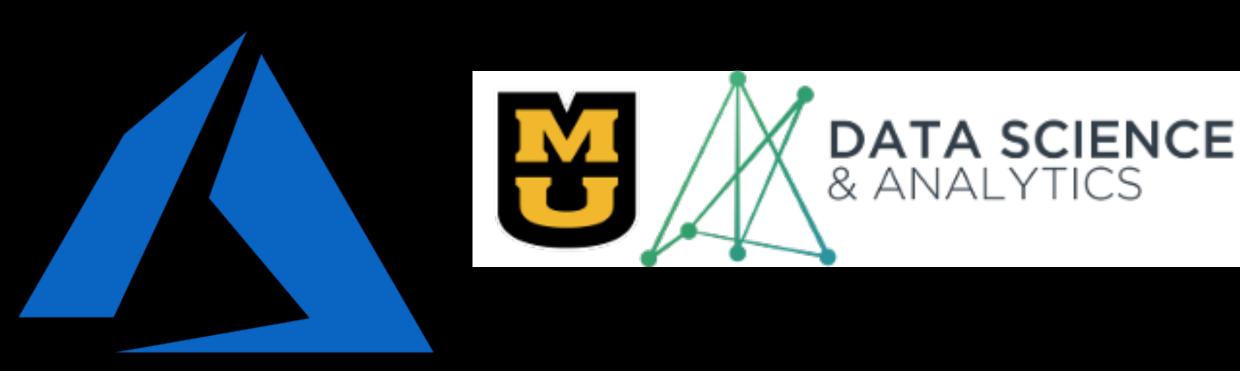
Data Science Powered by Kubernetes and Microservices in Azure



Outline

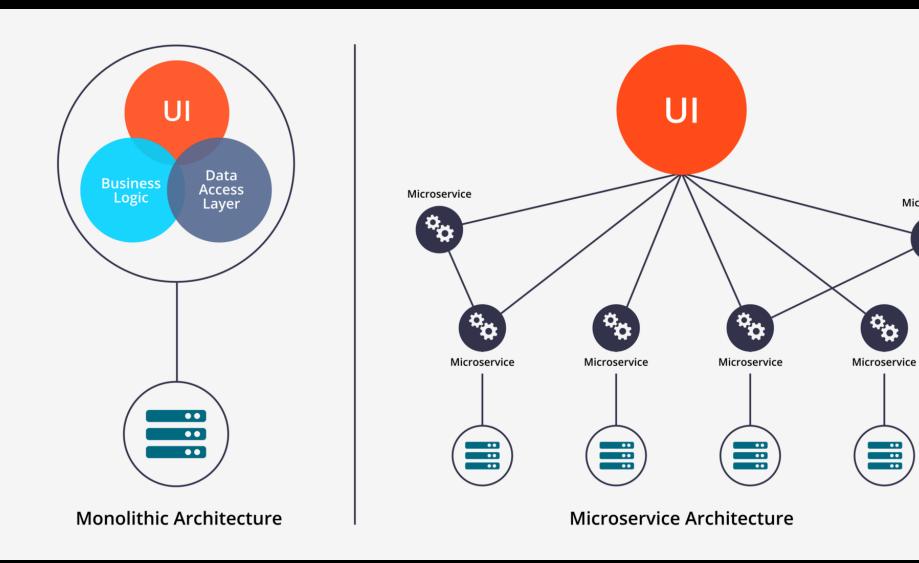
- What are Microservices
- Managing Microservices
- Azure Cloud Computing
- Hands On Lab: Time Series Analysis of Crypto-Currencies

Extract Transform and Load (ETLing)

- Extraction: Accessing data sources (APIs, CSVs, IoT devices)
 - Usually requires multiple data sources
 - Data is never clean
- Transforming: Cleaning the data
 - Fixing empty values
 - Normalizing values
- Loading: Getting the data to a usable storage locations
 - Picking the correct data store (SQL, NoSql, TimeSeries, Graph)

Microservices (SOA)

