
KubeCon
CloudNativeCon
North America 2024

KubeCon | CloudNativeCon

North America 2024

Experience in Designing & Implementing a Cloud Native Framework for Farm Data Analytics

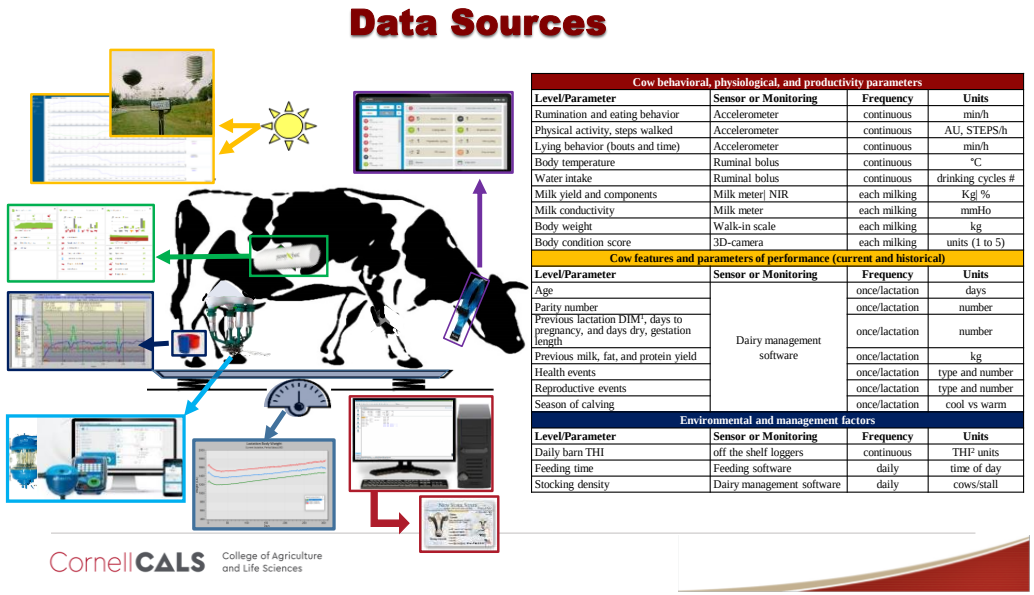
Braulio Dumba, IBM Research
Gloire Rubambiza, Cornell University

Motivation

- The promise of data-driven farming or digital agriculture (DA)
  - Sustainable intensification of farm yields and efficiency
  - Financial, environmental, and societal impacts

Motivation

- Digital agriculture platform: commercial dairy farm
  - 17 months processing tens of gigabytes of data on 1500 cows



Cow behavioral, physiological, and productivity parameters			
Level/Parameter	**Sensor or Monitoring**	**Frequency**	**Units**
Rumination and eating behavior	Accelerometer	continuous	min/h
Physical activity, steps walked	Accelerometer	continuous	AU, STEPS/h
Lying behavior (bouts and time)	Accelerometer	continuous	min/h
Body temperature	Ruminal bolus	continuous	°C
Water intake	Ruminal bolus	continuous	drinking cycles #
Milk yield and components	Milk meter| NIR	each milking	Kg| %
Milk conductivity	Milk meter	each milking	mmHo
Body weight	Walk-in scale	each milking	kg
Body condition score	3D-camera	each milking	units (1 to 5)
Cow features and parameters of performance (current and historical)			
Level/Parameter	**Sensor or Monitoring**	**Frequency**	**Units**
Age	Dairy management software	once/lactation	days
Parity number		once/lactation	number
Previous lactation DIM[1], days to pregnancy, and days dry, gestation length		once/lactation	number
Previous milk, fat, and protein yield		once/lactation	kg
Health events		once/lactation	type and number
Reproductive events		once/lactation	type and number
Season of calving		once/lactation	cool vs warm
Environmental and management factors			
Level/Parameter	**Sensor or Monitoring**	**Frequency**	**Units**
Daily barn THI	off the shelf loggers	continuous	THI[2] units
Feeding time	Feeding software	daily	time of day
Stocking density	Dairy management software	daily	cows/stall

Image credit: Martin Perez

Motivation

- Digital agriculture platform: commercial dairy farm
  - AI/ML to identify cow heath status

