

# Accelerating Humanitarian Relief with Kubernetes

Building a highly available native spark pipeline on k8s

Erik Schlegel - Microsoft – Principal Engineering Manager

Christoph Schittko - Microsoft – Principal Engineer

# Us

## **Erik Schlegel**

- Principal Engineering Manager
- Focus on emerging technology projects with innovative partners
- [github.com/erikschlegel](https://github.com/erikschlegel)
- [@erikschlegel1](https://twitter.com/erikschlegel1)

## **Christoph Schittko**

- Principal Engineer
- Chronic Early Adopter
- [github.com/xtophs](https://github.com/xtophs)
- [@CTheArchitect](https://twitter.com/CTheArchitect)

# What we do

**PARTNER** with innovators to unblock and accelerate their ideas leveraging both open source and Microsoft technology.

**GENERALIZE** some or all of the solution as patterns, reference designs, or open source software.

**SCALE** our learning and patterns through product team collaboration, blogs, code stories, and open source communities.





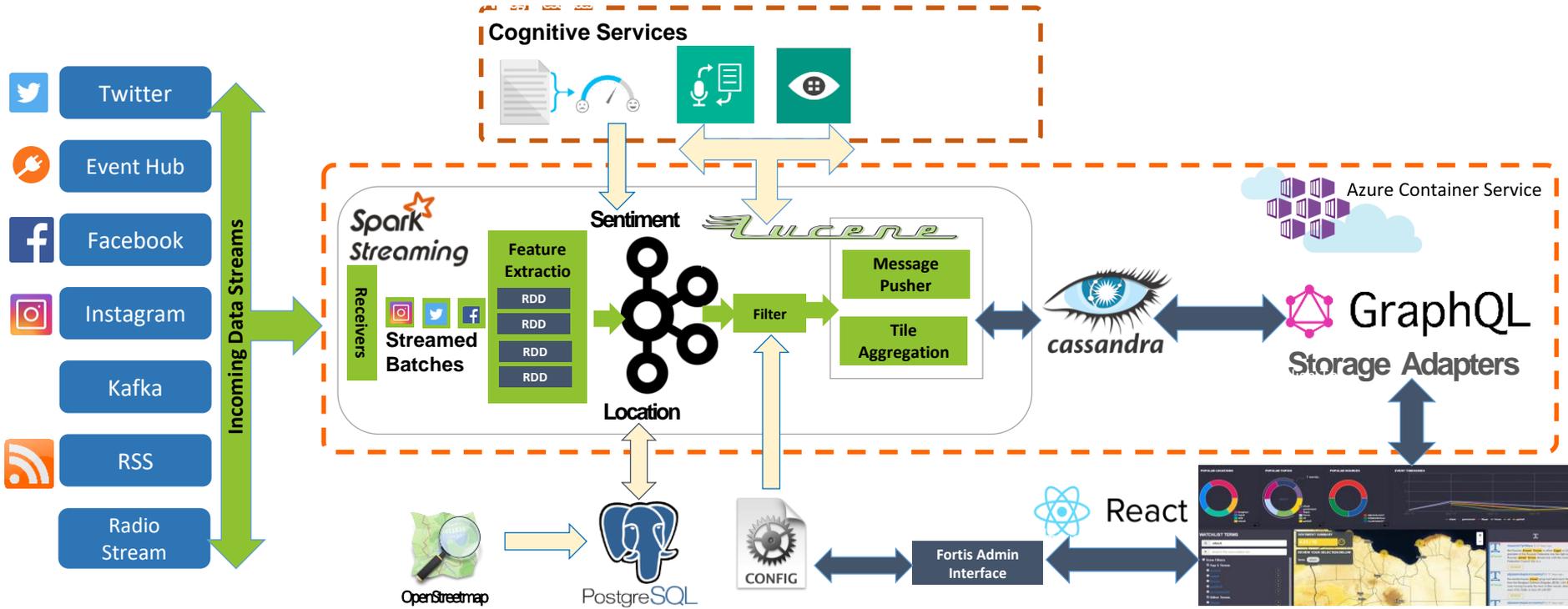
# UN Challenges

- Humanitarian aid plans are manually composed
- Resulting in imprecise aid relief plans
- Aid planners monitor 400+ data sources daily
- Impacted areas often have limited accessibility
- Slow turnaround

