



Hewlett Packard
Enterprise

GALERA  CLUSTER

Galera on Kubernetes

Running synchronous MySQL
replication on Kubernetes

Patrick Galbraith, ATG
April 2016



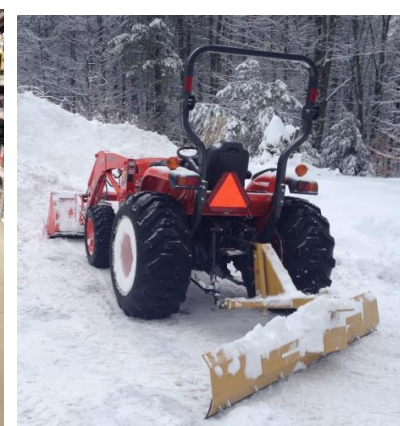
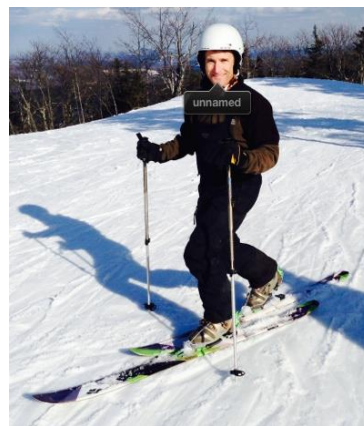
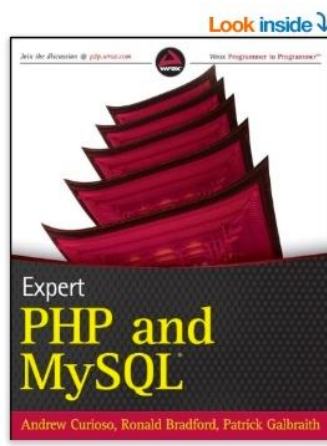
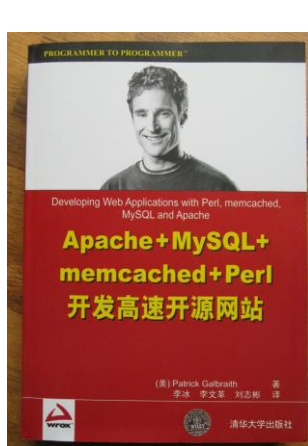
kubernetes

HPE ATG

HPE's (HP Enterprise) Advanced Technology Group for Open Source and Cloud embraces a vision that is two steps ahead of today's solutions. We use this vision to drive product adoption and incubate technologies to advance HPE. Through Open Source initiatives we foster collaboration across HPE and beyond.

About the speaker

- Patrick Galbraith
- HP Advanced Technology Group
- Has worked at Blue Gecko, MySQL AB, Classmates, Slashdot, Cobalt Group, US Navy, K-mart
- MySQL projects: memcached UDFs, DBD::mysql, federated storage engine
- Family
- Outdoors



Purpose of this talk – why are you here?

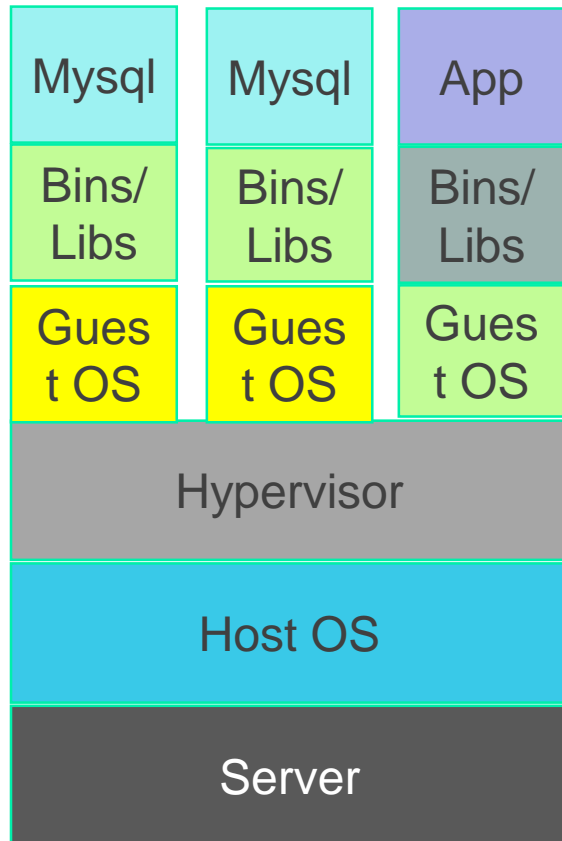
- Docker
 - Containers vs. Virtualization
 - Simple Docker usage
 - Clustered Docker
- Kubernetes
 - Understand what Kubernetes is
 - Using Kubernetes to do work, in this case what it means to MySQL users
- Galera
 - Synchronous replication – excellent solution to clustering MySQL