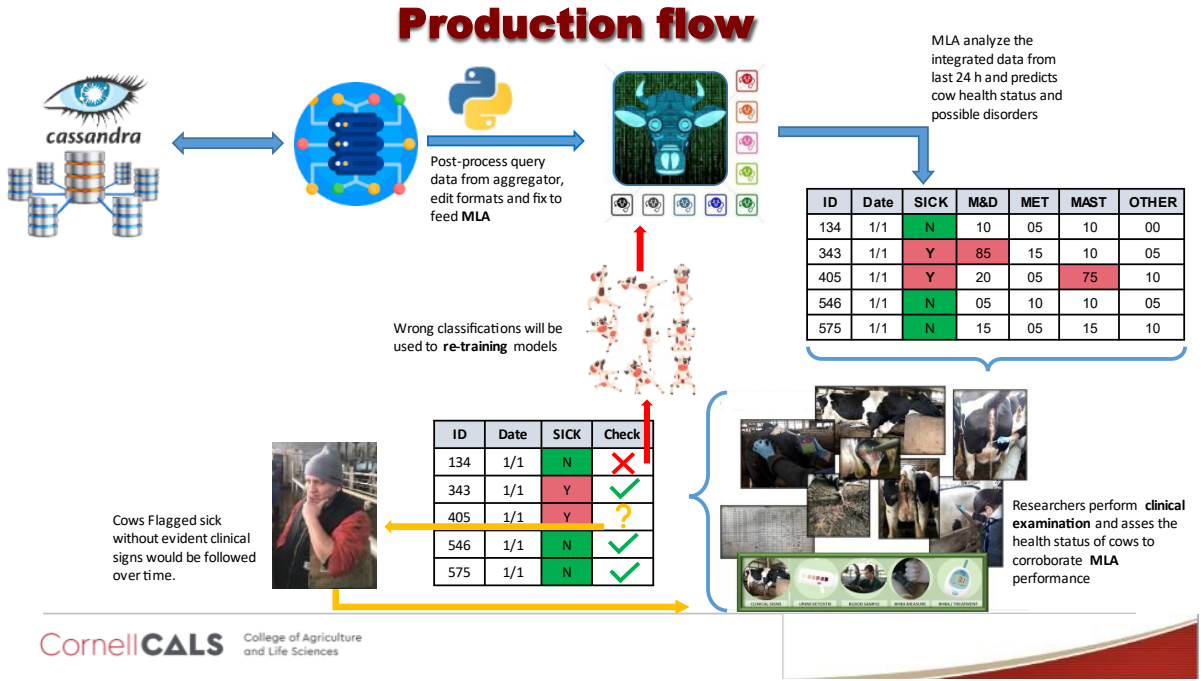


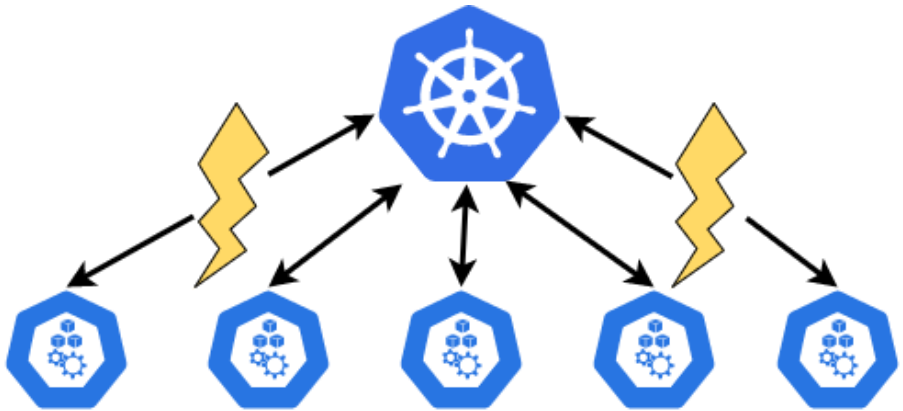
Image credit: Martin Perez

Challenges Managing & Scaling DA

- Managing multiple fields/clusters
- Intermittent connectivity
- Privacy concerns

Challenges Managing & Scaling DA

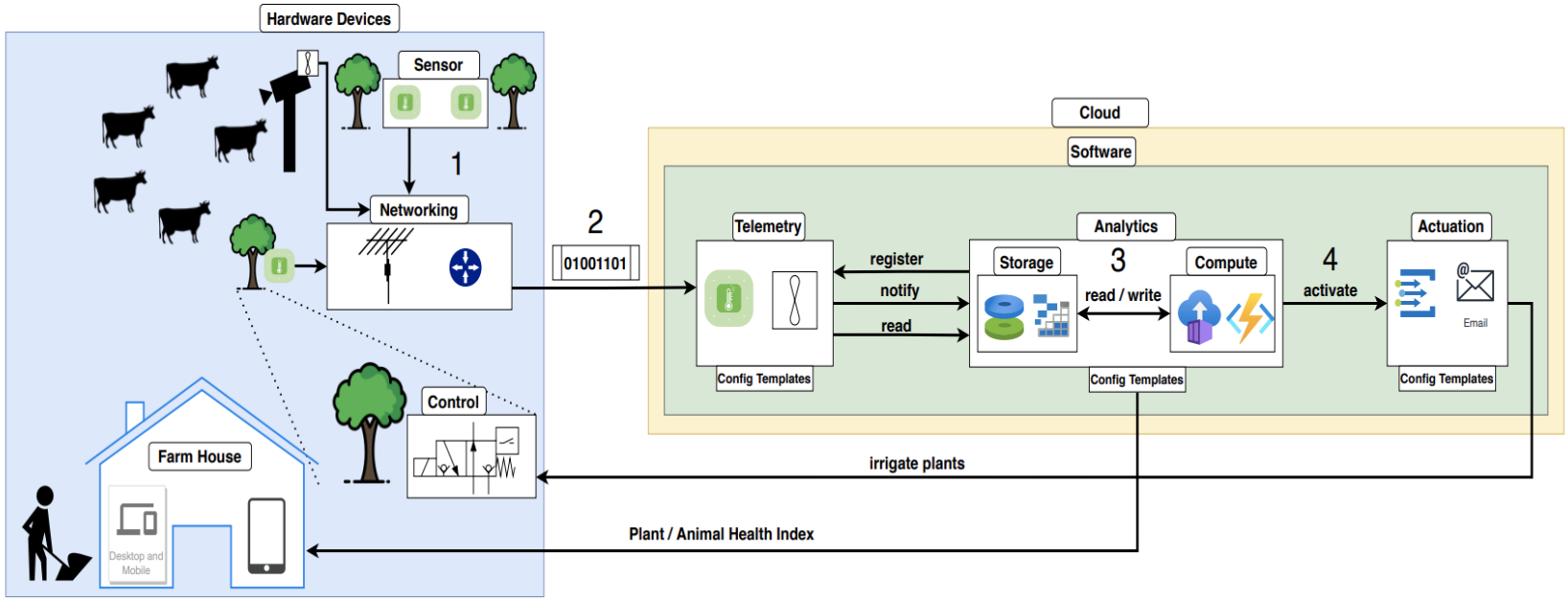
- Managing multiple fields/clusters
- Intermittent connectivity
- Privacy concerns



Similar issues in cloud-native environments

Software Defined Farms - Background

Architecture to transmit and analyze data from networked plants/animals on rural farms



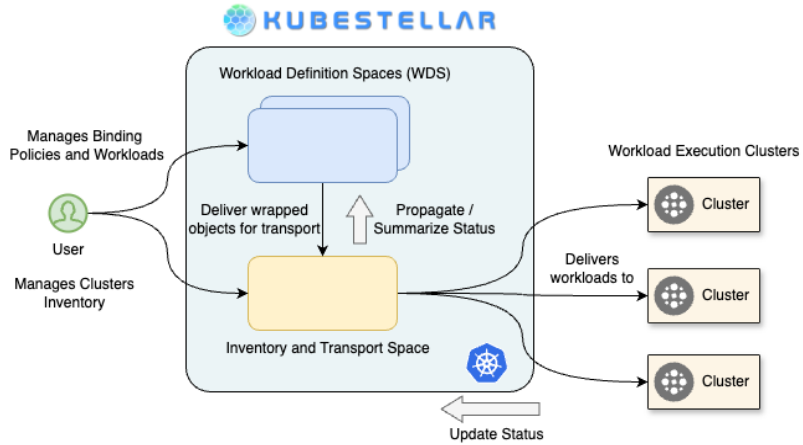
KubeStellar - Background

- A CNCF Sandbox project:
- Deploy, configure and manage workloads across multiple k8s clusters
  - Use “native” k8s interfaces
  - Policy-based placement, workload customization, status summarization
- Find more info: https://docs.kubestellar.io/

