KNI IE Architecture document

- Overall Architecture
- Platform Architecture
    - Deployments to AWS
    - Deployments to GCP
- APIs
- Hardware and Software Management
- Licensing

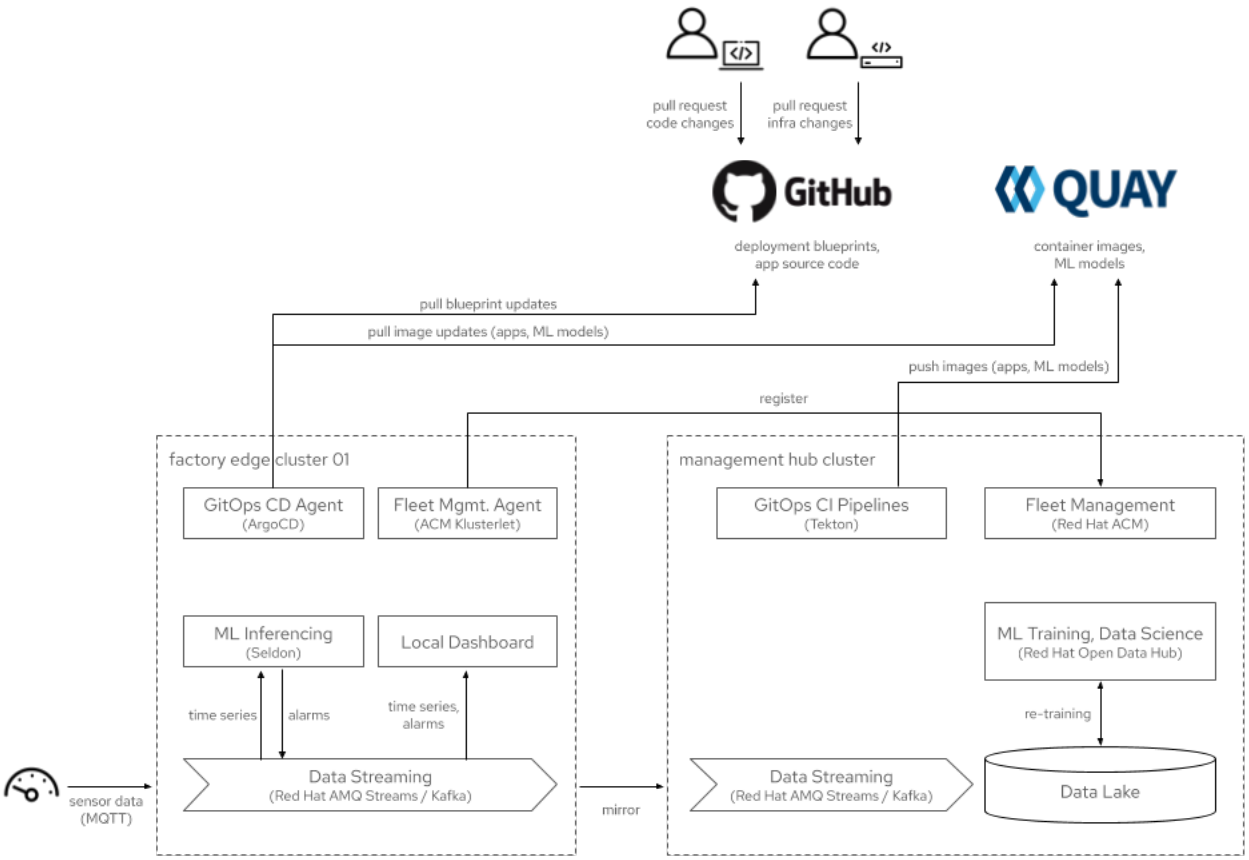
Overall Architecture

The KNI Industrial Edge blueprint consists of at least two sites:

- a management hub and
- one or more factory edge sites.

The management hub consists of a 3-master, 3-worker cluster running Open Cluster Management for managing edge site clusters, applying upgrades, policies, etc. to them as well as OpenDataHub that allows streaming data from factory edge clusters to be stored in a data lake for re-training of machine learning models. OpenDataHub also deploys Jupyter Notebooks for data scientists to analyse data and work on models. Updated models can be distributed out to the factory via the same GitOps mechanisms used also for updates of the clusters and their workloads. The management hub also deploys Tekton pipelines, which will eventually be used for GitOps based management of edge sites, but is not yet implemented in this release.