What are containers?

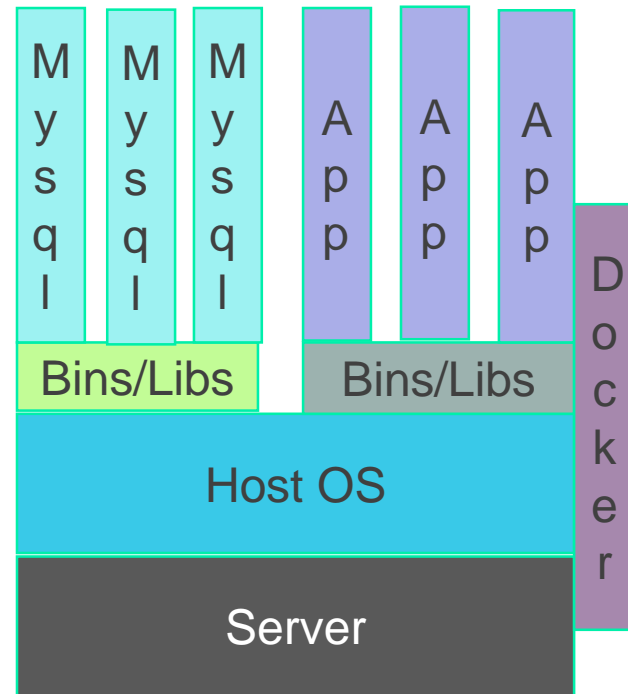
- Operating-system-level isolation
- Encapsulated, hermetically sealed applications
- Portable. And did I mention, portable?!
- Small footprint of container images
- Fast to launch!
- Use of host OS and Kernel
- Execution consists of time to startup application in question
- LXC, Docker, Solaris Zones, BSD Jails, Parallels Virtuozzo, OpenVZ, ...

VM vs. Containers

VM



Containers



What is Docker?

- Set of tools for managing containers
- Command line tool that doubles as a daemon
- Kernel namespaces – the core ingredient to containers working
 - PID
 - IPC
 - uts (what will be seen by a group of processes)
 - Mount
 - Network
 - User
 - Cgroups (control groups) -- limit, account and isolate resource usage (CPU, memory, disk I/O, etc.) of process groups
- Originally used lxc, now defaults to Libcontainer but meant for any containerization mechanism
- Much more light weight than VMs
- Encapsulated application containers in a relatively isolated but lightweight operating environment
- Written in Go

Docker – common terms and usage

- Dockerfile
- EXPOSE ports
- Entrypoints and CMD
- `docker build`
- `docker push`
- `docker run`
- `docker inspect`
- `docker exec`
- `docker commit`

