KubeCon | CloudNativeCon

North America 2022

BUILDING FOR THE ROAD AHEAD

DETROIT 2022

Smart Green Computing Cloud-Native Operations

KubeCon | CloudNativeCon
North America 2022
BUILDING FOR THE ROAD AHEAD
DETROIT 2022

October 24-28, 2022

William Caban
Telco Chief Architect
Red Hat

Federico Rossi
Senior Solution Architect
Red Hat

Smart Green Computing
Cloud-Native Operations

The European Union and the United States have set up a target of at least 50% - 55% net reduction in greenhouse gas emissions by 2030. But, with the sprawling of the cloud-native workloads and the increased demand for resources: are we doing enough?

Many community efforts and open source projects enable the observability of the power consumption from software resources to hardware resources. How can we combine the visibility provided by these tools to achieve the organization's sustainability goals?

In this talk, we combine CNCF projects and other open source communities tools to create and continuously improve Machine Learning models for cluster operations. These ML models consider a holistic view of a system: from application runtimes, node metrics, cluster metrics, and network metrics to the tracing of the interactions among the distributed components. These ML models are used for the "smart operations" of the distributed systems aligning to the organization's carbon and power optimization goals.

Smart Green Computing Cloud-Native Operations

William Caban
Global Telco Chief Architect
linkedin.com/in/williamcaban
@williamcaban

Federico Rossi
Senior Solution Architect
linkedin.com/in/ferossi-rh

Smart Green Computing
Cloud-Native Operations

Goal-driven Workload Management (1)

KubeCon CloudNativeCon North America 2022

CO2/Energy Aware Workload Placement & Scheduling

Goal-driven Scheduling
E.g. scheduling workloads to achieve organization's **goal of reduction of CO2 emissions, or reduction of power, or achieve higher energy efficiency,** or to maintain certain latency, etc.

Governance
Workload placement following local or cross-border regulations

Continuous Re-evaluation
E.g. Near-Real Time reassessment of placement decision and proactively schedule/deschedule workload.

Traditional Scheduling

Workload A	Workload B	Workload C
CO2: 1 Units Replicas: 3	**CO2: 3 units** Replicas: 3	**CO2: 5 Units** Replicas: 3

~~Goal: < 12 CO2 units per node/cluster~~

Source: **Carbon**	Source: **Petroleum**	Source: **Natural Gas**
CO2: **2.23 lb p/kWh**	CO2: **2.13 lb p/kWh**	CO2: **0.91 lb p/kWh**
T_CO2: 9 x 2.23 = **20.07**	T_CO2: 9 x 2.13 = **19.17**	T_CO2: 9 x 0.91 = **8.19**

Goal-driven Workload Management (2)

CO2/Energy Aware Workload Placement & Scheduling

Goal-driven Scheduling
E.g. scheduling workloads to achieve organization's **goal of reduction of CO2 emissions, or reduction of power, or achieve higher energy efficiency,** or to maintain certain latency, etc.

Governance
Workload placement following local or cross-border regulations

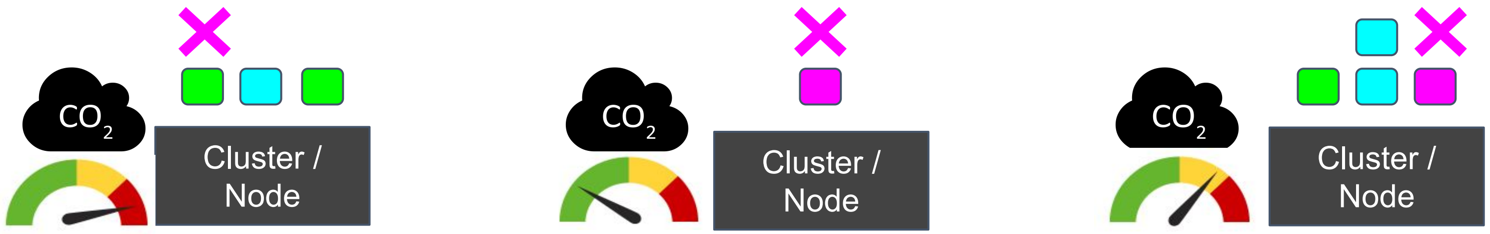
Continuous Re-evaluation
E.g. Near-Real Time reassessment of placement decision and proactively schedule/deschedule workload.

Goal-driven (e.g. CO2) scheduling/rescheduling

Workload A	Workload B	Workload C
CO2: 1 Units Replicas: 3	**CO2: 3 units** Replicas: 3	**CO2: 5 Units** Replicas: 3

?