KUBESTELLAR

KubeStellar: Architecture Overview

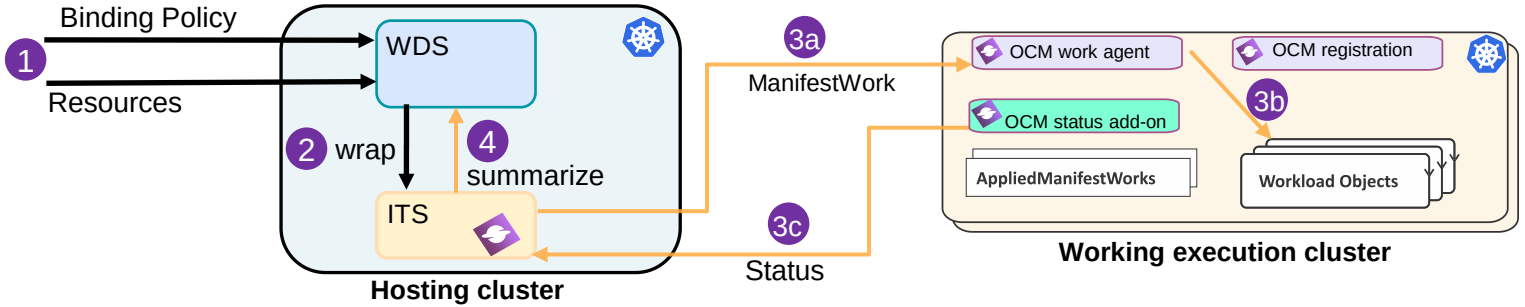
- Based on "spaces" abstraction for isolation and multi-tenancy
  - Space: behaves like a regular k8s cluster. Exposes k8s API endpoint
  - KS can use different types of spaces (e.g., vCluster, api-server only, etc.)
- Pluggable transport framework
  - Propagate resources from WDS to execution cluster(s)
  - Currently using OCM (https://open-cluster-management.io/)
- Collect & summarize statuses from WECs
- Customize resources deployed to WEC



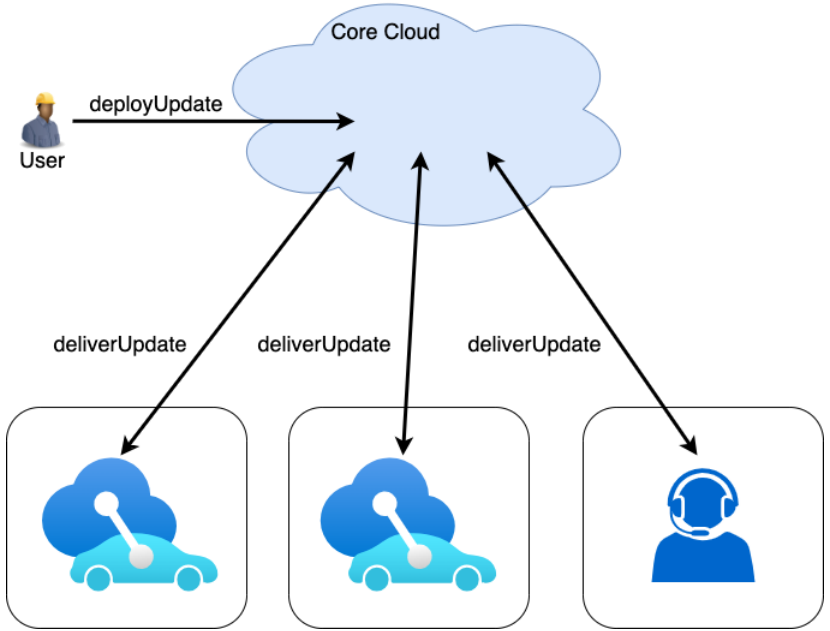
WDS: Workload Distribution Space,
ITS: Inventory and Transport Space,
WEC: Workload Execution Cluster

KubeStellar Example: Deploy a workload

1. User:
   - Deploys workload/resources into WDS
   - Defines the binding policy (desired placement)
2. KS: Transport controller pushes resources into the inventory space (ITS)
   a. Wrap resources into a container object (e.g., ManifestWork)
3. Transport mechanism (OCM)
   a. Distribute wrapped resources into the WEC
   b. Unwrap & deploy resources in the WEC
   c. Returned status through the KS transport status plugin
4. KS: Propagates status into WDS (user facing objects)



A Cloud-Native Framework for Data Analytics



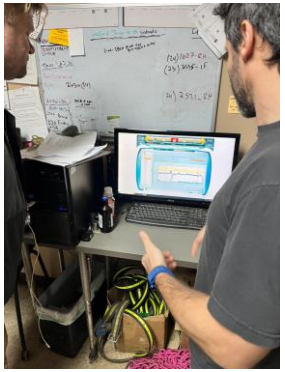
Rubambiza, G.; Dumba, B.; Anderson, Weatherspoon, H. ***EdgeRDV**: A Framework for Edge Workload Management at Scale*.

IEEE Edge '23

Edge Computing at Chick-fil-A. Chambers *et al*. 2022

A Cloud-Native Framework for Data Analytics

Approach: Augment the Software-Defined Farm (SDF) with agricultural domain expertise and cloud native abstractions to manage multiple farm clusters