Dockerfile

- https://github.com/CaptTofu/percona_xtradb_cluster_docker

Running a docker container

```
$ docker run \  
  --name mybox \  
  -e MYSQL_ROOT_PASSWORD=secret \  
  -d \  
  mysql/mysql-server --log-bin --server-id=100$
```

```
$ cat minimal.cnf  
[mysqld]  
user=mysql  
log-bin=mysql-bin  
server-id=100
```

```
$ docker run \  
  --name mybox \  
  -e MYSQL_ROOT_PASSWORD=secret \  
  -d --hostname mybox \  
  -v $PWD/minimal.cnf:/etc/my.cnf  
  mysql/mysql-server
```

Clustered Docker

- Kubernetes -- <http://kubernetes.io>
- CoreOS -- <https://coreos.com/>
- Mesos + Marathon -- <http://mesos.apache.org/> Apache project, Zookeeper, etc
- Project Atomic -- <http://www.projectatomic.io/> -- RH/Fedora/Centos designed for ru
- Docker Openstack -- <https://wiki.openstack.org/wiki/Docker> Hypervisor Driver for C
- Swarm/Compose/Machine
- RancherOS <http://rancher.com/rancher-os> Minimalist Linux, Docker daemon runs
- Flocker -- <https://clusterhq.com>
- Spotify Helios -- <https://github.com/spotify/helios> -- Zookeeper
- Deis (<http://deis.io>)
- Maestro (<https://github.com/toscanini/maestro>)
- Shipyard (<http://shipyard-project.com>)
- ... others to come!