Microservice



Microservices - Advantages

Multiple languages

Fault Resistant

Small Teams

Scalability

Microservices Disadvantages

Complexity

Networking

Communication

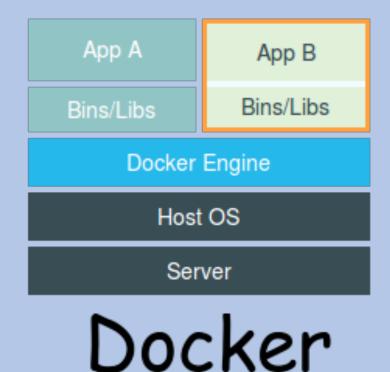
Knowledge Overhead



docker

VMs vs Docker





docker run

Image

is instance of

Docker Container

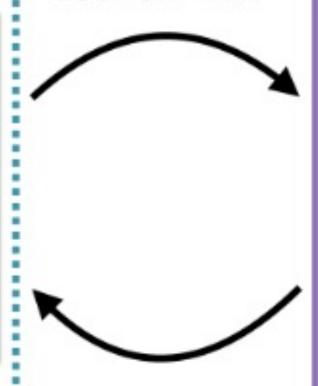
docker ps -a to view all containers

Host OS

docker images to view images

Image Registry

i.e. hub.docker.com



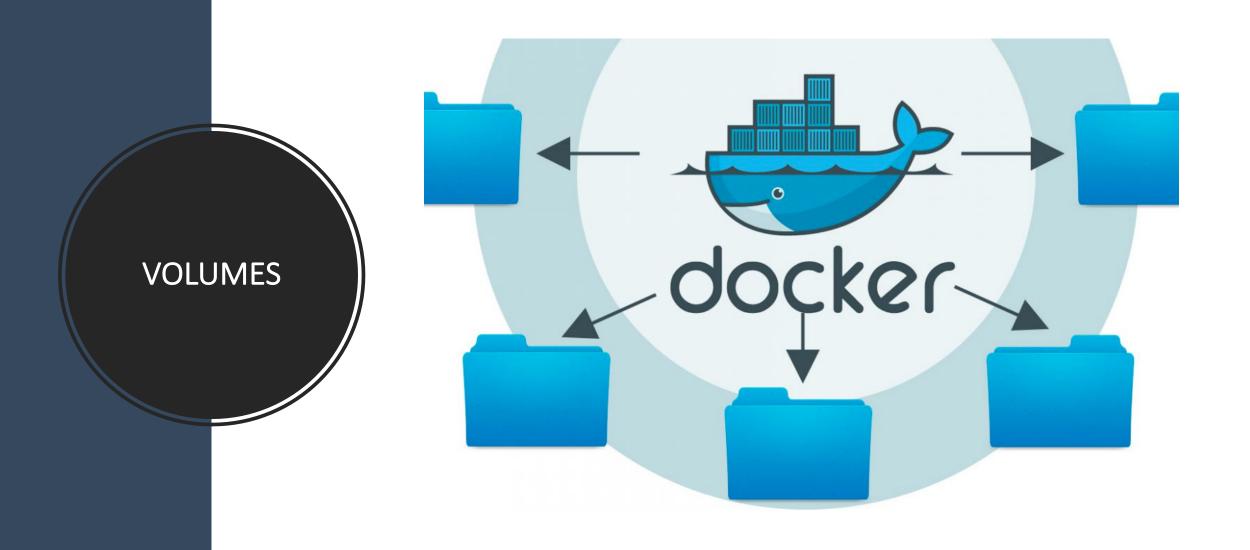


DOCKERFILE

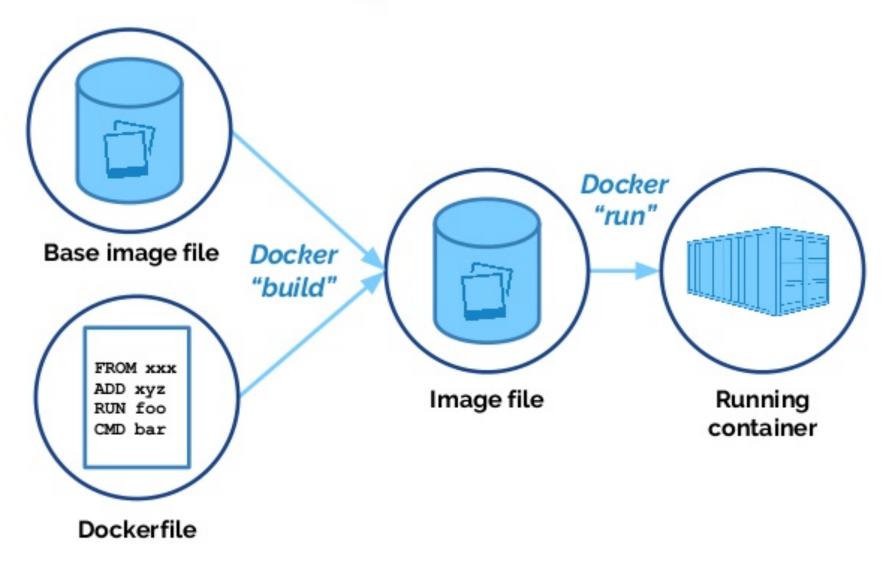








Docker images and containers







