Embracing Network Softwarization for Enhanced Efficiency and Performance
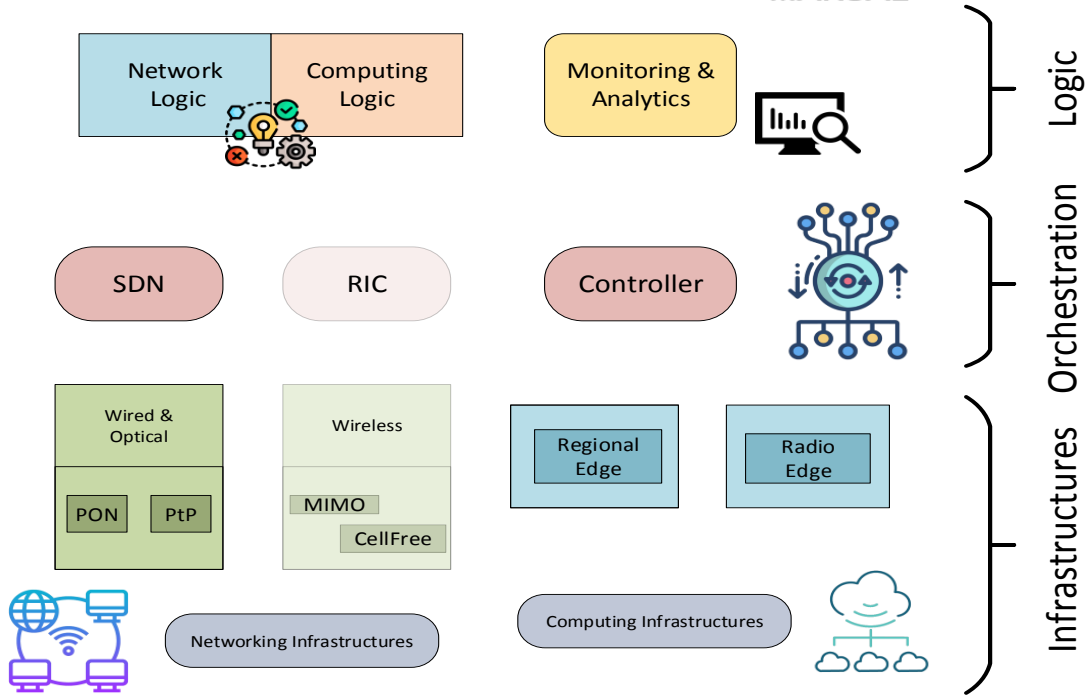
Fotis Kouzinos, NTUA