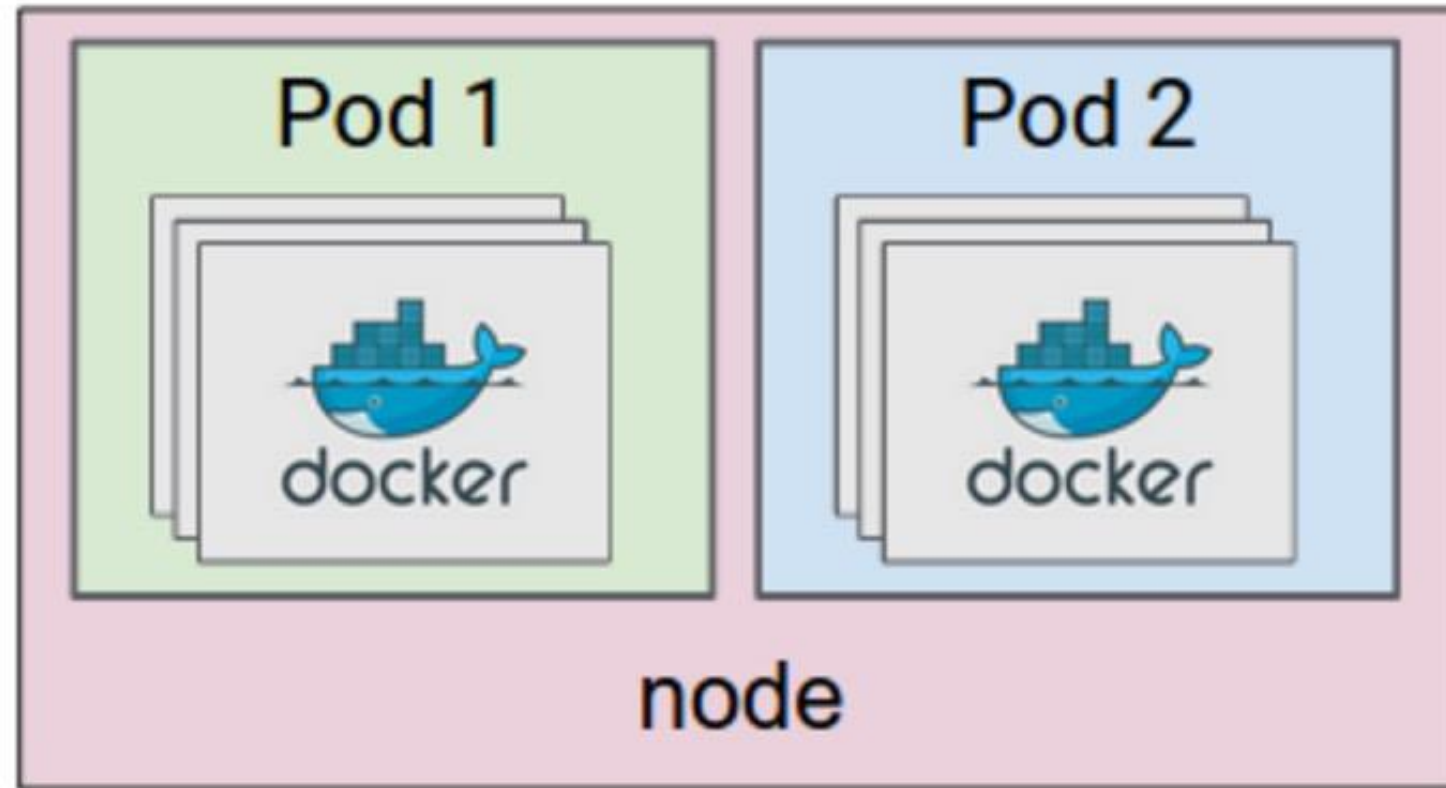
Kubernetes

- “An open source system for managing containerized applications across multiple hosts”
- Lean
- Portable – will run cloud, bare metal, hybrid, etc
- Extensible – using modular design allowing for plug-ability and hooks
- Self-healing – auto-placement, auto-restart, auto-replication
- Google engineering bring good work to the Open-source world

Kubernetes concepts

- Pod
 - Group of closely-related containers on the same host
- Service
 - Virtual abstraction
 - Basic load-balancer
 - Single consistent access point to a pod
- Replication controller
 - Defines pods to be horizontally scaled
 - Uses a label query for identifying what containers to run
 - Maintains specified number of replicas of a particular thing to run
 - Dynamic resizing
- Label
 - Key/value tag to mark work units a part of group
 - Management and action targeting
- Definition file – YAML/json describing a pod, service, or replication controller