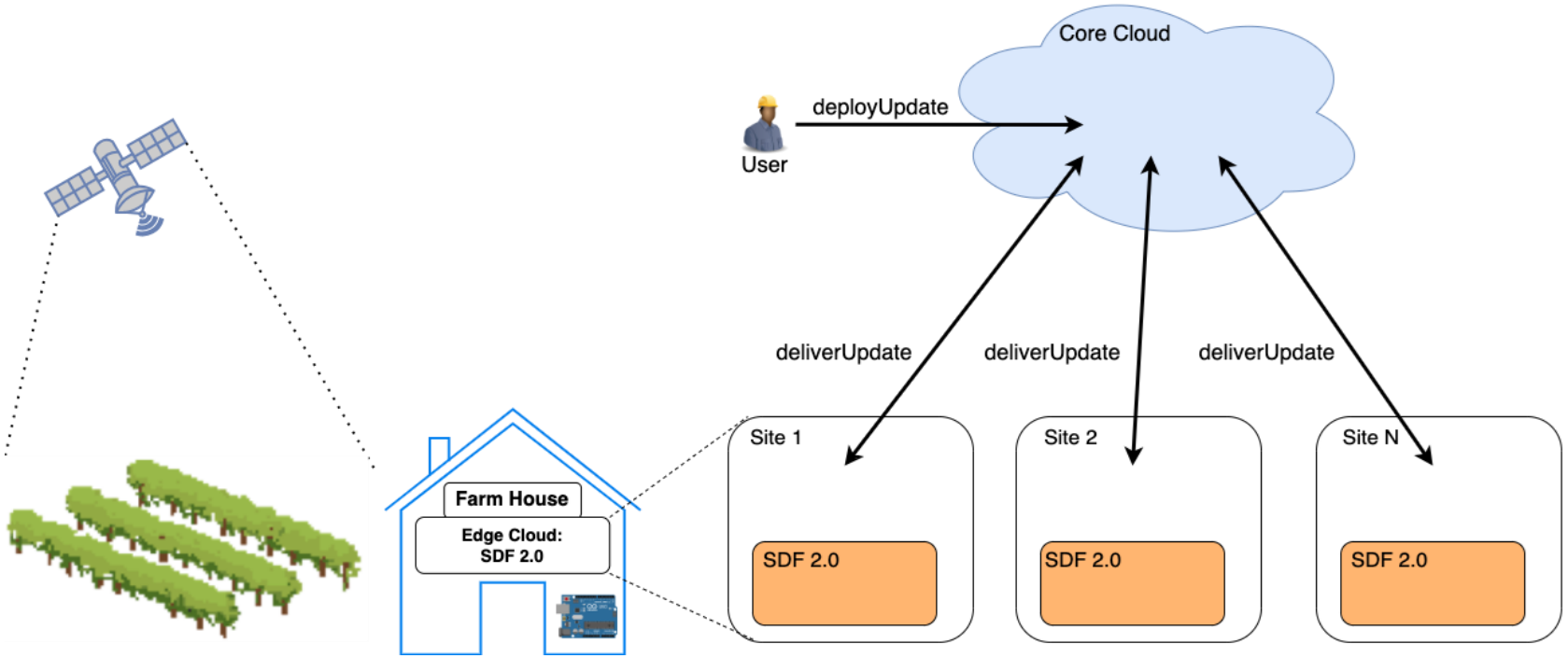


Cornell Dairy Research Center
Feb 2023

Dakota Shy Winery
May 2023

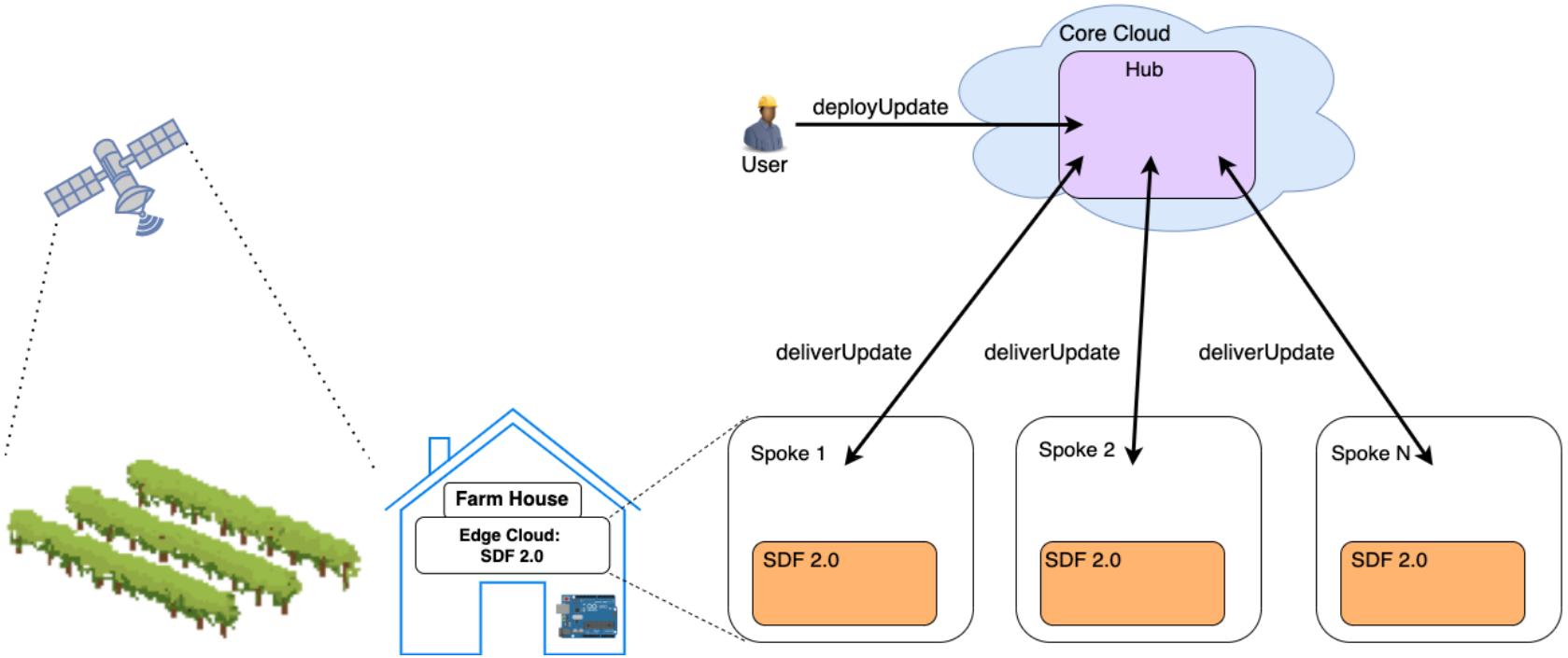
Jasmine Vineyards
June 2023

A Cloud-Native Framework for Data Analytics



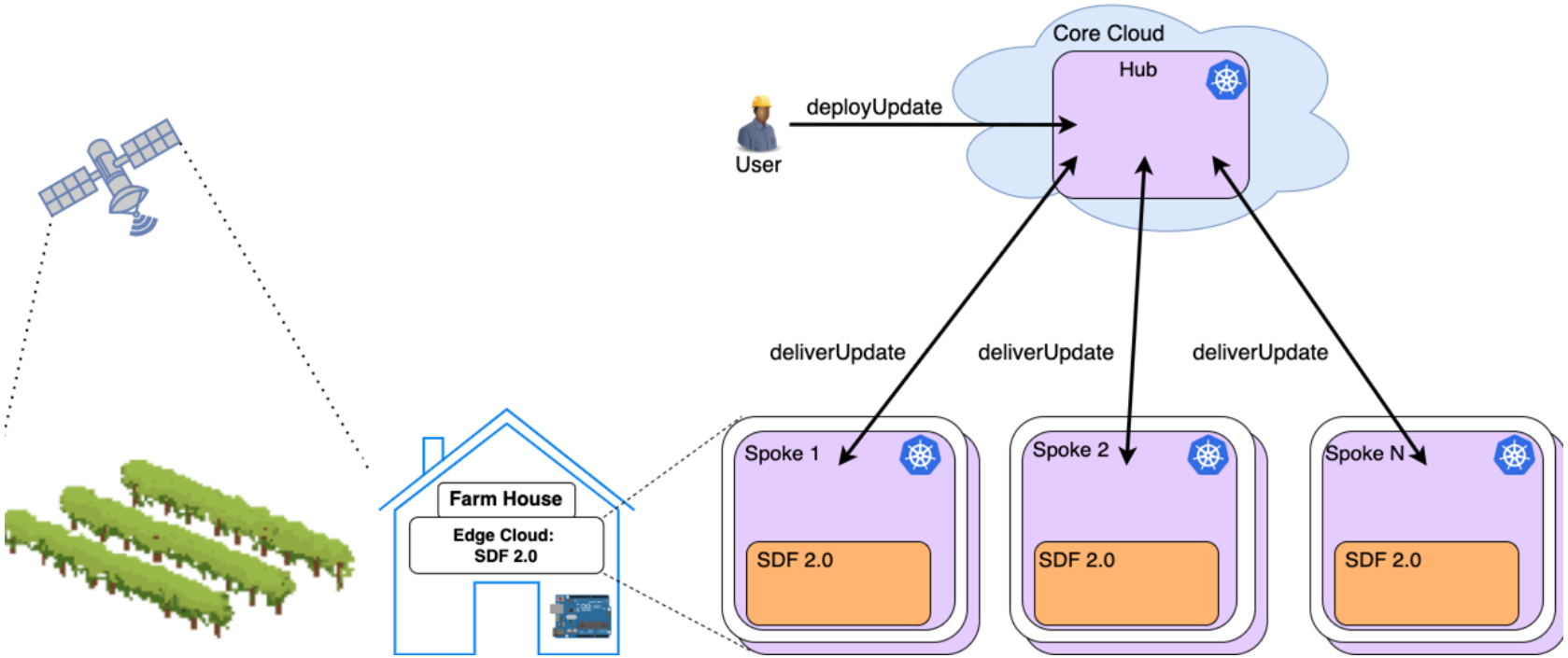
EdgeRDV: A Framework for Edge Workload Management at Scale. Rubambiza *et al.* 2023