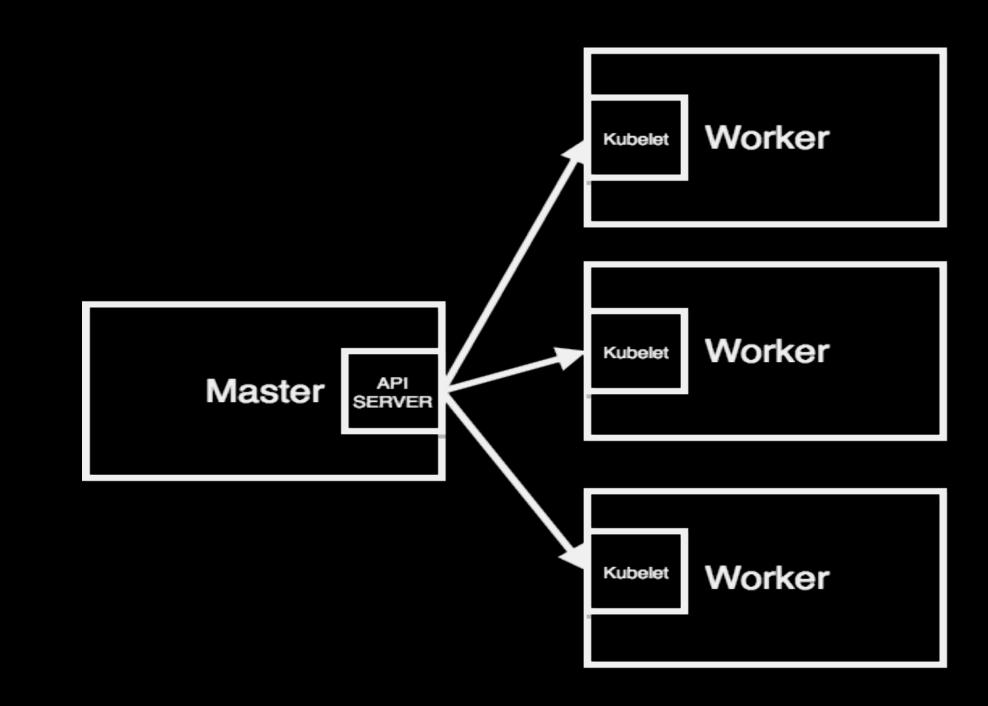


kubernetes

Container Orchestration

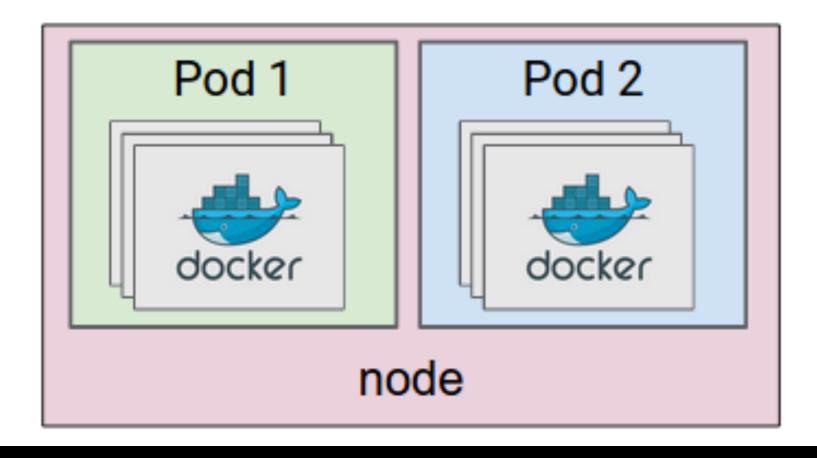
Benefits

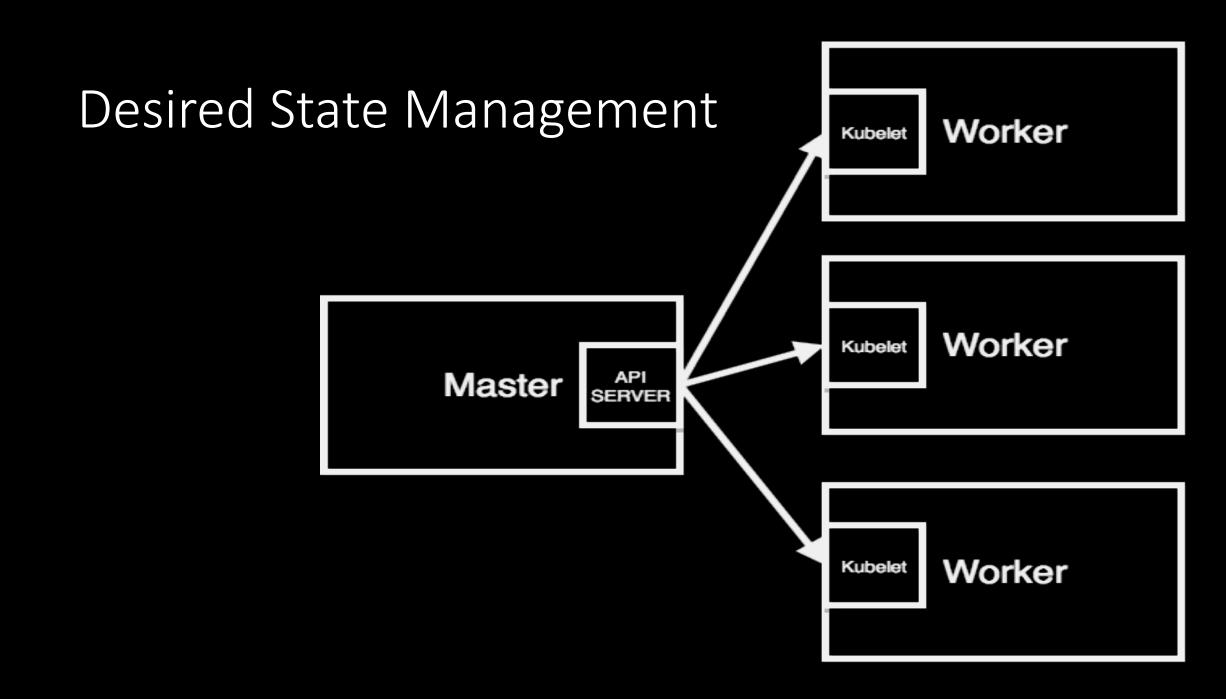
- High scaling without worrying about the hardware
- Health checking cycles
- Easy rolling updates
- Heterogeneous Infrastructure