A factory edge site consists of a 3-node cluster of schedulable masters. Edge sites are deployed from a minimal blueprint that contains Open Cluster Management's klusterlet agent. When the edge cluster comes up, the klusterlet registers the cluster with the management hub and installs a local ArgoCD instance that references the GitHub repo on which the definition of the services to be installed is hosted. Changes to the services will be automatically pulled into and applied to the edge cluster by ArgoCD. The edge cluster's services include Apache Camel-K for ingestion and transformation of sensor data, Kafka for streaming data, and MirrorMaker for replicating streaming data to the data lake on the management hub. Edge clusters also include Seldon runtime for ML models, which the edge clusters pull from the Quay container registry just like every other container image.



Platform Architecture

This blueprint currently runs on GCP and AWS, but is currently only tested against GCP.

Deployments to AWS

Deployments to GCP

Resources used for the management hub cluster:

nodes	instance type
3x masters	EC2: m4.xlarge, EBS: 120GB GP2
3x workers	EC2: m4.large, EBS: 120GB GP2

Resources used for the factory edge cluster:

nodes	instance type
3x masters	EC2: m4.xlarge, EBS: 120GB GP2

Deployments to GCP

Resources used for the management hub cluster:

nodes	instance type
3x masters	EC2: m4.xlarge, EBS: 120GB GP2
3x workers	EC2: m4.large, EBS: 120GB GP2

Resources used for the factory edge cluster:

nodes	instance type
3x masters	EC2: m4.xlarge, EBS: 120GB GP2

Deployments to Bare Metal

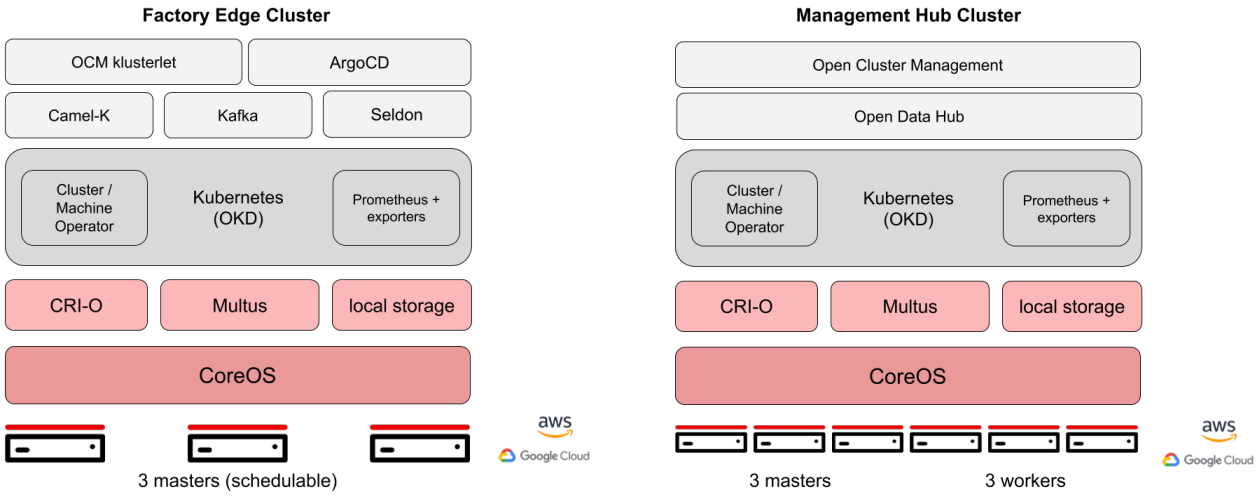
Resources used for the factory edge cluster:

nodes	requirements
3x masters	12 cores, 16GB RAM, 200GB disk free, 2 NICs (1 provisioning+storage, 1 cluster)

The blueprint validation lab uses 3 SuperMicro SuperServer 1028R-WTR (Black) with the following specs:

Units	Type	Description
2	CPU	BDW-EP 12C E5-2650V4 2.2G 30M 9.6GT QPI
8	Mem	16GB DDR4-2400 2RX8 ECC RDIMM
1	SSD	Samsung PM863, 480GB, SATA 6Gb/s, VNAND, 2.5" SSD - MZ7LM480HCHP-00005
4	HDD	Seagate 2.5" 2TB SATA 6Gb/s 7.2K RPM 128M, 512N (Avenger)
2	NIC	Standard LP 40GbE with 2 QSFP ports, Intel XL710

Software Platform Architecture



Release 4 components:

- OKD v4.5 GA
- CRI-O: v1.18
- Multus-cni: version:v4.3.3-202002171705, commit:d406b4470f58367df1dd79b47e6263582b8fb511
- Open Cluster Management: v2.0
- ArgoCD Operator: v0.0.11
- OpenShift Pipelines Operator: v1.1.1
- OpenDataHub Operator: v0.6.1

APIs

No specific APIs involved on this blueprint. It relies on Kubernetes cluster so all the APIs used are Kubernetes ones.

Hardware and Software Management

Licensing

- Apache license