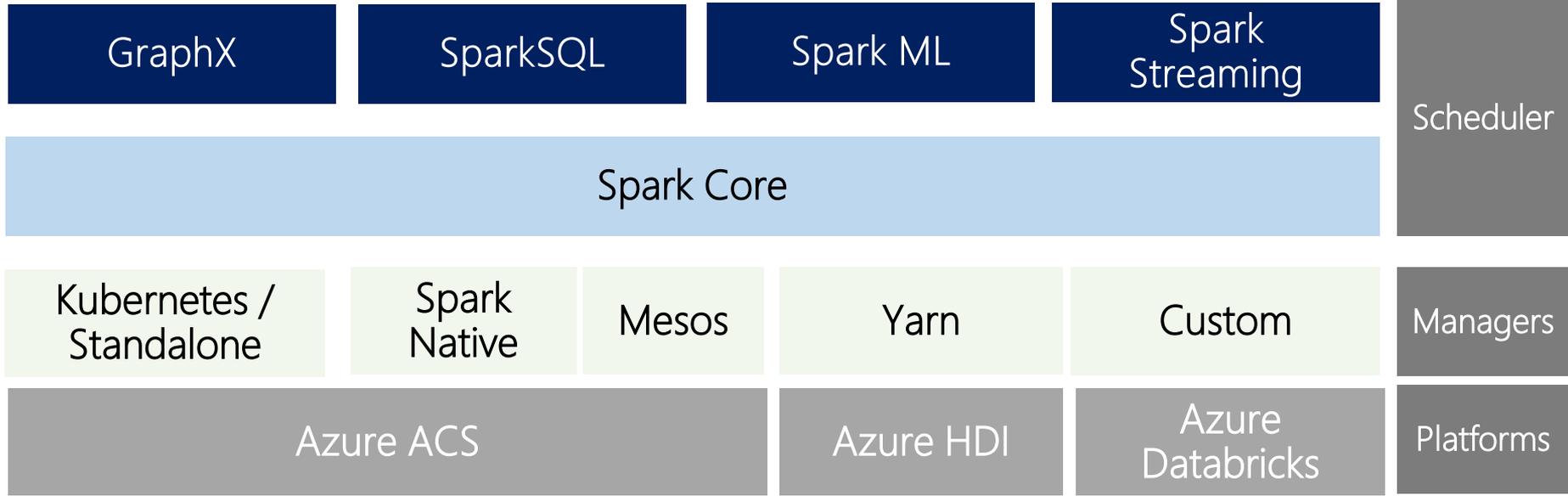
# Project Fortis: Goals

- Accelerate the construction of aid planning
- Improve its data accuracy
- Provide deeper insights and trends
- Real-time analytics
- More intelligence and insight to enable better forecasting