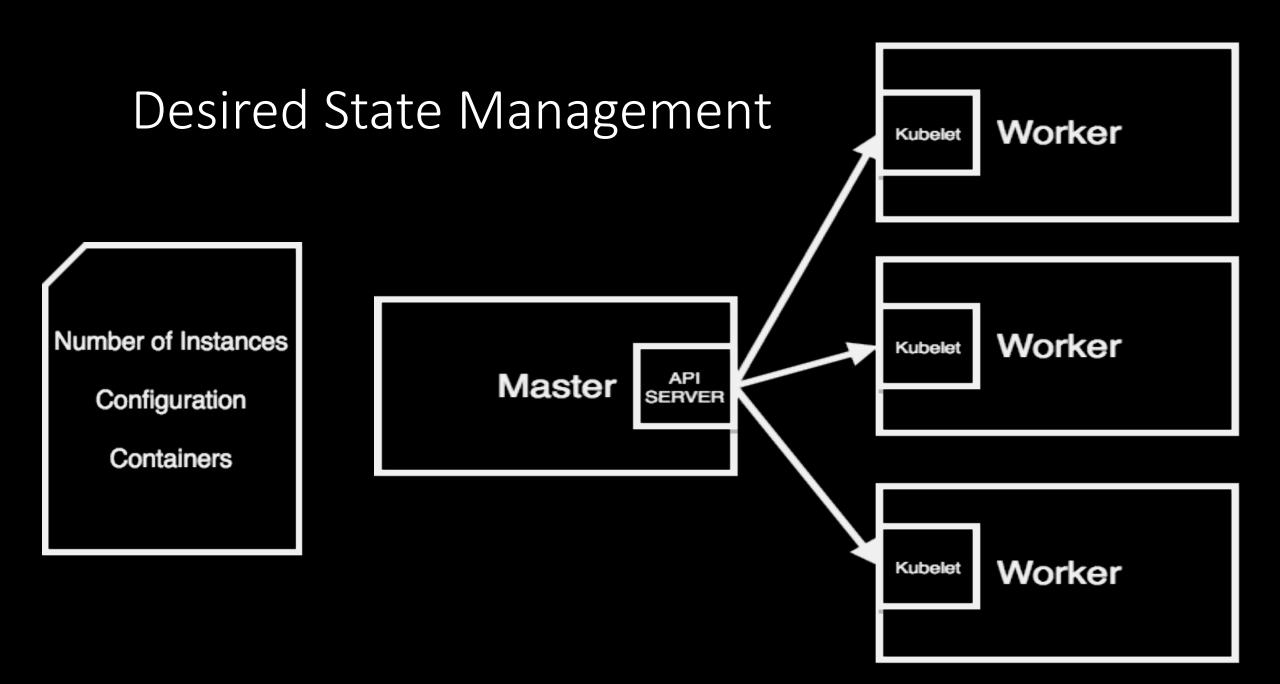


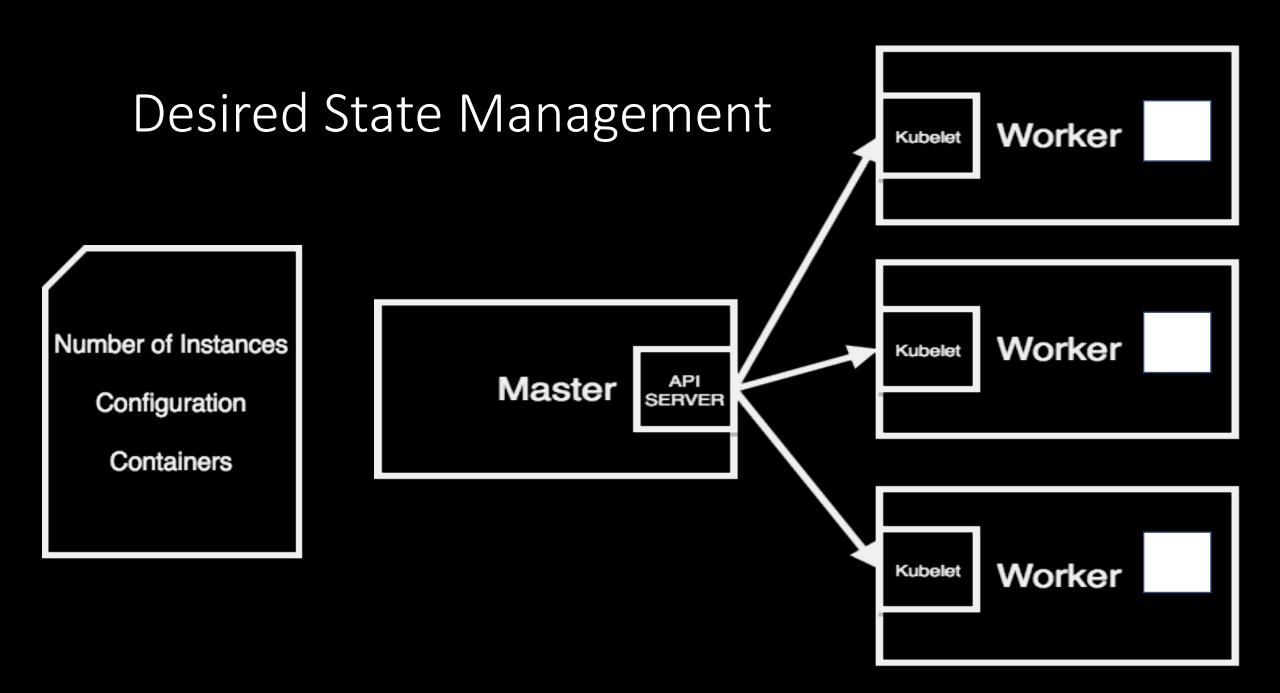
Kubernetes Pods