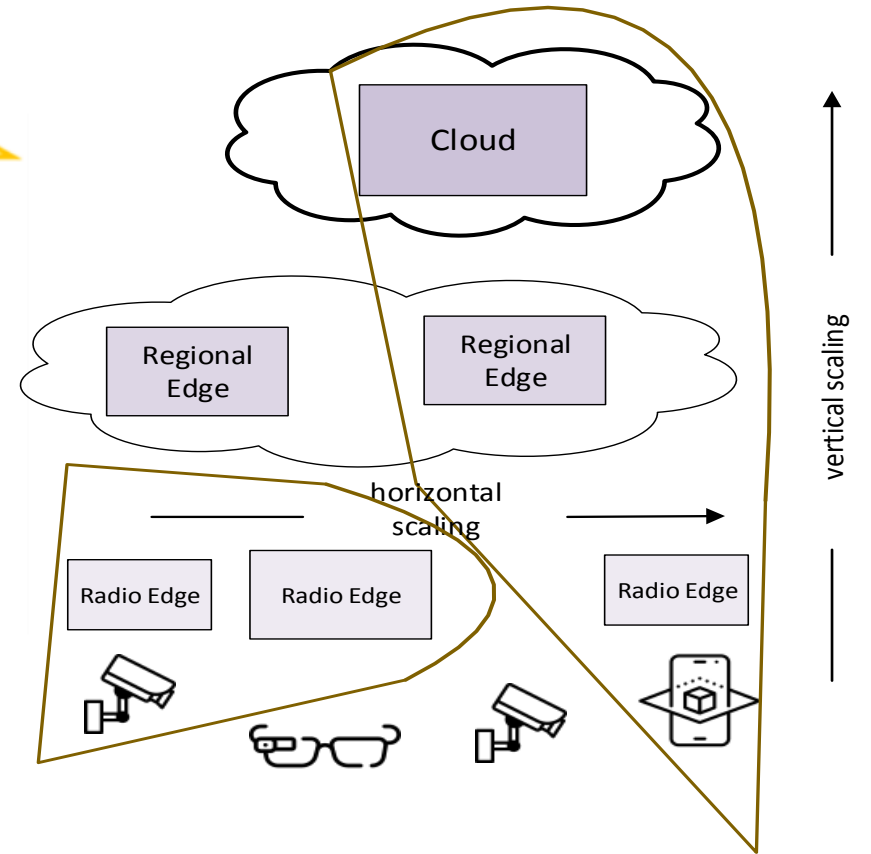# MARSAL Project

1ˢᵗ Major Challenge:
Realize a Virtual Elastic Infrastructure

1ˢᵗ Major Innovation: Design and implement orchestration, decision logic and analytic mechanisms for computing (edge & cloud) and networking (Fixed Access – FAN and Radio Access Network) resources

2ⁿᵈ Major Challenge: Provide an Elastic MEC platform for Cloud-Native applications

2ⁿᵈ Major Innovation: Develop a disaggregated MEC platform that manages application workload horizontally (across the edge) and vertically (from the edge to the cloud), while accessed by any UE

SERRANO Project

- Intent-based abstraction layer transforming distributed edge, cloud and HPC resources into a single borderless infrastructure.