# Functional Architecture



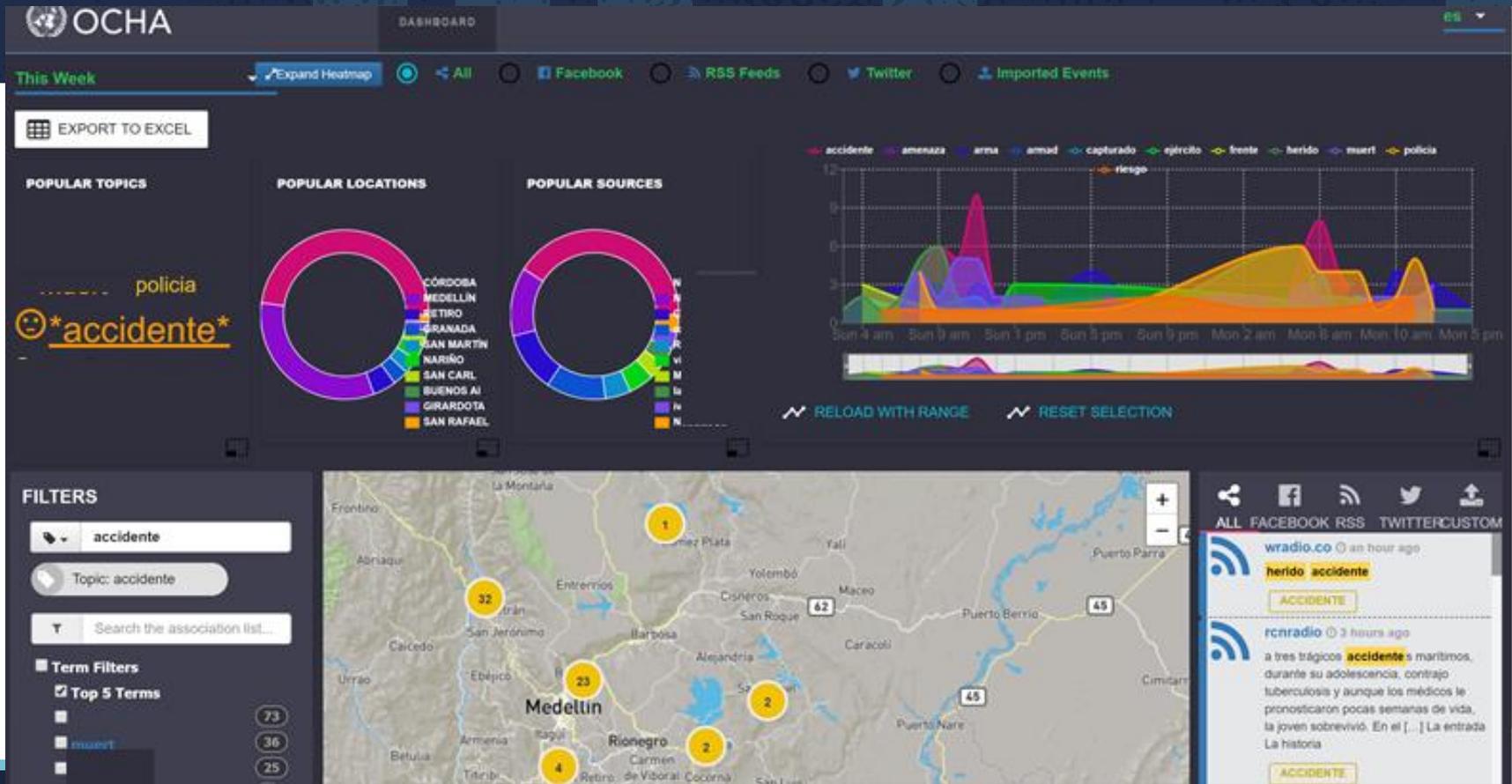
# Spark Cluster Managers



# Spark On Kubernetes

- Resource sharing / utilization across Spark and Cassandra
- Simplified deployment model(ie helm, deis, spark-submit, etc)
- Streamlined developer experience(ie kubectl, etc)
- High Availability – Kubernetes orchestrator replaces Zookeeper
- Elastic executor scaling with kubectl min/max replicas
- Simplify dev ops with distributed env/config map, monitoring & logging