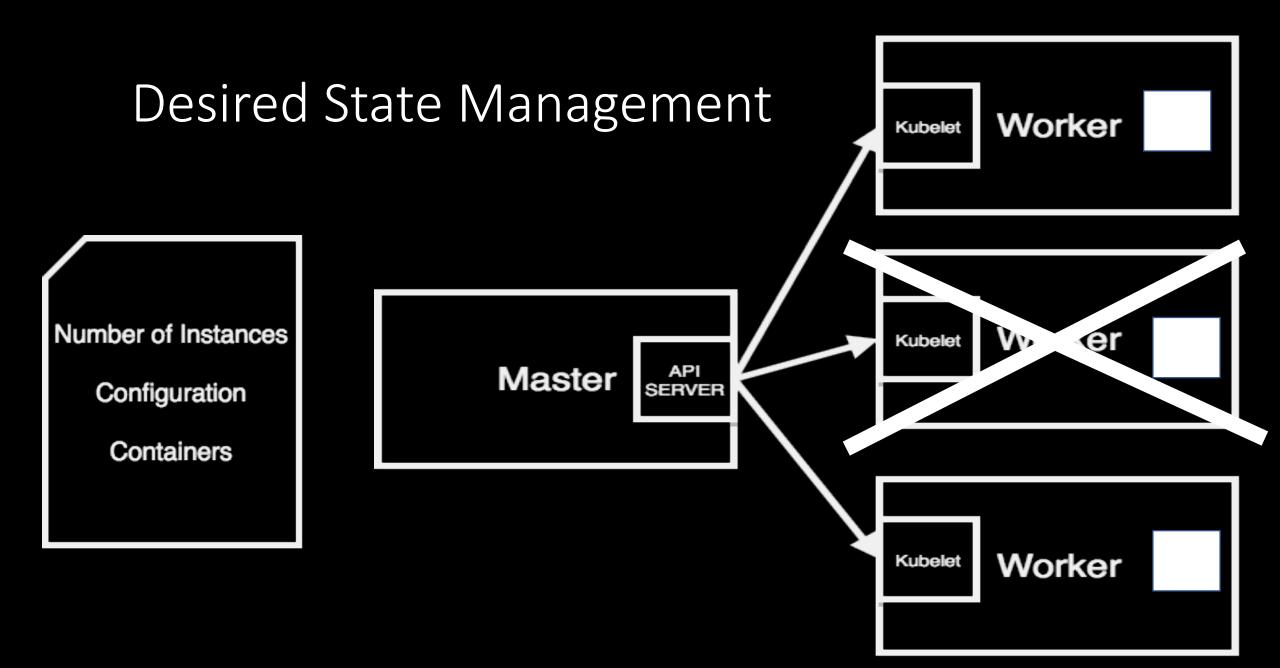
collections of containers that are co-scheduled