- Orchestrates resources that contain the HW and SW innovations developed in WP3, WP4.
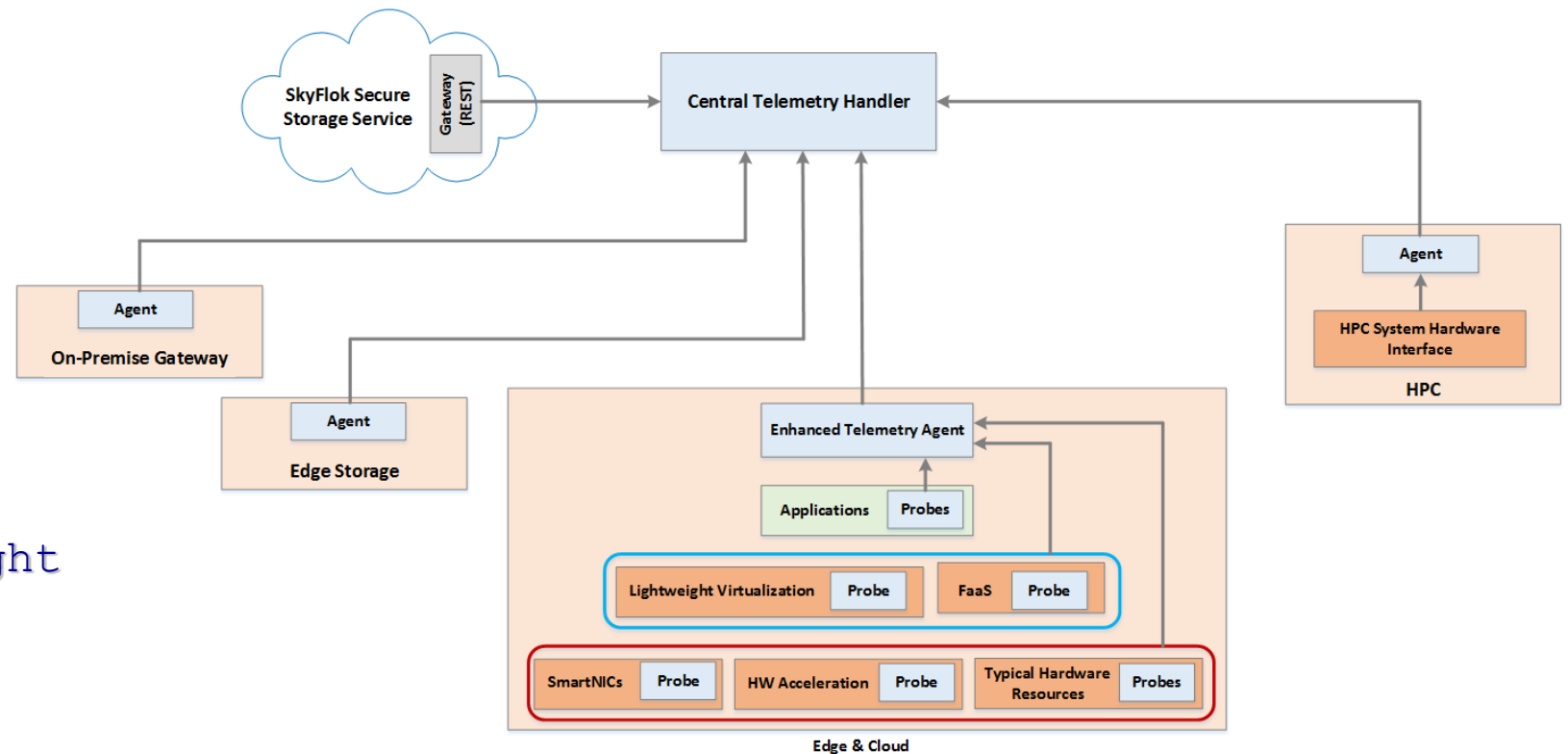
- WP pillars:
  - Cognitive end-to-end resource orchestration.
  - Hierarchical closed-looped control.
