KUBERNETES

What is it?

An open-source system for automating deployment, scaling, and management of containerized applications.

DATACENTER AS A COMPUTER

The software running on these systems, such as Gmail or Web search services, execute at a scale far beyond a single machine or a single rack: they run on no smaller a unit than clusters of hundreds to thousands of individual servers. Therefore, the machine, the computer, is this large cluster or aggregation of servers itself and needs to be considered as a single computing unit.

The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines

Why is it important?

- Runs anywhere (local, on-premise, hybrid, public cloud)
- Runs anything (stateless, stateful)
- Provides a declarative API for compute resources
  - CPU / RAM
  - Storage
  - Networking

DECLARE YOUR INFRASTRUCTURE

KUBERNETES WORKS TO RESOLVE THAT DECLARATION

KEY CONCEPTS

Pods

- Represents a single instance of an application
- Encapsulates resources
  - Compute (CPU/RAM)
  - Storage
- **All resources in a pod can talk to each other via localhost**

Services

- A powerful virtual network layer
- Can be IP based
  - Selectors/Labels are a better option
- Allows pods to be exposed to each other
- Allows pods to be exposed outside the cluster

Volumes

- Represents a block storage unit
- Backed by lots of storage providers
  - s3, gcp, nfs, cephfs, etc.
- Both ephemeral and persistent volumes supported

Namespaces

- Allows you to 'virtualize' your cluster
- Can separate your cluster by environment or use case
- Can separate your cluster by authenticated user
- Can define quotas to manage resource usage

LIVE DEMO

OUR APP

- Database with a persistent volume
- Web application that connects to the database
  - Horizontally scalable :)
- Load balancer with HTTPS
  - cant be demoed locally :(