A Cloud-Native Framework for Data Analytics



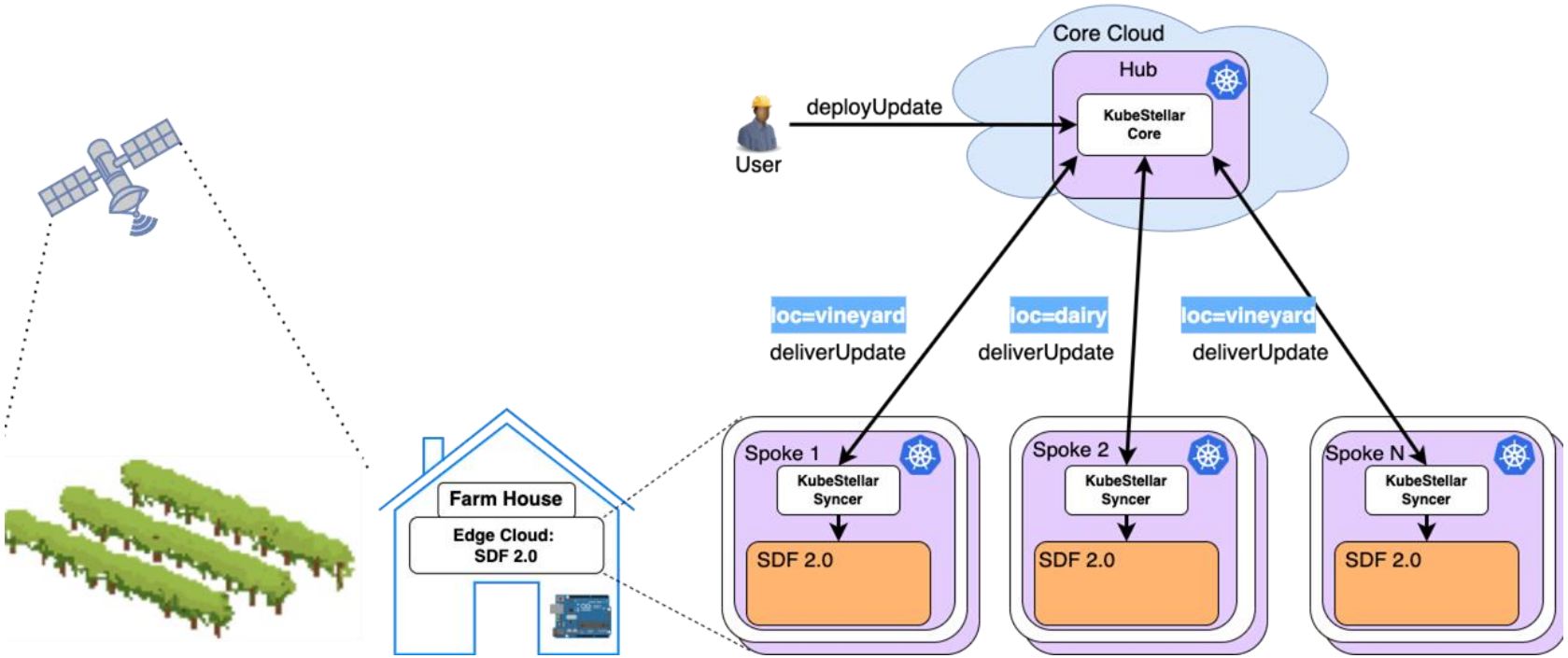
EdgeRDV: A Framework for Edge Workload Management at Scale. Rubambiza *et al.* 2023

A Cloud-Native Framework for Data Analytics



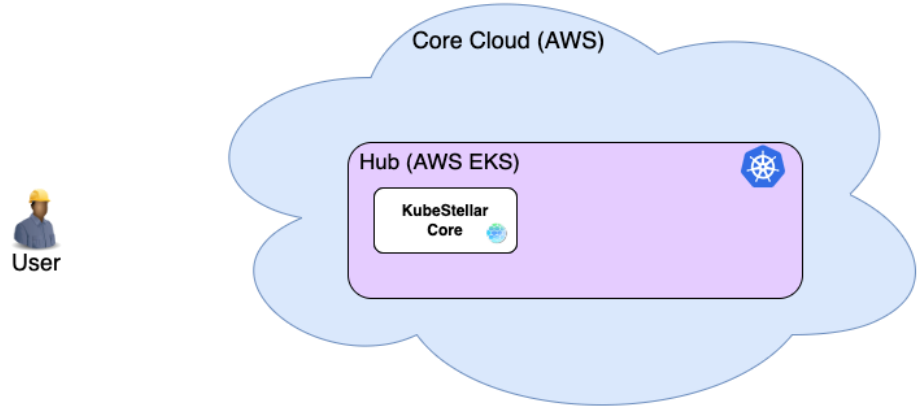
EdgeRDV: A Framework for Edge Workload Management at Scale. Rubambiza *et al.* 2023