  - Transparent application deployment.
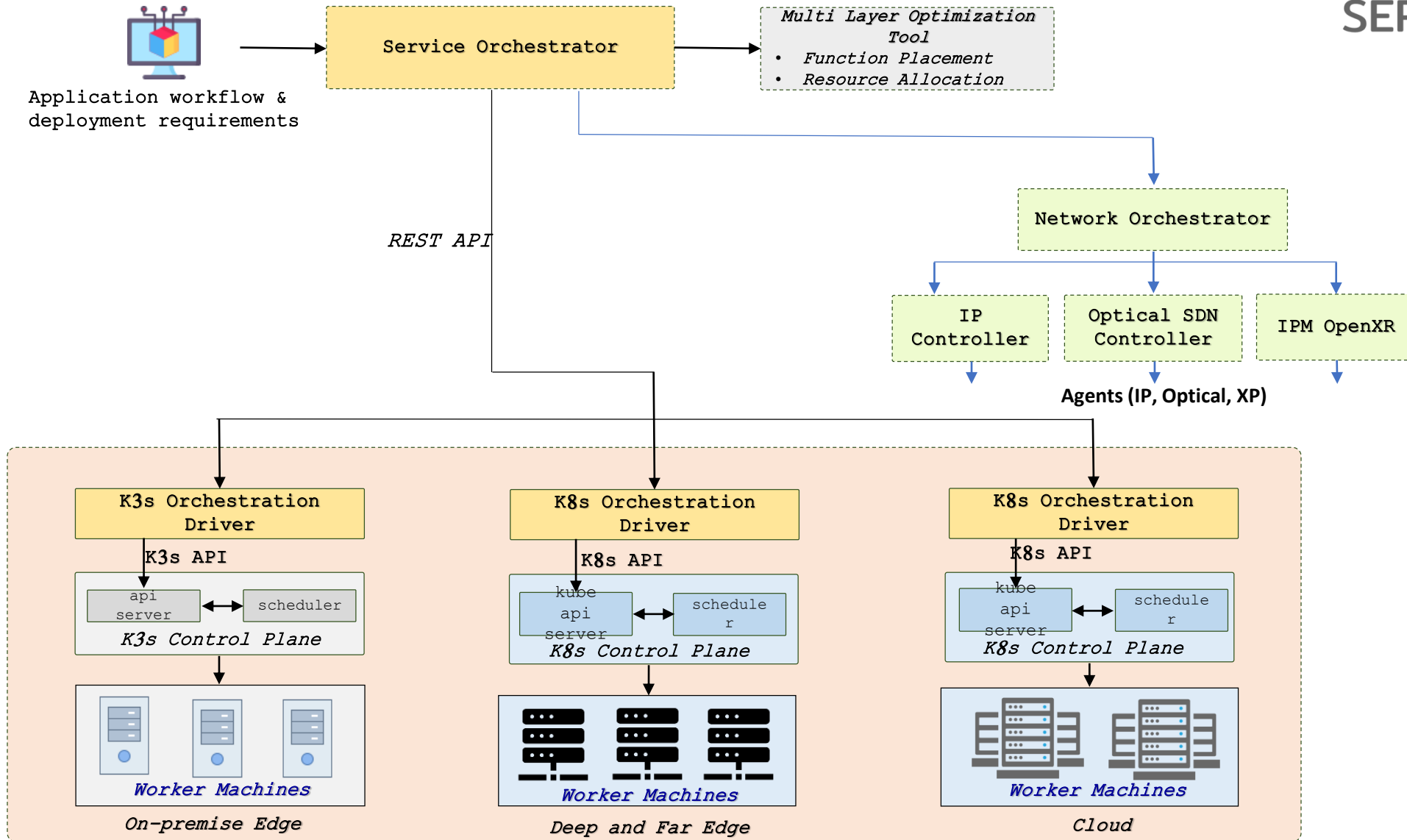 - Automatic and continuous adaptation.
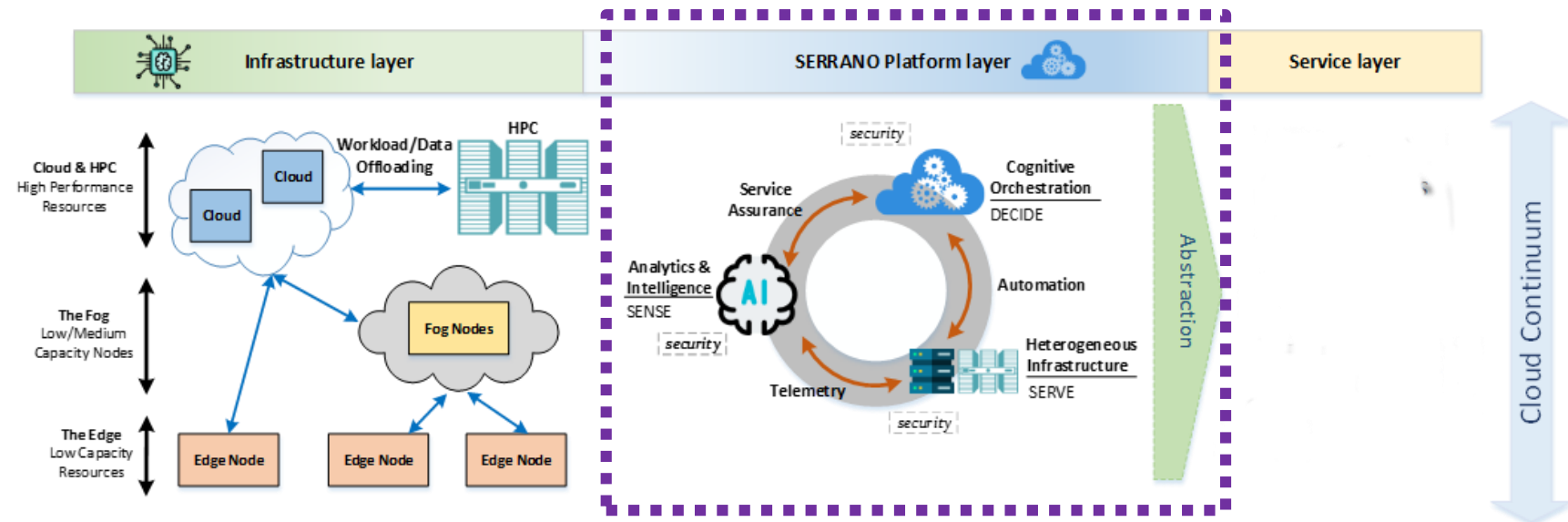
Cloud and Network Telemetry Handler