# UN OCHA Deployment

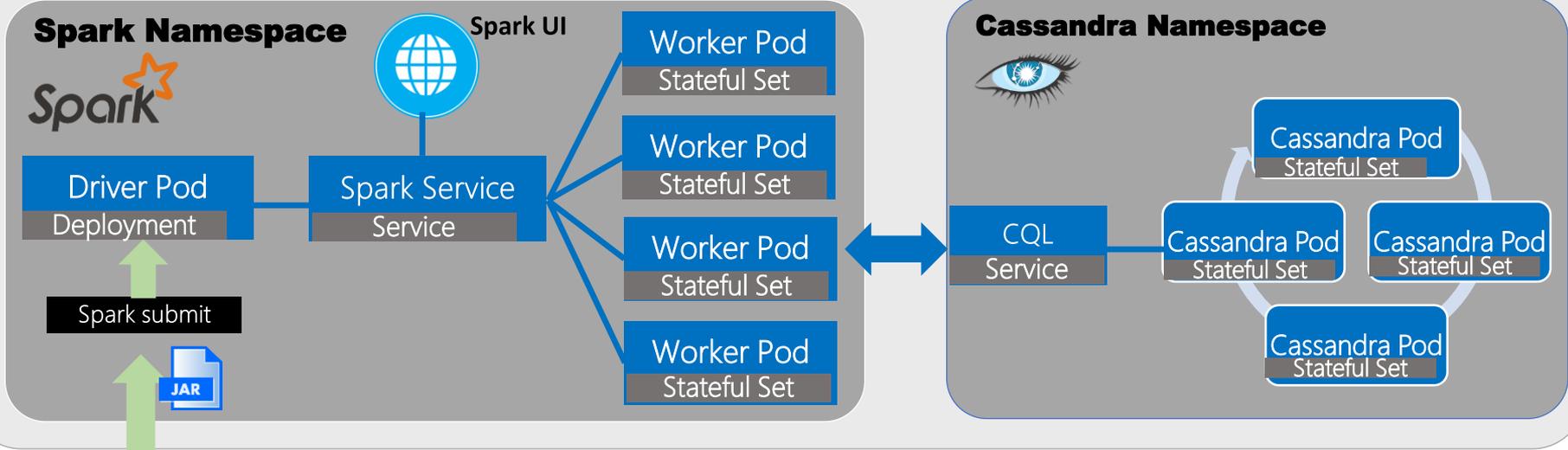


# Fortis k8s Setup

## Kubernetes Cluster



Azure Container Service



## Deployment Process

Blob Store

Published JAR

Assembly

Build

Published Release



Travis CI+

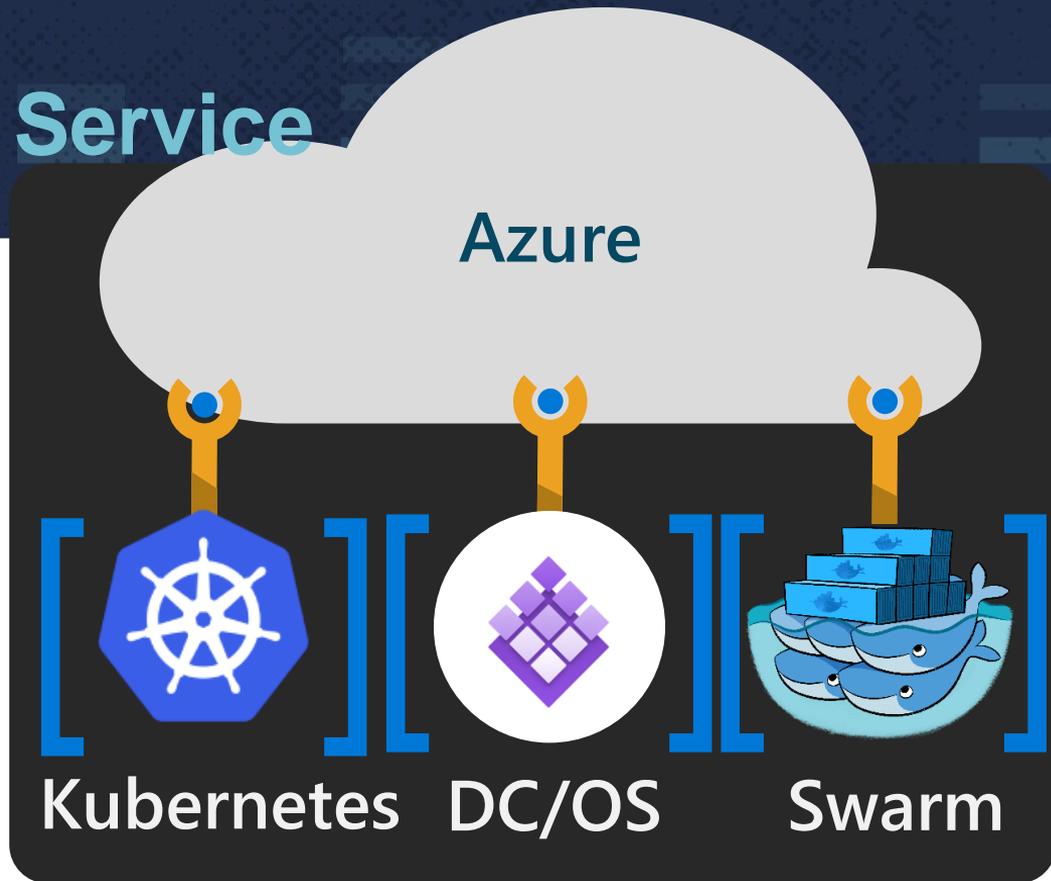


# Azure Container Service

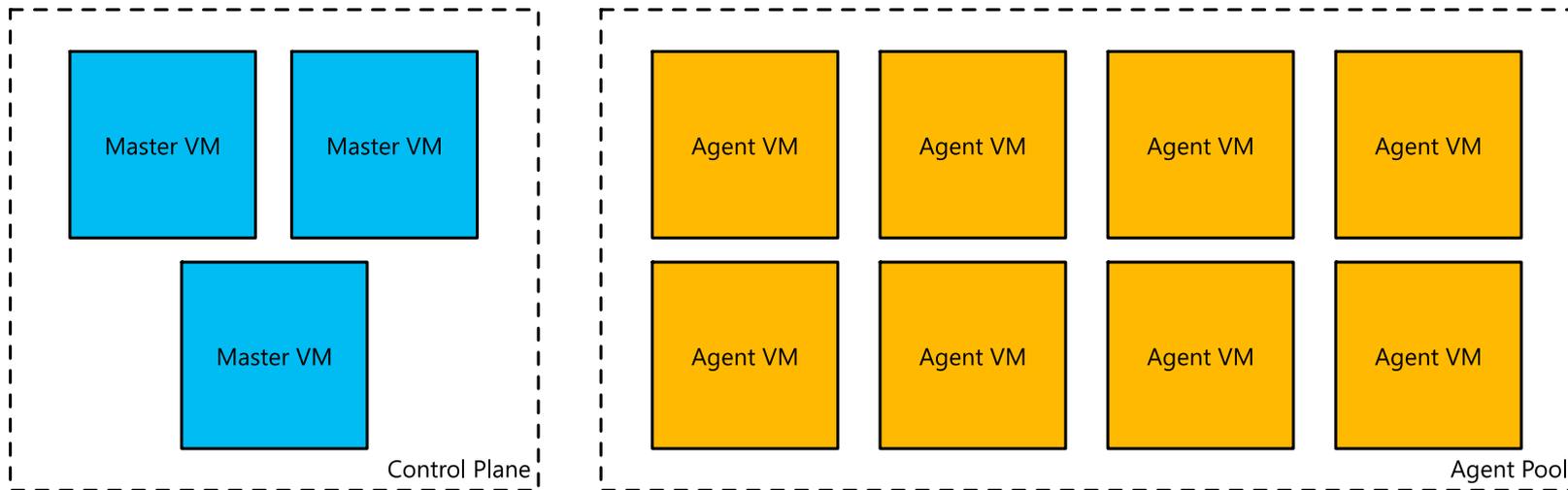