Kubernetes pod



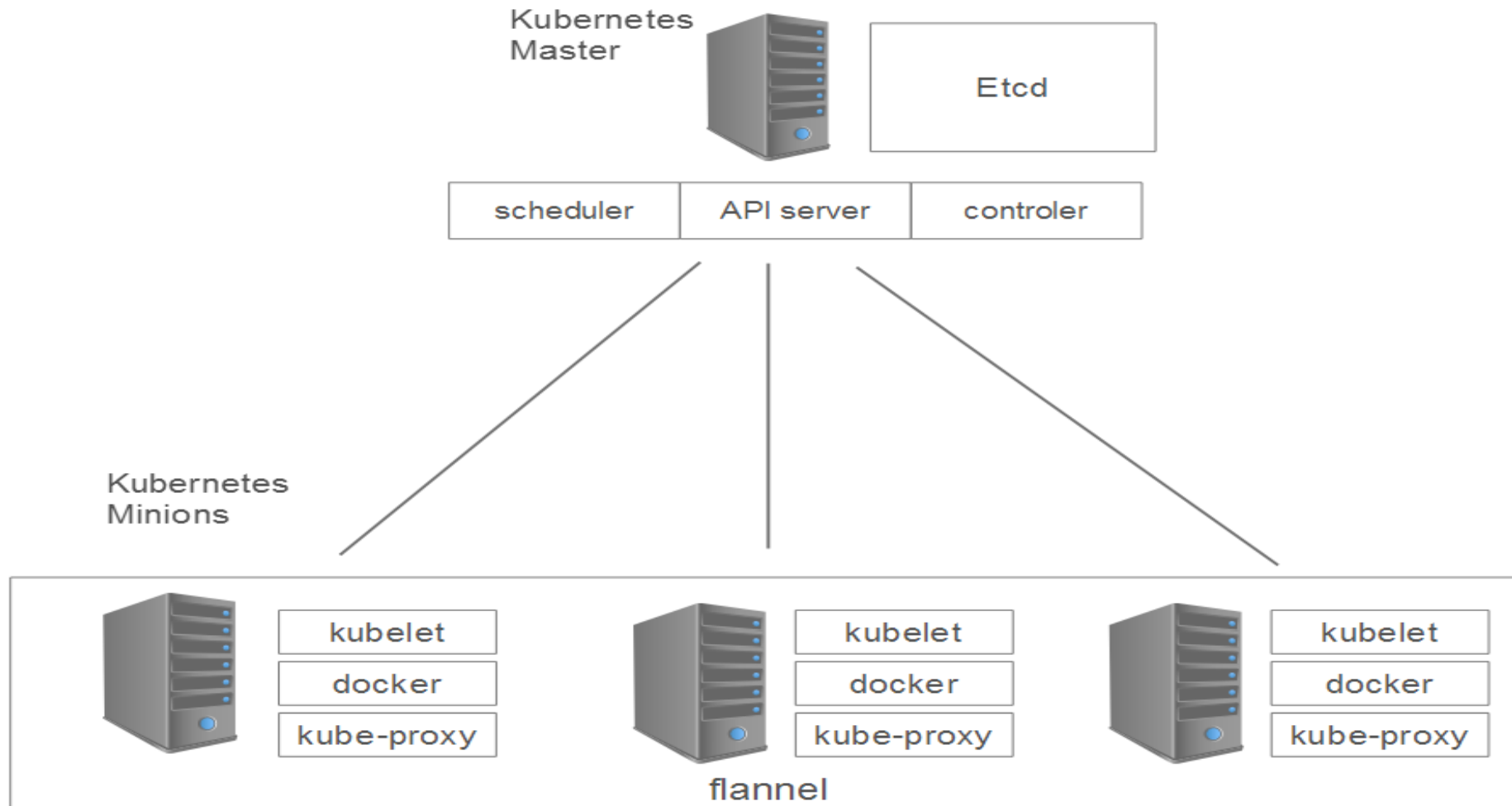
Kubernetes Master

- `kube-apiserver` – API Server (RESTful)
 - primary management for cluster
 - reconciles etcd entries with deployed containers
- `kube-controllermanager` — Controller Manager Server
 - Handle replication processes defined by replication tasks
 - Writes details to etcd
 - Monitors changes and implements procedure to reflect the change
- `kube-scheduler` -- Scheduler Server
 - Assigns workloads to specific minions in cluster taking into account service's operating requirements and infrastructure environment
- `kube-register` -- Register Server