- Inventory, monitoring and telemetry information
- Storage services:
  - Cloud storage
  - Edge storage
  - On-premise Gateway
  - HPC platforms
- Edge & Cloud:
  - Typical resources
  - GPUs & FPGAs
- SERRANO enhancements
  - hardware kernels
  - trusted and lightweight virtualization
  - HW acceleration abstractions
  - serverless
- Applications

SkyFlok Secure Storage Service — Gateway (REST) → Central Telemetry Handler

Agent — On-Premise Gateway

Agent — Edge Storage

Agent — HPC System Hardware Interface — HPC

Enhanced Telemetry Agent

Applications | Probes

Lightweight Virtualization | Probe   FaaS | Probe

SmartNICs | Probe   HW Acceleration | Probe   Typical Hardware Resources | Probes

Edge & Cloud

Cloud-native App orchestration over distributed K8s/K3s platforms

Service Orchestrator - Architecture

Resource Orchestrator:
- Orchestration API: includes Access Interface and Dispatcher
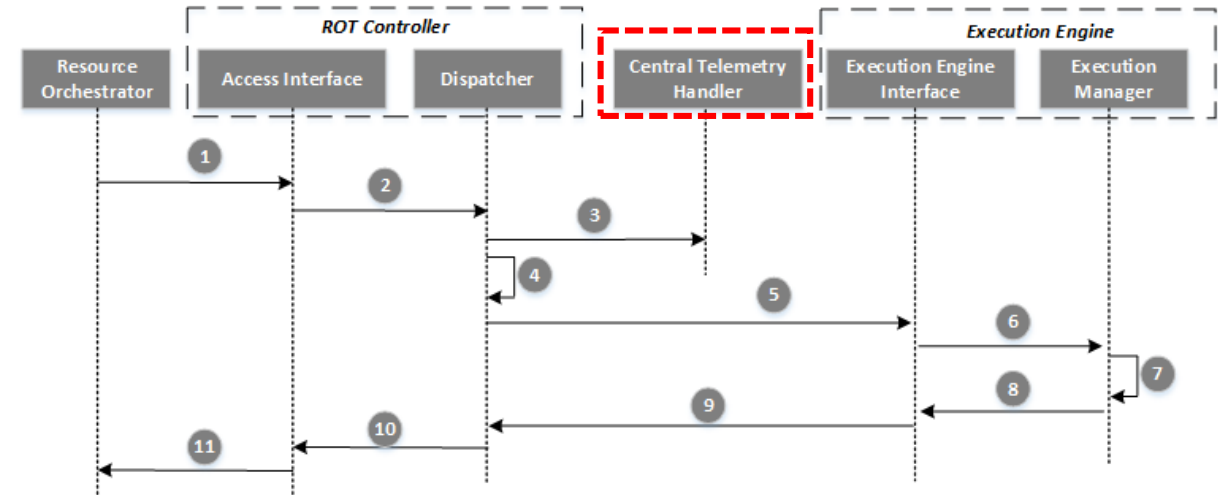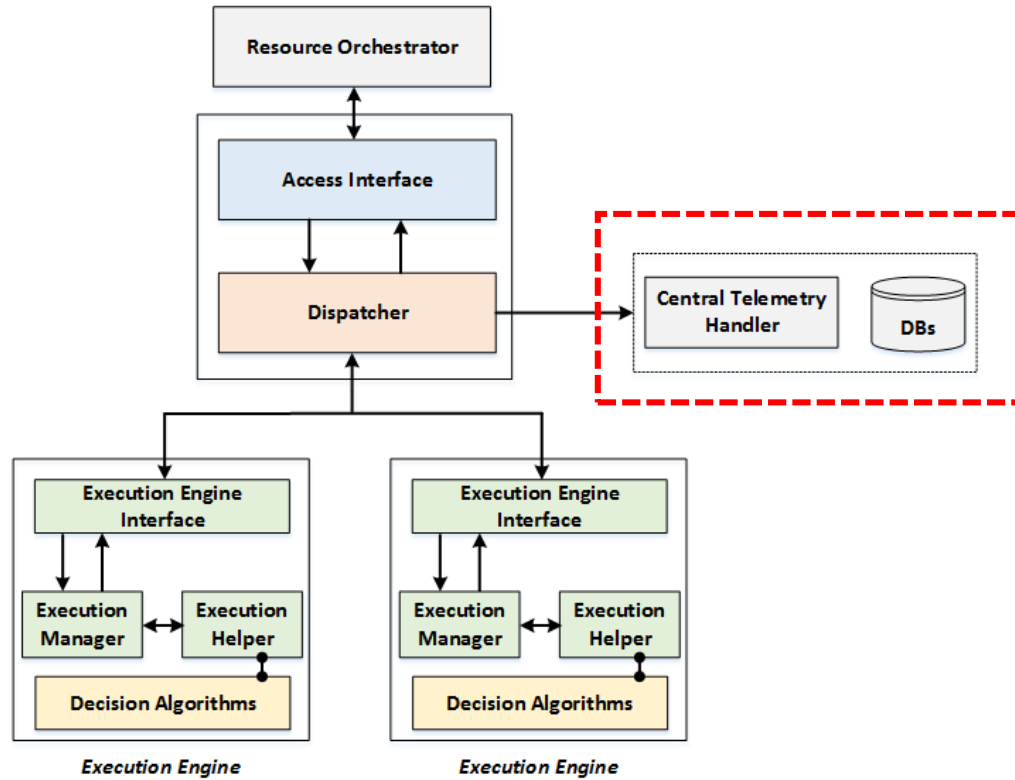- Orchestration Manager: includes Orchestration Manager