Standard Docker tooling and API support

Provisioning of DC/OS, Docker, and K8s

Linux and Windows Server containers



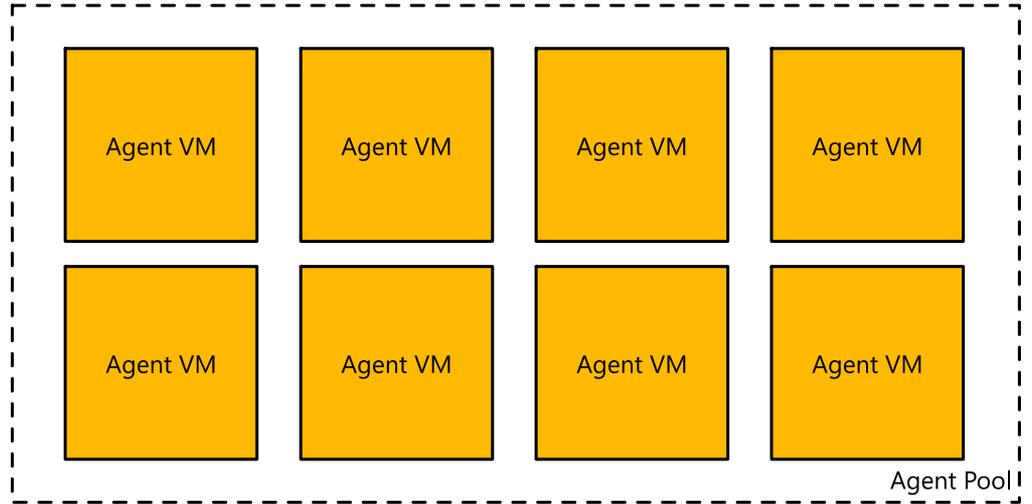
# Kubernetes with ACS



```
$ az acs create -g myResourceGroup -n myCluster \  
  --generate-ssh-keys --orchestrator-type kubernetes \  
  \ Running ..
```