A Cloud-Native Framework for Data Analytics

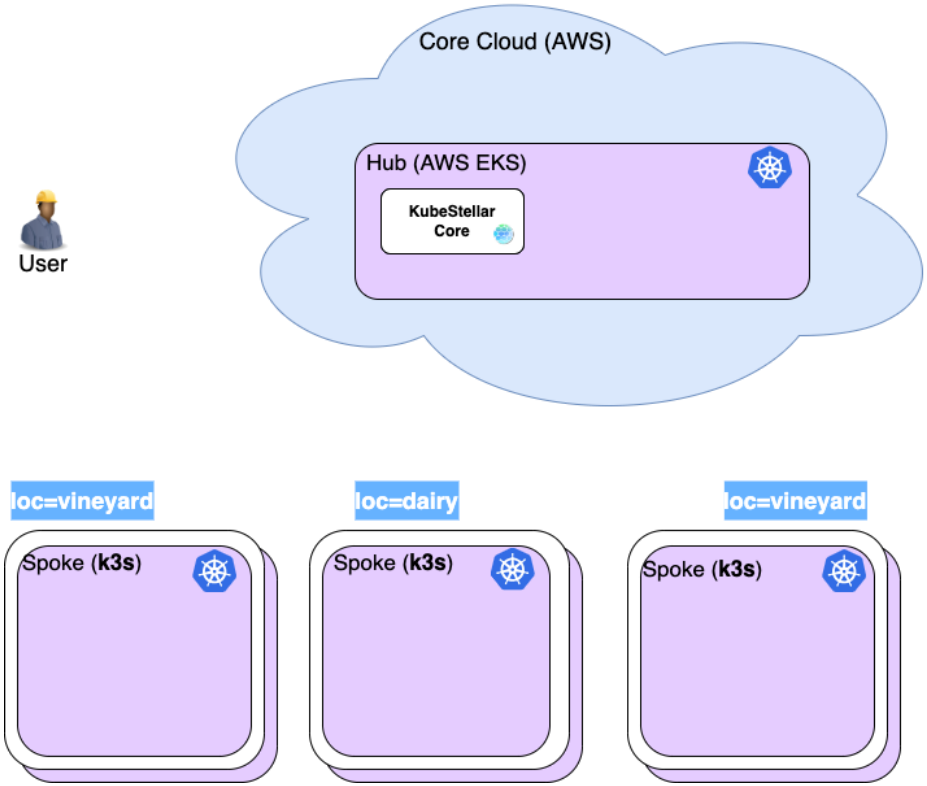


https://docs.kubestellar.io/release-0.14/readme/. The KubeStellar Authors. 2024

Cloud-Native Framework: Implementation

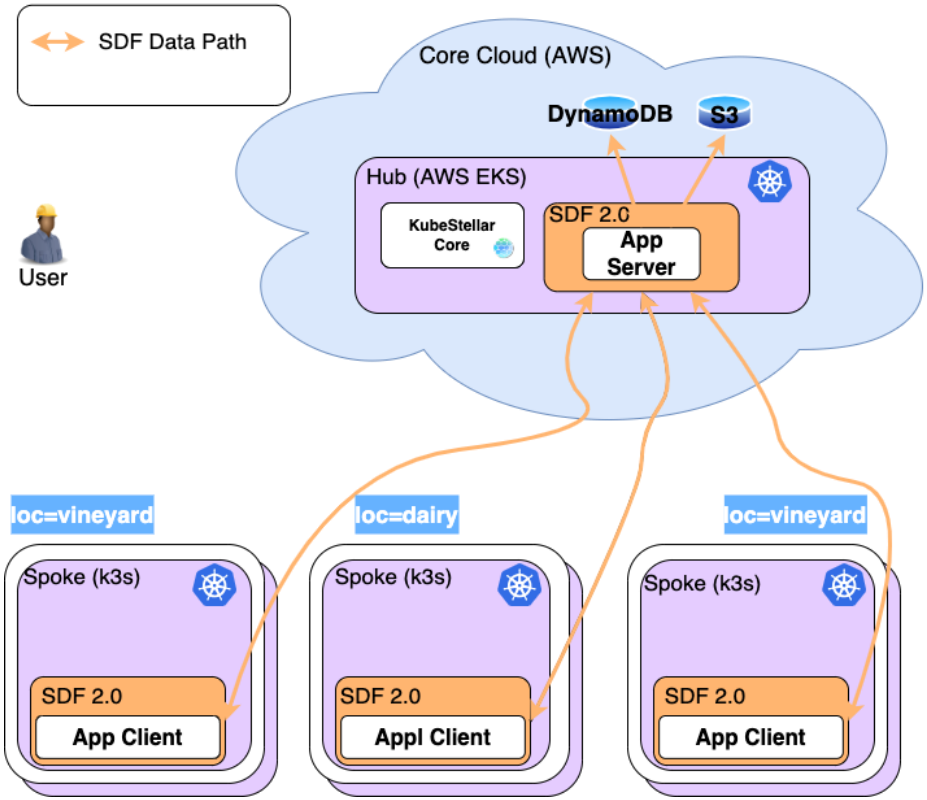


Cloud-Native Framework: Implementation

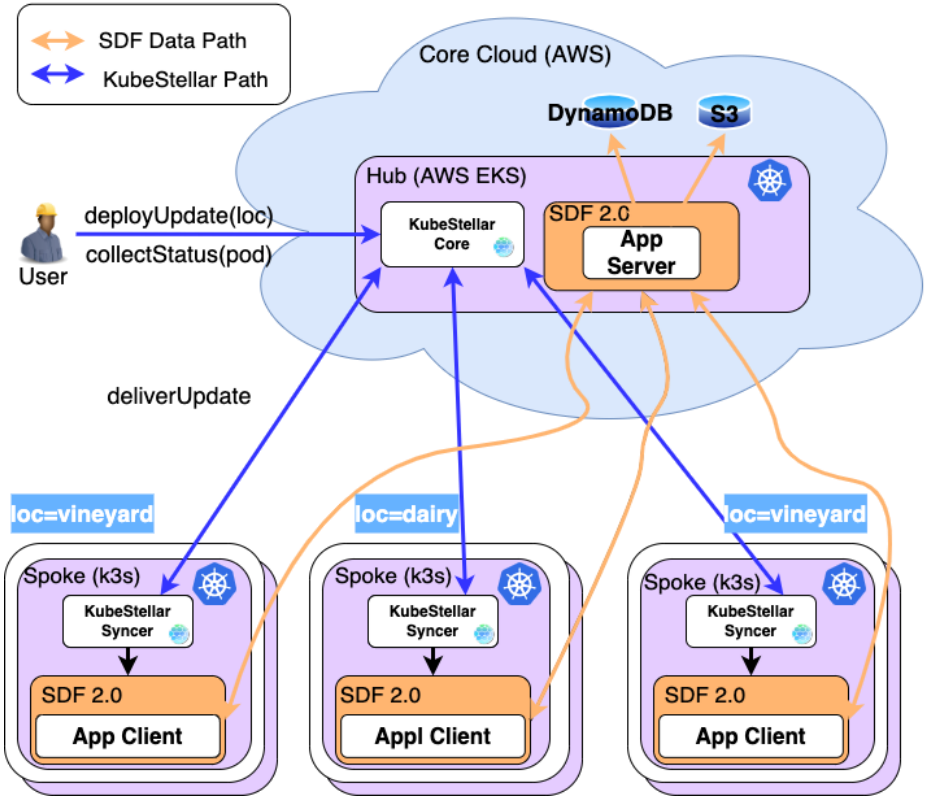


Lightweight Kubernetes Distributions: A Performance Comparison of MicroK8s, k3s, k0s, and Microshift. Koziolek and Eskadani. 2023

Cloud-Native Framework: Implementation

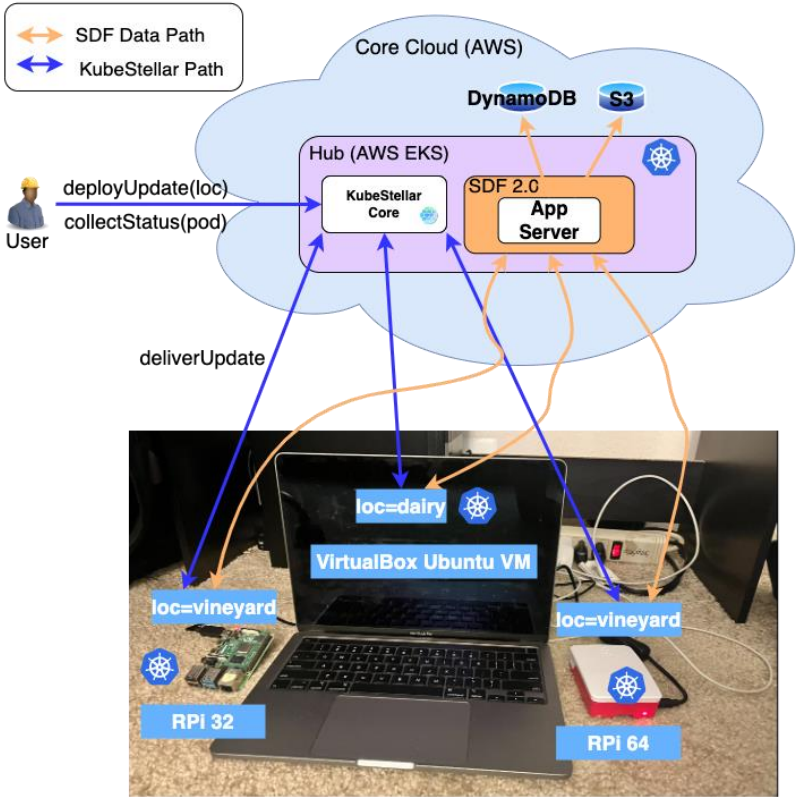


Cloud-Native Framework: Implementation