Orchestration Drivers (k8s, k3s) different service.

Datastore, based on etcd, stores orchestration-related data for entire platform

Available REST API for managing applications and resources

Decision Engine

PMDS Interfaces
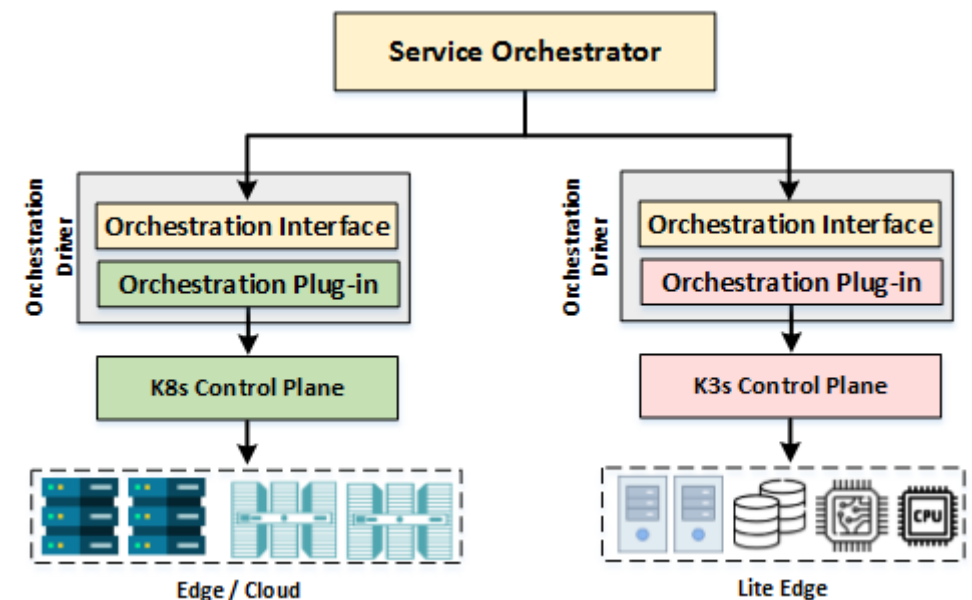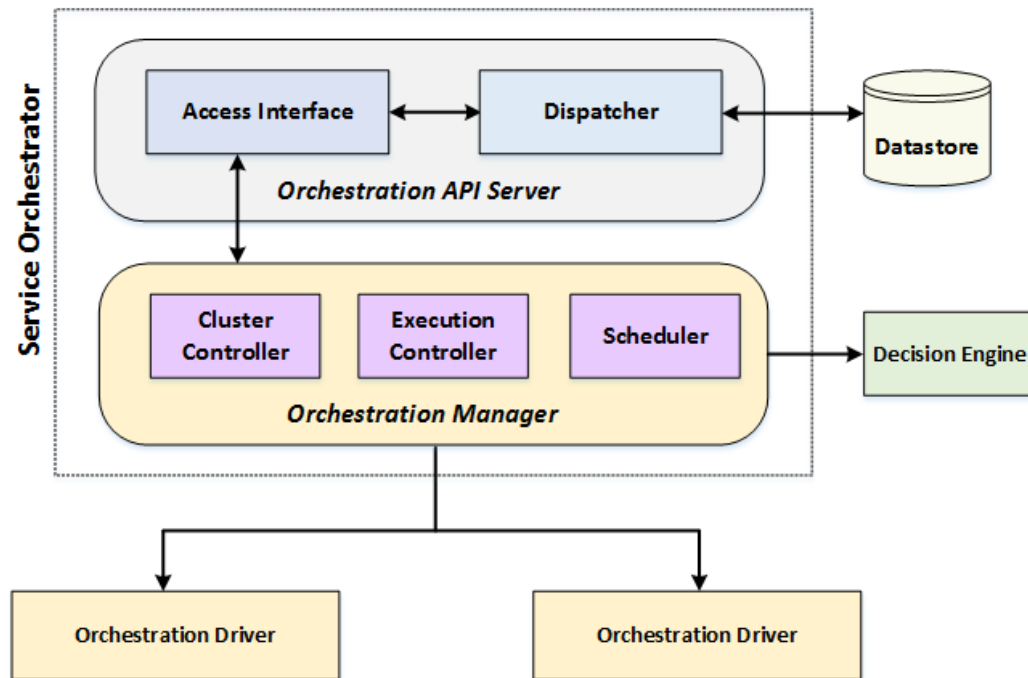
GET `/api/v1/pmds/nodes/{cluster_uuid}` Retrieve historical telemetry data for the worker nodes within a K8s cluster. ⌄