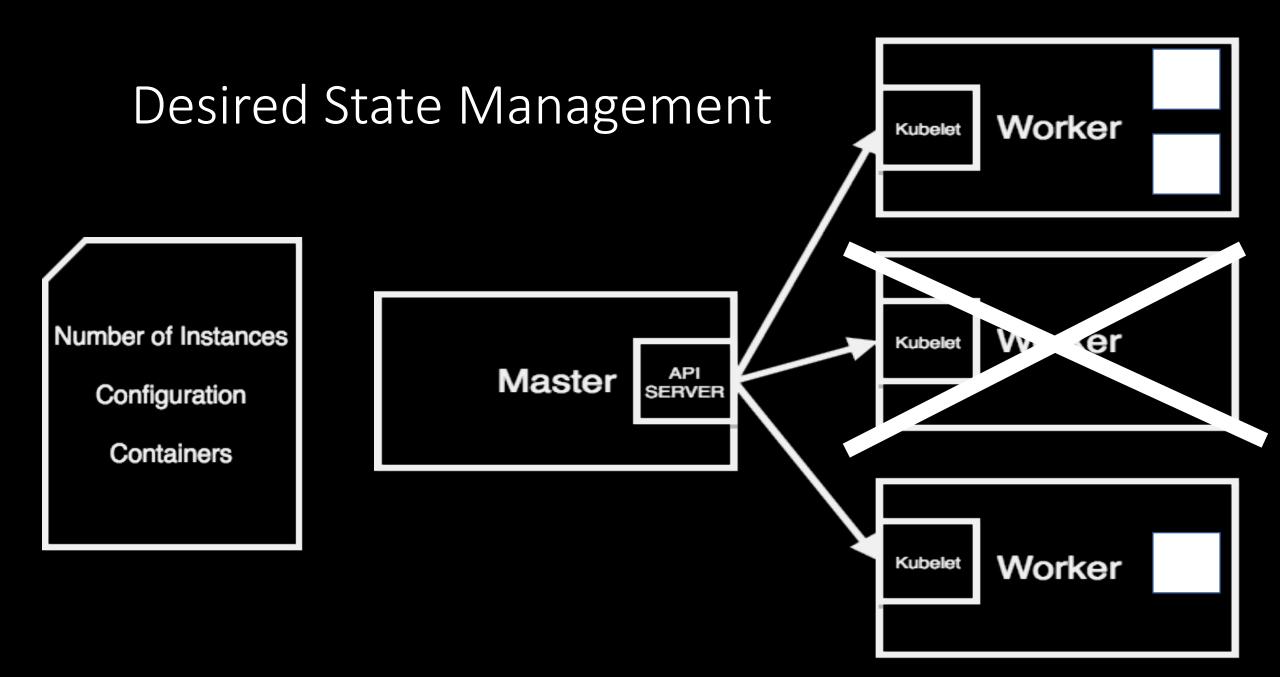




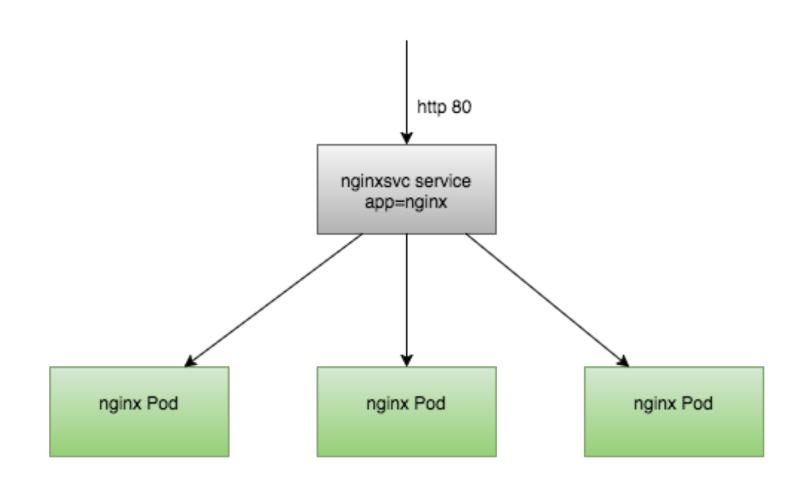


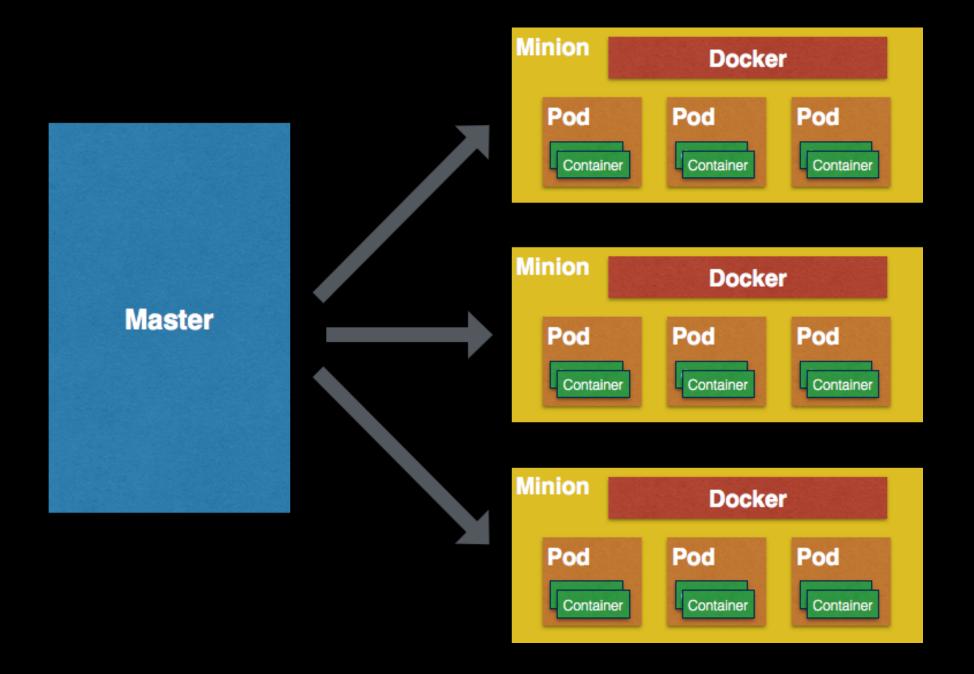






Kubernetes Services





More Reasons to Love

- Rolling Updates
- Care about the application not how its run
- Worry "less" about scaling
- Package manager (HELM)