Goal: < 12 CO2 units per node/cluster



Source: **Carbon**	Source: **Petroleum**	Source: **Natural Gas**
CO2: **2.23 lb p/kWh**	CO2: **2.13 lb p/kWh**	CO2: **0.91 lb p/kWh**
T_CO2: **5** x 2.23 = **11.15**	T_CO2: **5** x 2.13 = **10.65**	T_CO2: **12** x 0.91 = **10.92**
		(Max) 13 x 0.91 = 11.83

Domain-aware Workload Management

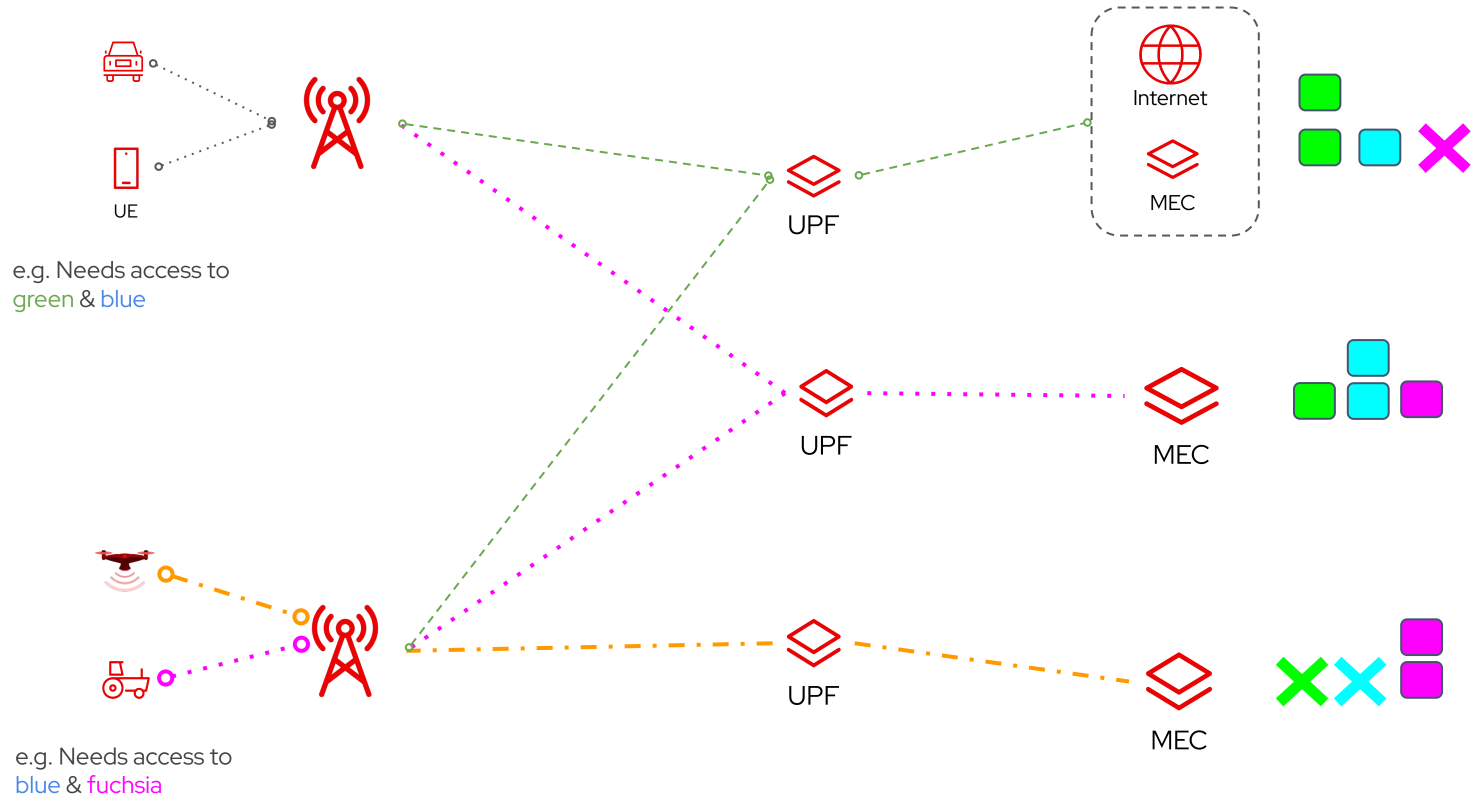
Domain Aware Workload Placement & Scheduling