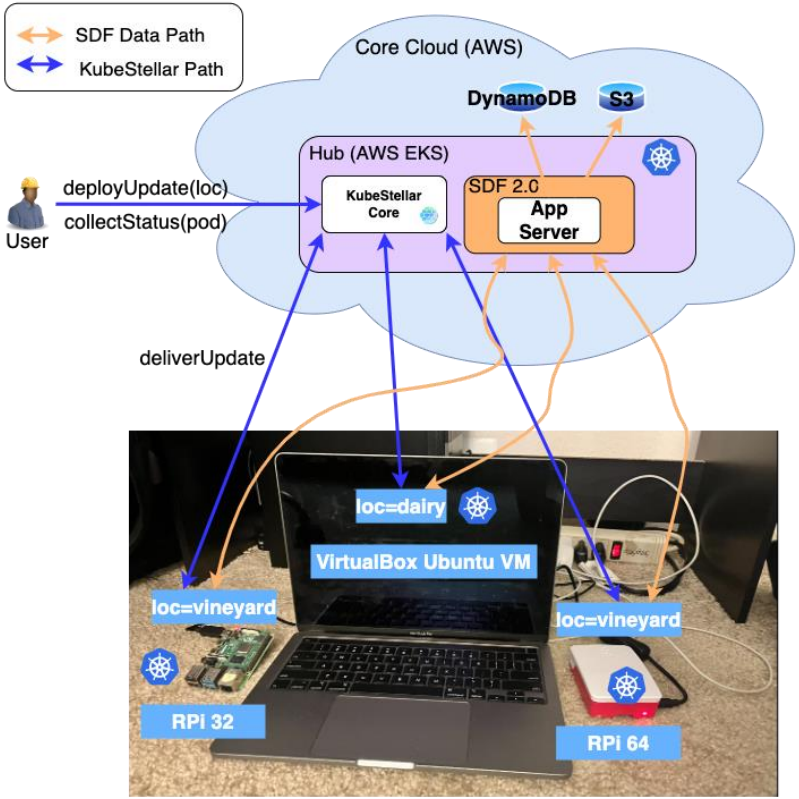
Cloud-Native Framework: Evaluation

- **Setup**
  - Hub: AWS East-2 (Virginia)
  - Vineyards: Raspberry Pis (32/64-bit)
  - Dairy: Ubuntu virtual machine
- **Metrics**
  - Workload update latency
  - End-to-end request latency
  - Qualitative cluster provisioning



Cloud-Native Framework: Evaluation

- **Setup**
  - Hub: AWS East-2 (Virginia)
  - Vineyards: Raspberry Pis (32/64-bit)
  - Dairy: Ubuntu virtual machine
- **Metrics**
  - **Workload update latency**
  - End-to-end request latency
  - Qualitative cluster provisioning