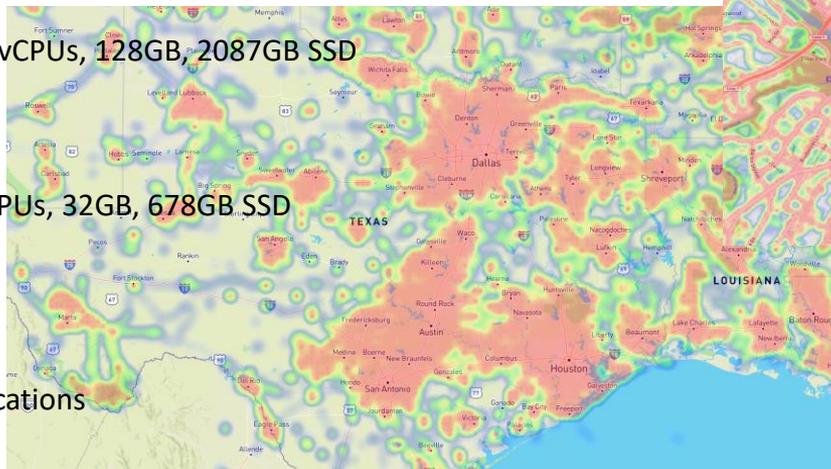


# Kubernetes with AKS



# Networking Benchmark

- Spark (k8s)
  - 64 Core Cluster
    - Single DC
    - Azure L16s, 16vCPUs, 128GB, 2087GB SSD
- Cassandra (VMs)
  - 24 Core Cluster
    - Single DC
    - Azure L4s, 4vCPUs, 32GB, 678GB SSD
- Cassandra Config
  - Replica Factor: 2
  - LOCAL\_QUORUM
- Activity Data
  - 20M activities, 20B locations



Azure CNI

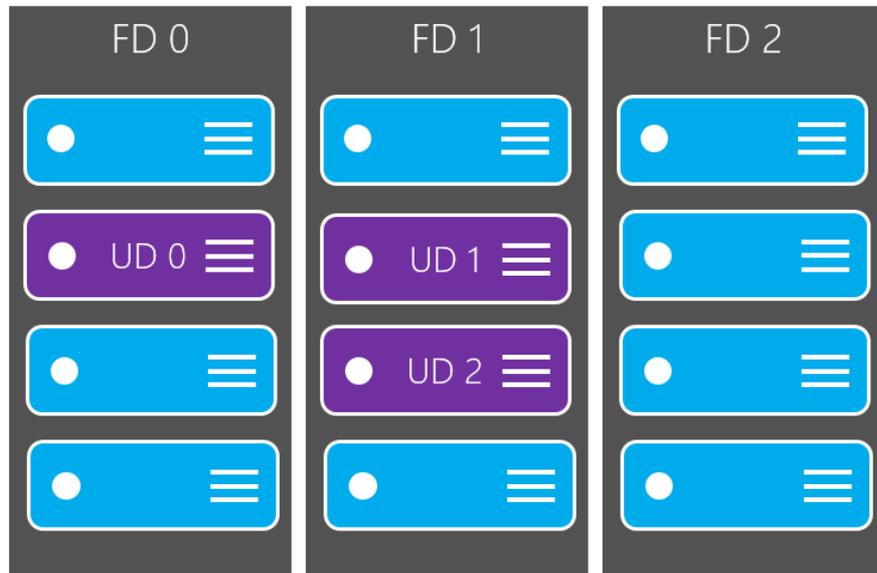
29.86 hrs

Calico

33.76 hrs

# Cassandra Azure HA

- 1 Pod Per Node
  - Node Affinity
- Azure Fault Domain Aware
  - Customized Docker Image
  - Azure Metadata Service
    - Azure Fault Domain: Rack
    - Azure Region: DC
  - GossipingPropertyFileSnitch