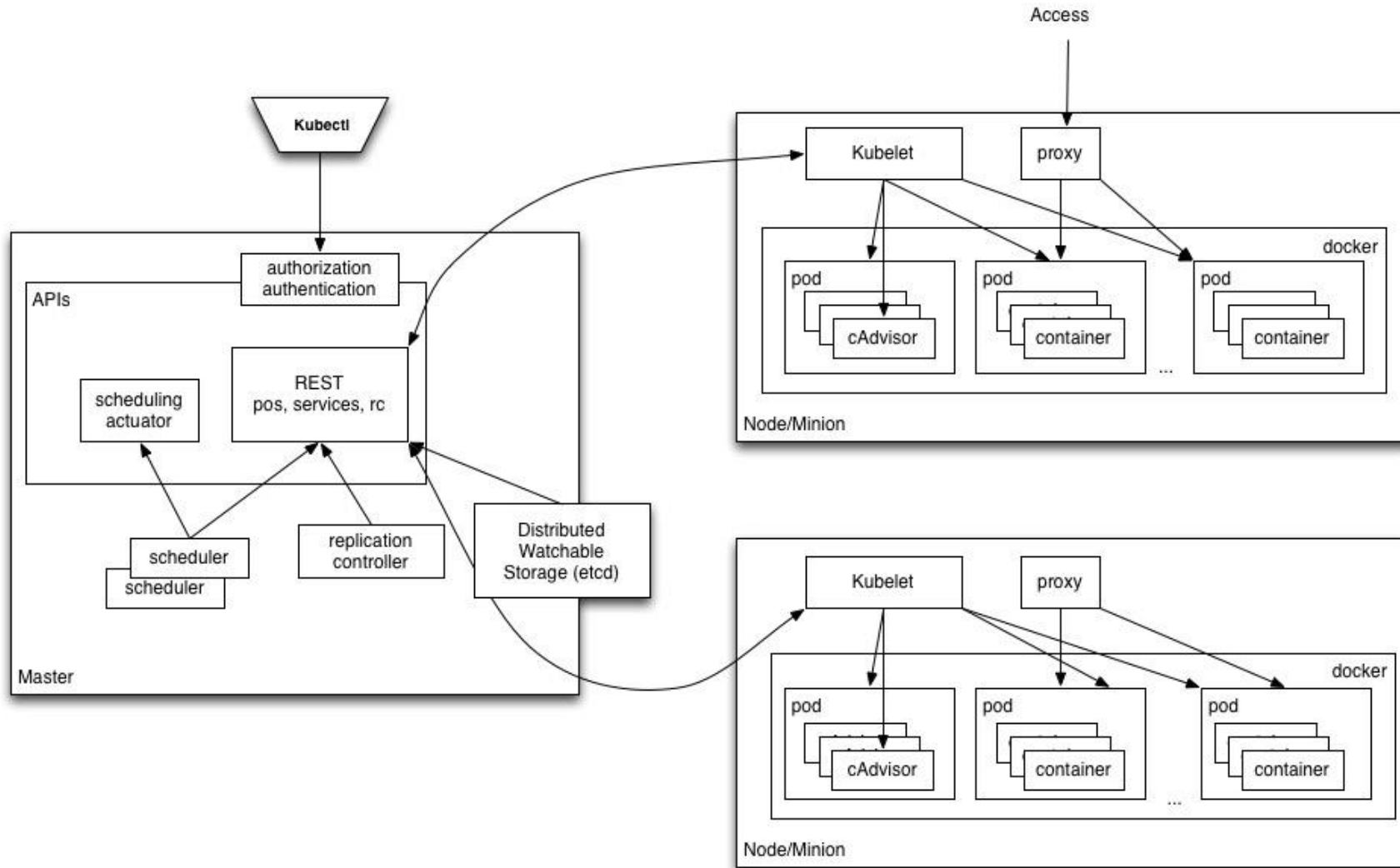
Kubernetes minion

- `kubelet`
 - Communicates with the master, relaying information to/from
 - Reads and updates etcd
 - Receives work in a manifest that defines the workload and operating parameters.
 - Assumes responsibility for the state of work on minion
- `kube-proxy`
 - Ensures network environment is accessible but isolated.
 - Makes services available externally by forwarding requests to containers.
 - Can perform rudimentary load balancing.

Kubernetes Basic Setup



Kubernetes diagram



Kubernetes usage

- Pod configuration file – YAML or JSON
- Service configuration file
- Replication controller configuration file
- `export KUBERNETES_API=http://kube-master:8080`
- `kubectl create -f mysql_master.json`
- `kubectl create -f mysql_master_service.json`
- ...

How can I run Kubernetes?

- <https://github.com/CaptTofu/vagrant-kubernetes-cluster.git>
- <https://github.com/Samsung-AG/kraken.git>
- <http://kubernetes.io/docs/hellonode/>
- <https://github.com/pires/kubernetes-vagrant-coreos-cluster>
- <https://github.com/TheNewNormal/kube-solo-osx>

Other deployment Strategies for Kubernetes

- Helm -- <http://helm.sh/> -- Package Manager for k8s
- Deis v2 – Builds, deploys, 12-factor PaaS
- Kupak
- RedSpread
- KPM