GET `/api/v1/pmds/pvs/{cluster_uuid}` Provide historical telemetry data for the available Persitient Volumes (PVs) within a K8s cluster. ⌄

GET `/api/v1/pmds/deployments/{cluster_uuid}` Provide historical telemetry data for the available Deployments within a K8s cluster. ⌄

GET `/api/v1/pmds/pods/{cluster_uuid}` Provide historical telemetry data for the available Pods within a K8s cluster. ⌄

GET `/api/v1/pmds/edge_storage_devices/{cluster_uuid}` Provide historical telemetry data for the available SERRANO Edge Storage devices within a K8s cluster. ⌄

GET `/api/v1/pmds/serrano_deployments/{deployment_uuid}` Provide historical telemetry data for a specific application deployed through the SERRANO orchestration mechanisms. ⌄

```
pmds_service_query_serrano_deployments("649decae-63ec-40cb-9c5f-eb16f5b93590", format="compact", start="-15m")
```

```
---> position-service-classifier-training-7466869887-b4b5f
{'cpu_usage': '26367782n', 'group_id': 's3', 'memory_usage': '245656Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:04:14.738953+00:00'}
{'cpu_usage': '2054355893n', 'group_id': 's3', 'memory_usage': '1899136Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:08:14.702984+00:00'}
{'cpu_usage': '999527171n', 'group_id': 's3', 'memory_usage': '247748Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:12:14.755065+00:00'}
{'cpu_usage': '151913n', 'group_id': 's3', 'memory_usage': '155368Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:16:14.884757+00:00'}


---> position-service-data-manager-df985f58d-fwghk
{'cpu_usage': '105824037n', 'group_id': 's1', 'memory_usage': '84664Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:04:14.744515+00:00'}
{'cpu_usage': '79932931n', 'group_id': 's1', 'memory_usage': '84844Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:08:14.708450+00:00'}
{'cpu_usage': '90799170n', 'group_id': 's1', 'memory_usage': '84440Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:12:14.760960+00:00'}
{'cpu_usage': '50509519n', 'group_id': 's1', 'memory_usage': '84428Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:16:14.889931+00:00'}


---> position-service-model-inference-85dfb84975-c8jv2
{'cpu_usage': '424217130n', 'group_id': 's2', 'memory_usage': '216684Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:04:14.749158+00:00'}
{'cpu_usage': '585385005n', 'group_id': 's2', 'memory_usage': '307244Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:08:14.712578+00:00'}
{'cpu_usage': '683851129n', 'group_id': 's2', 'memory_usage': '307432Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:12:14.765617+00:00'}
{'cpu_usage': '592624658n', 'group_id': 's2', 'memory_usage': '174504Ki', 'phase': 'Running', 'restarts': 0, 'time': '2023-06-09T11:16:14.894321+00:00'}
```

THANK YOU!

MERRY CHRISTMAS!