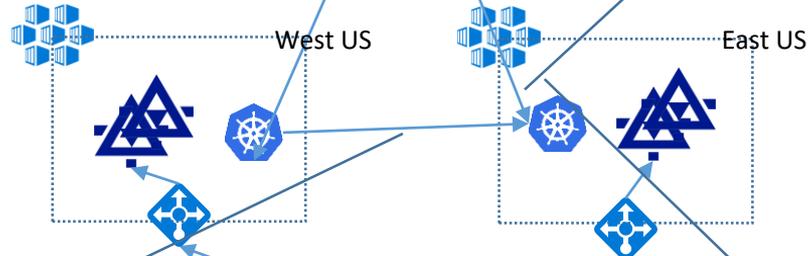


# Multi Azure DC Architecture

Admin

Single Script  
Deployment across 2  
clusters

Seed Node discovery  
across clusters



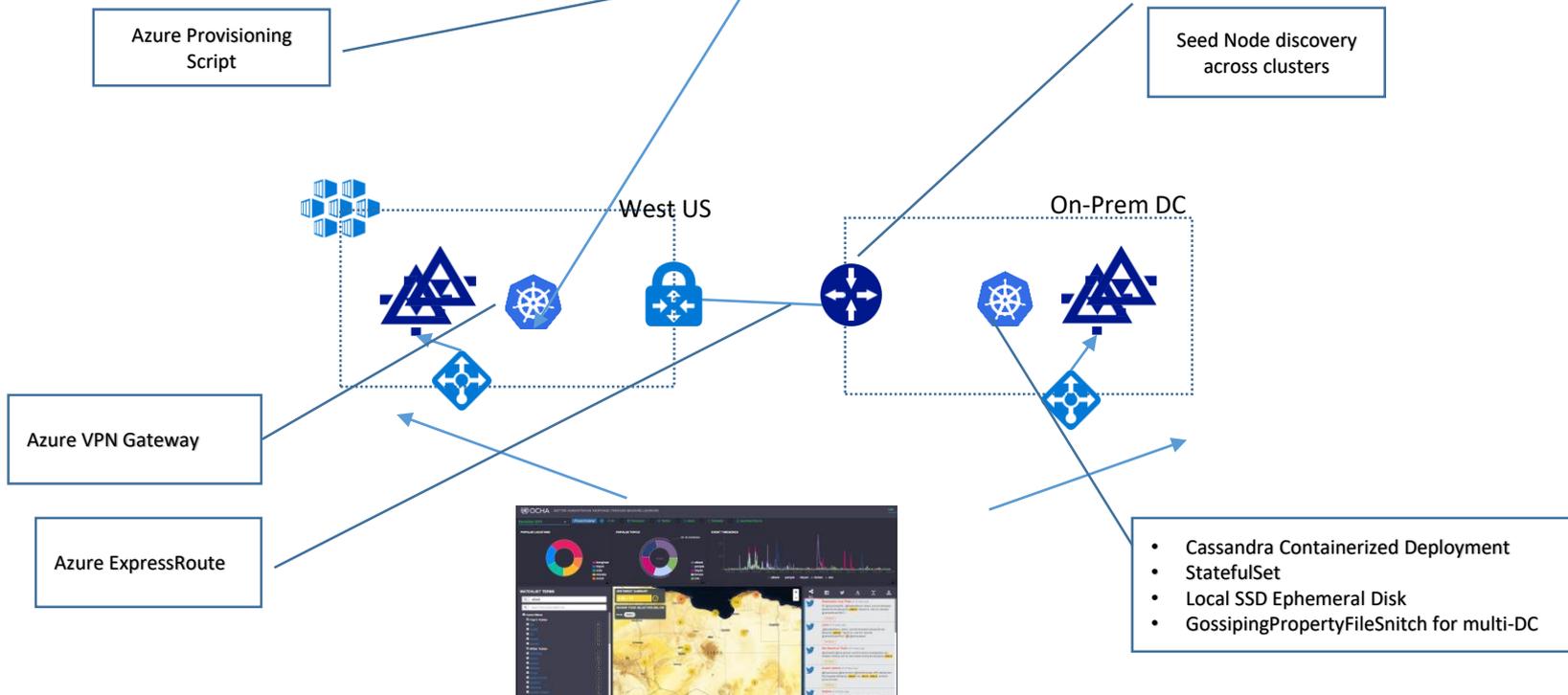
Global VNET Peering  
10 – 40ms latency

- Cassandra Containerized Deployment
- StatefulSet
- Local SSD Ephemeral Disk
- GossipingPropertyFileSnitch for multi-DC