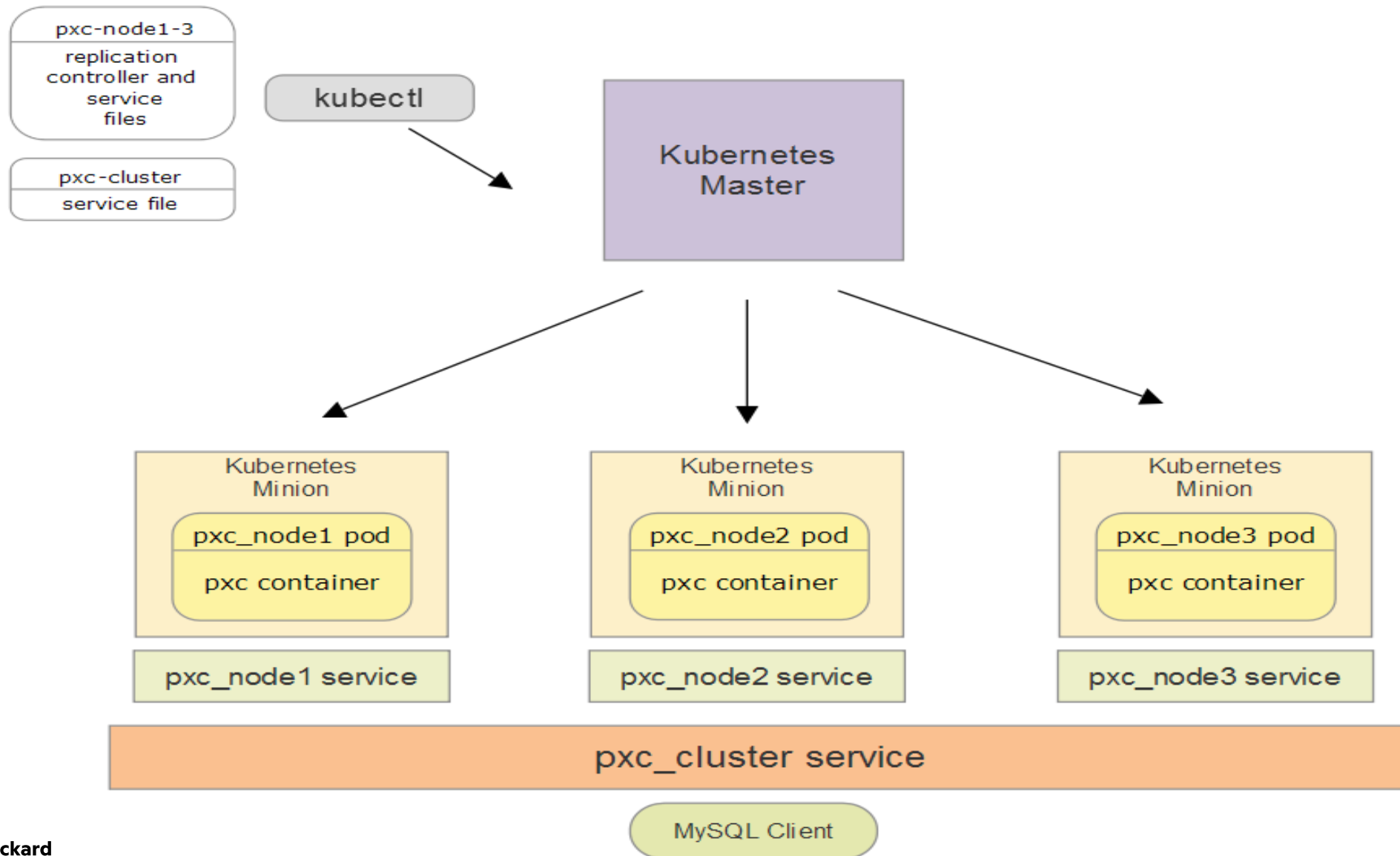
Galera Replication on Kubernetes

- One pod per node (pxc_node1, pxc_node2, and pxc_node03)
- Utilizes SkyDNS
- Single container per pod
- Secret sauce
 - The pod configuration file passes environment variables to set root password, sst u
 - Entrypoint script:
 - `wsrep cluster address on pxc_node1 set to gcomm://`
 - `pxc_node2 set to gcomm://pxc_node1`
 - `pxc_node3 set to gcomm://pxc_node1,pxc_node2`