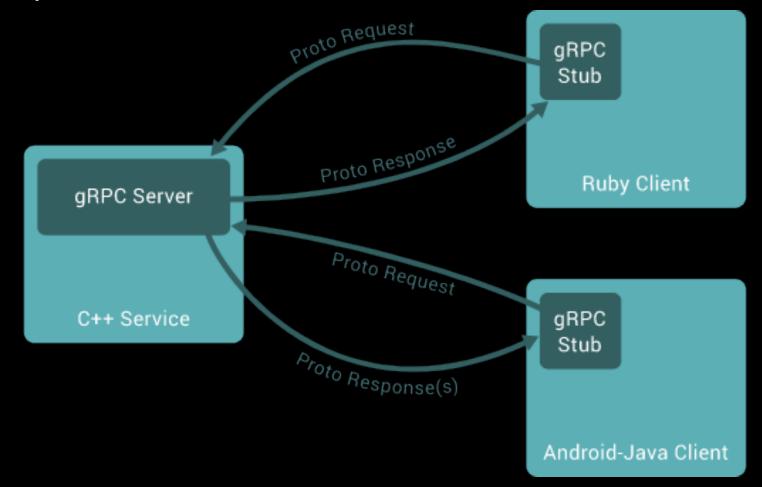
DEMOS

Kubernetes – Basic Operations

Microservice Communication

• Http works good for many tasks but we can do better.





Code: https://github.com/samkreter/DSA-workshop

Twitter: @samkreter

Github: samkreter

Hands On Lab: Time Series Analysis of Crypto-Currencies and Precious Metals

- Architecture Overview
- Technologies



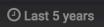












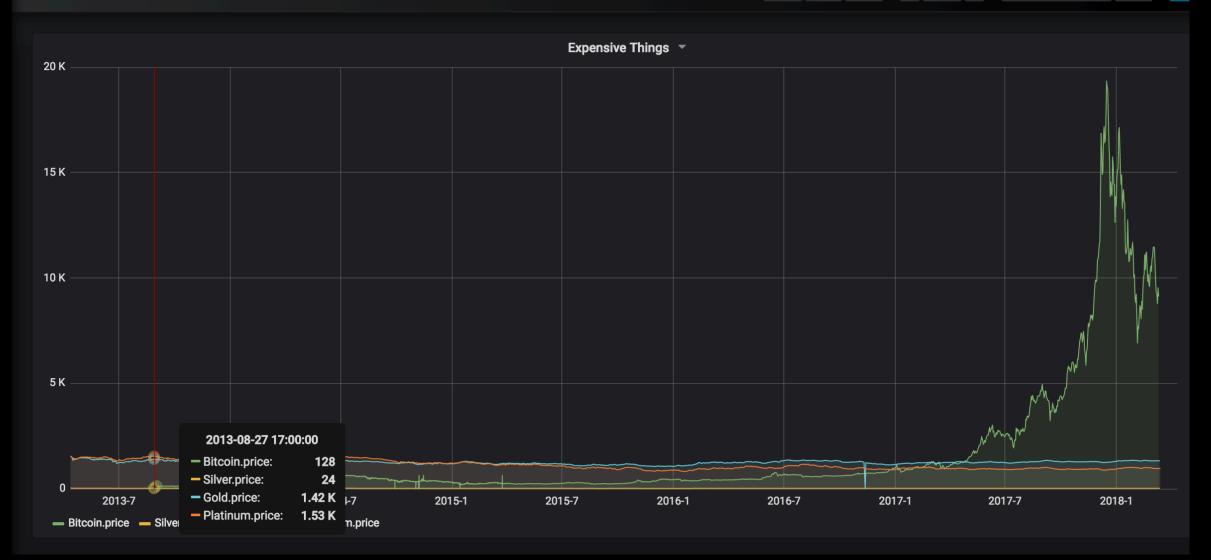
















Streaming Architecture Overview Source Cluster Steaming API InfluxDB Web browser UI