# Hybrid Cloud Setup

Admin

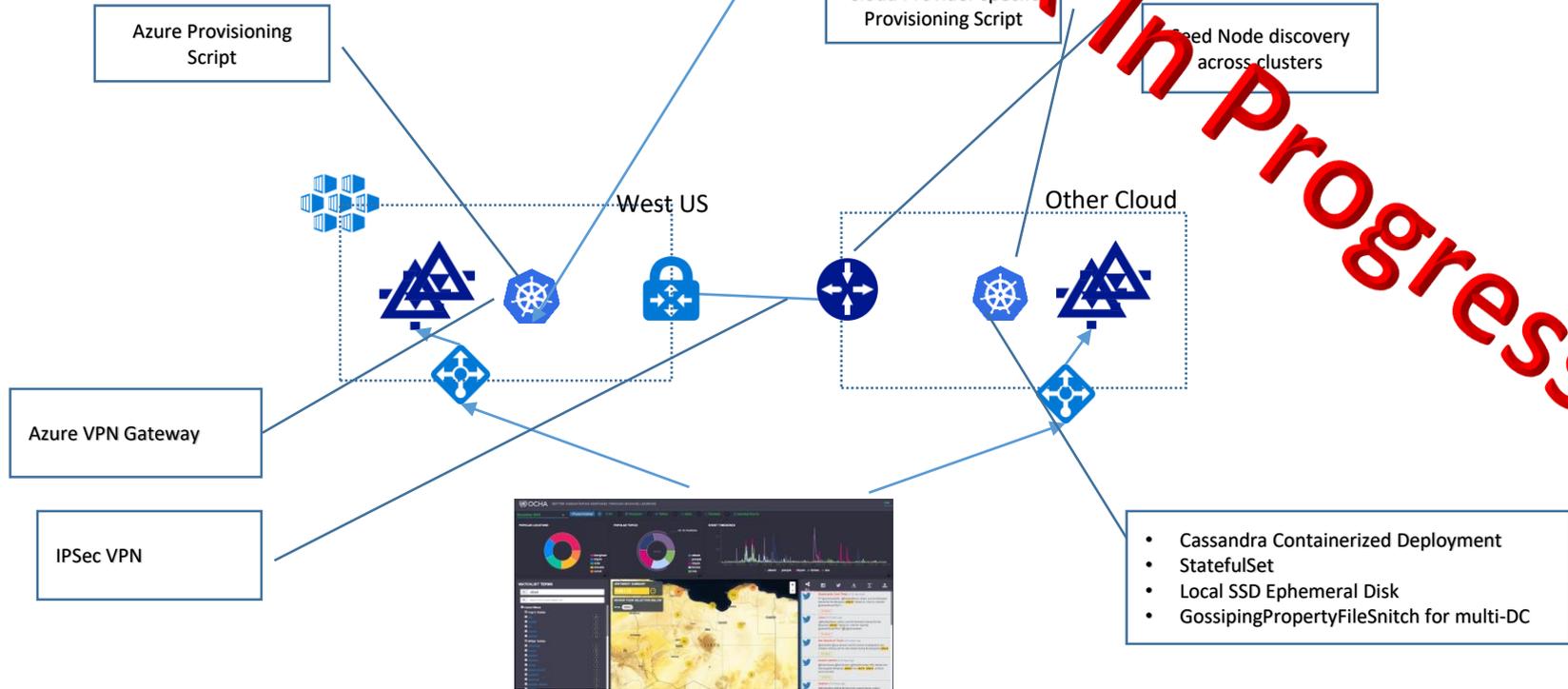


# Multi Cloud Setup

Admin



**Work In Progress**



# K8s Federation or DIY

Federation	DIY
In Flux	You can do it!
Limited Production Use	Leverage built-in HA (cloud, packages)
Automated DNS Management	Leverage Cloud Services (Azure Traffic Manager, DNS)
Azure Support not GA	

# Code

- Cassandra / Spark Helm Charts: [github.com/CatalystCode/charts](https://github.com/CatalystCode/charts)
- Multi-DC Cassandra: [github.com/xtophs/acs-k8s-cassandra-multi-dc](https://github.com/xtophs/acs-k8s-cassandra-multi-dc)
- Fortis: [github.com/CatalystCode/project-fortis-pipeline](https://github.com/CatalystCode/project-fortis-pipeline)
- Cross-Cloud Networking: <https://github.com/dcasati/openbsd-on-azure>
- Heatmap Calculations: <https://github.com/timfpark/heatmap>
- Spark Speech To Text:  
<https://www.microsoft.com/developerblog/2017/11/01/building-a-custom-spark-connector-for-near-real-time-speech-to-text-transcription/>
- Azure Container Service Engine: <https://github.com/azure/acs-engine>

# Questions