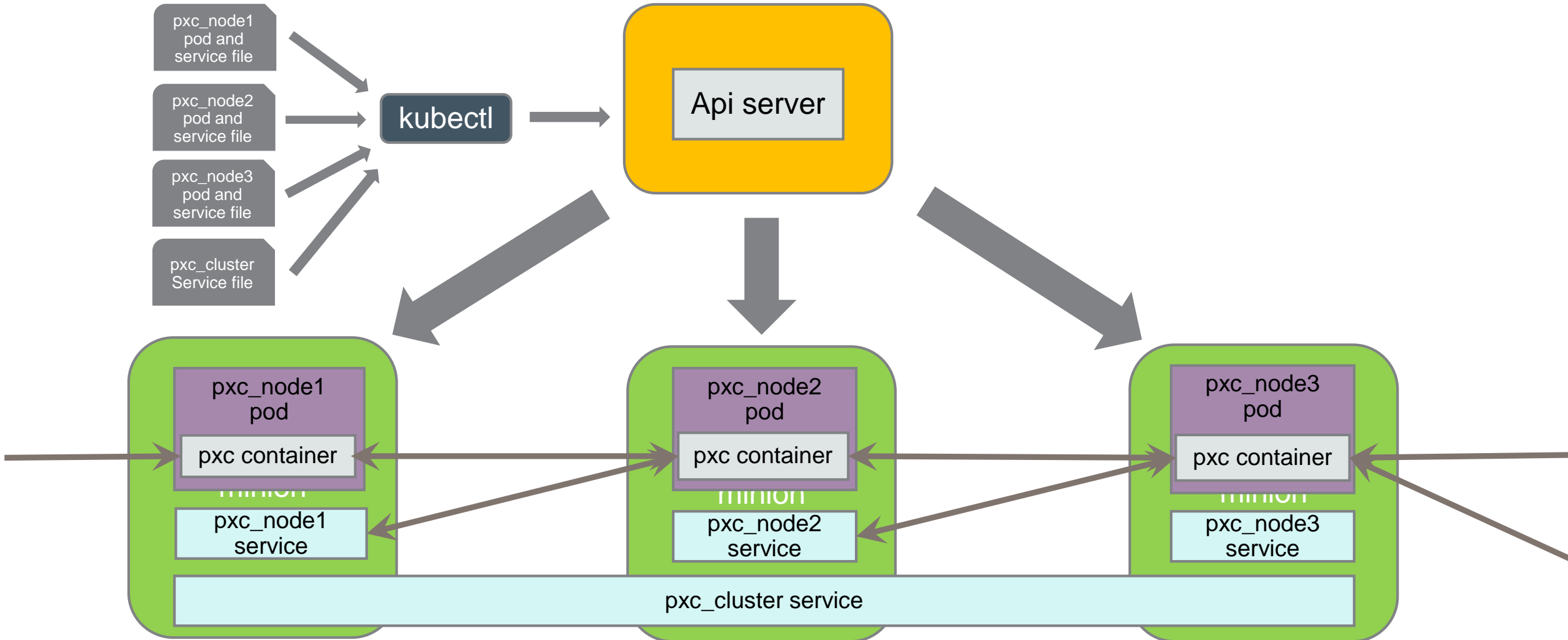
Galera replication on Kubernetes

<https://github.com/kubernetes/kubernetes/tree/master/examples/mysql-galera>

Galera on Kubernetes



Galera on Kubernetes Process



Demo

Demo: set up Kubernetes with Ansible