swarm-1 ~]$ docker service ls
ID                  NAME                                MODE                REPLICAS
                                  PORTS
5rme9l2t03vi        galaxy_galaxy-htcondor              replicated          1/1
axy-htcondor:18.01b
z6seh1nyhjep        galaxy_galaxy-htcondor-executor     replicated          2/2
axy-htcondor-executor:18.01b
vjq5y0dd5ni5        galaxy_galaxy-htcondor-executor-big replicated          1/1
axy-htcondor-executor:18.01b
```

```
[cloud-user@swarm-1 ~]$ docker service ps galaxy_galaxy-htcondor-executor | grep Running
59zcdu1us630        galaxy_galaxy-htcondor-executor.1       quay.io/bgruening/galaxy-htcondor-executor:18.01b   swarm-3
.novalocal   Running             Running 2 months ago
4dygqif85wi2        galaxy_galaxy-htcondor-executor.2       quay.io/bgruening/galaxy-htcondor-executor:18.01b   swarm-3
.novalocal   Running             Running 2 months ago
```

```
[cloud-user@swarm-1 ~]$ docker service scale galaxy_galaxy-htcondor-executor=5
galaxy_galaxy-htcondor-executor scaled to 5
overall progress: 5 out of 5 tasks
1/5: running   [==================================================>]
2/5: running   [==================================================>]
3/5: running   [==================================================>]
4/5: running   [==================================================>]
5/5: running   [==================================================>]
verify: Service converged
[cloud-user@swarm-1 ~]$ docker service ps galaxy_galaxy-htcondor-executor | grep Running
59zcdu1us630        galaxy_galaxy-htcondor-executor.1       quay.io/bgruening/galaxy-htcondor-executor:18.01b   swarm-3
.novalocal   Running             Running 2 months ago
4dygqif85wi2        galaxy_galaxy-htcondor-executor.2       quay.io/bgruening/galaxy-htcondor-executor:18.01b   swarm-3
.novalocal   Running             Running 2 months ago
t1cadorzttuh        galaxy_galaxy-htcondor-executor.3       quay.io/bgruening/galaxy-htcondor-executor:18.01b   swarm-2
.novalocal   Running             Running 15 seconds ago
av7mt5t7i6jk        galaxy_galaxy-htcondor-executor.4       quay.io/bgruening/galaxy-htcondor-executor:18.01b   swarm-2
.novalocal   Running             Running 15 seconds ago
920zv43mmaox        galaxy_galaxy-htcondor-executor.5       quay.io/bgruening/galaxy-htcondor-executor:18.01b   swarm-2
.novalocal   Running             Running 15 seconds ago
```

```
[cloud-user@swarm-3 ~]$ docker exec 53463399ae3a condor_status
Name               OpSys      Arch   State     Activity LoadAv Mem   ActvtyTime

2b928ae59394       LINUX      X86_64 Unclaimed Idle      0.160 1024  0+00:00:03
5ed8939b3952       LINUX      X86_64 Unclaimed Idle      0.000 2048 77+05:43:00
2073aae2920d       LINUX      X86_64 Unclaimed Idle      0.160 1024  0+00:00:03
a4740170e2cf       LINUX      X86_64 Unclaimed Idle      0.000 1024 77+05:46:51
c3d66659047a       LINUX      X86_64 Unclaimed Idle      0.000 1024 77+05:46:55
c55e3685352c       LINUX      X86_64 Unclaimed Idle      0.160 1024  0+00:00:03

                     Total Owner Claimed Unclaimed Matched Preempting Backfill  Drain

       X86_64/LINUX     6     0       0         6       0          0        0      0

              Total     6     0       0         6       0          0        0      0
```


Q&A