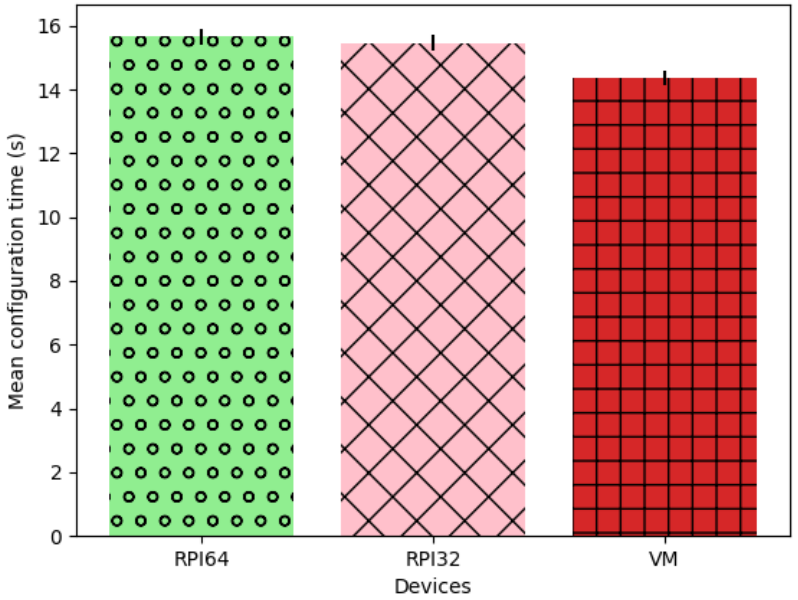


Cloud-Native Framework: Evaluation

Minimizing update latency

- **Experiment**:
  - To estimate the latency to create a pod for 50 consecutive updates

- **Observation(s)**:
  - Average latency is **15s** regardless of client hardware/operating system

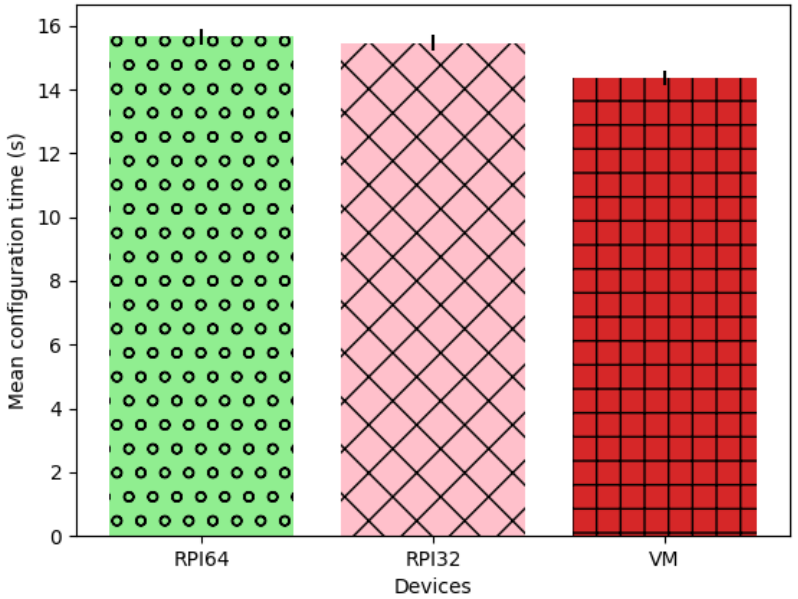


Cloud-Native Framework: Evaluation

Minimizing update latency

- **Experiment**:
  - To estimate the latency to create a pod for 50 consecutive updates

- **Observation(s)**:
  - Average latency is **15s** regardless of client hardware/operating system
  - SDF 2.0 **reduces update** deployment from **24h to 15s**



Lightweight Kubernetes Distributions: A Performance Comparison of MicroK8s, k3s, k0s, and Microshift. Koziolek and Eskadani. 2023

Cloud-Native Framework: Lessons Learned

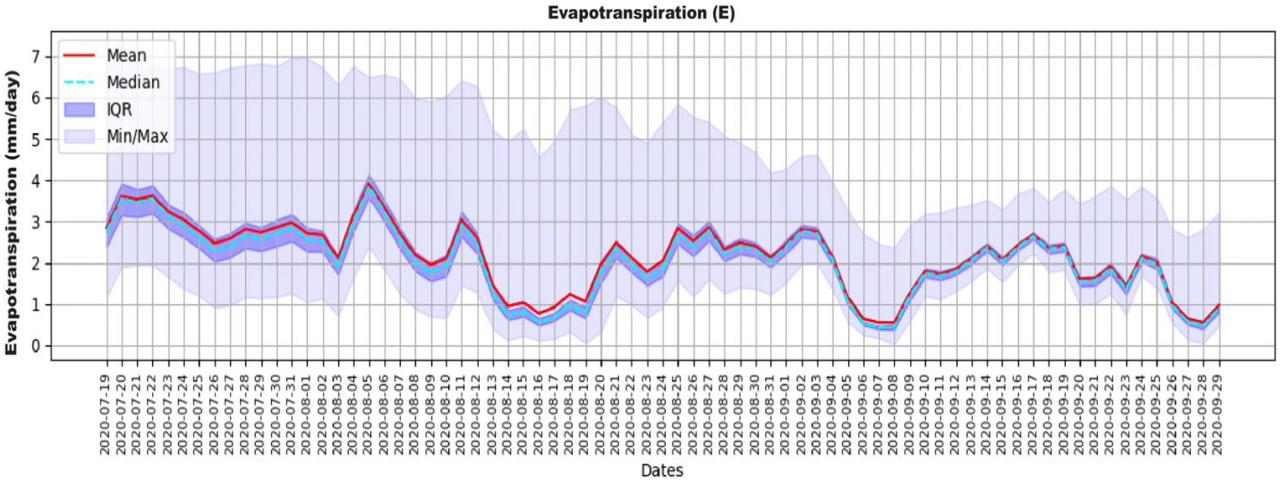


Image credit: Fernando Romero Galvan

- Farm resilience through disconnected operations
- Better understanding of farm privacy challenges
- Empower stakeholders with timely insights

Conclusion & Next Steps

- Presented a platform based on the hub-and-spoke model to enable a cloud-native data analytics in digital agriculture
- Implemented using KubeStellar and SDF platforms
- Next steps:
  - Field deployment in Northern California
  - Extend platform to enable privacy preserving AI models training and sharing across farms