Governance
Workload placement following local or cross-border regulations

Domain-aware Scheduling
E.g. scheduling specialized workloads (e.g. 5G UPF, MEC apps) to achieve organization's goal while maintaining the strict constraints imposed by the app or service requirements (e.g. latency, throughput, number of sessions)

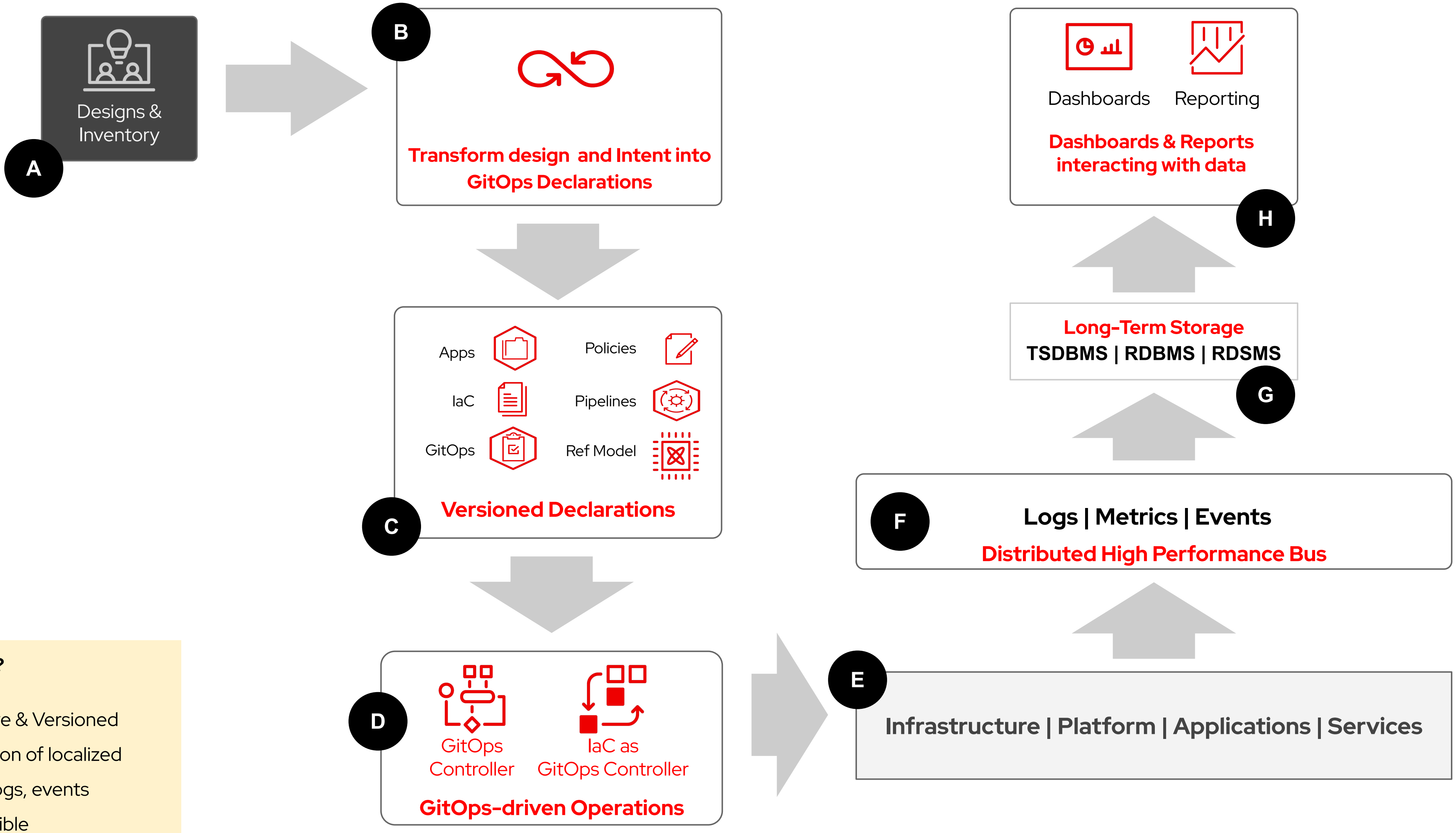
Continuous Re-evaluation
E.g. Near-Real Time reassessment of placement decision and proactively schedule/deschedule workload.

Workload Placement for 5G Networks



Composable Experimentation & Operations Stack Design

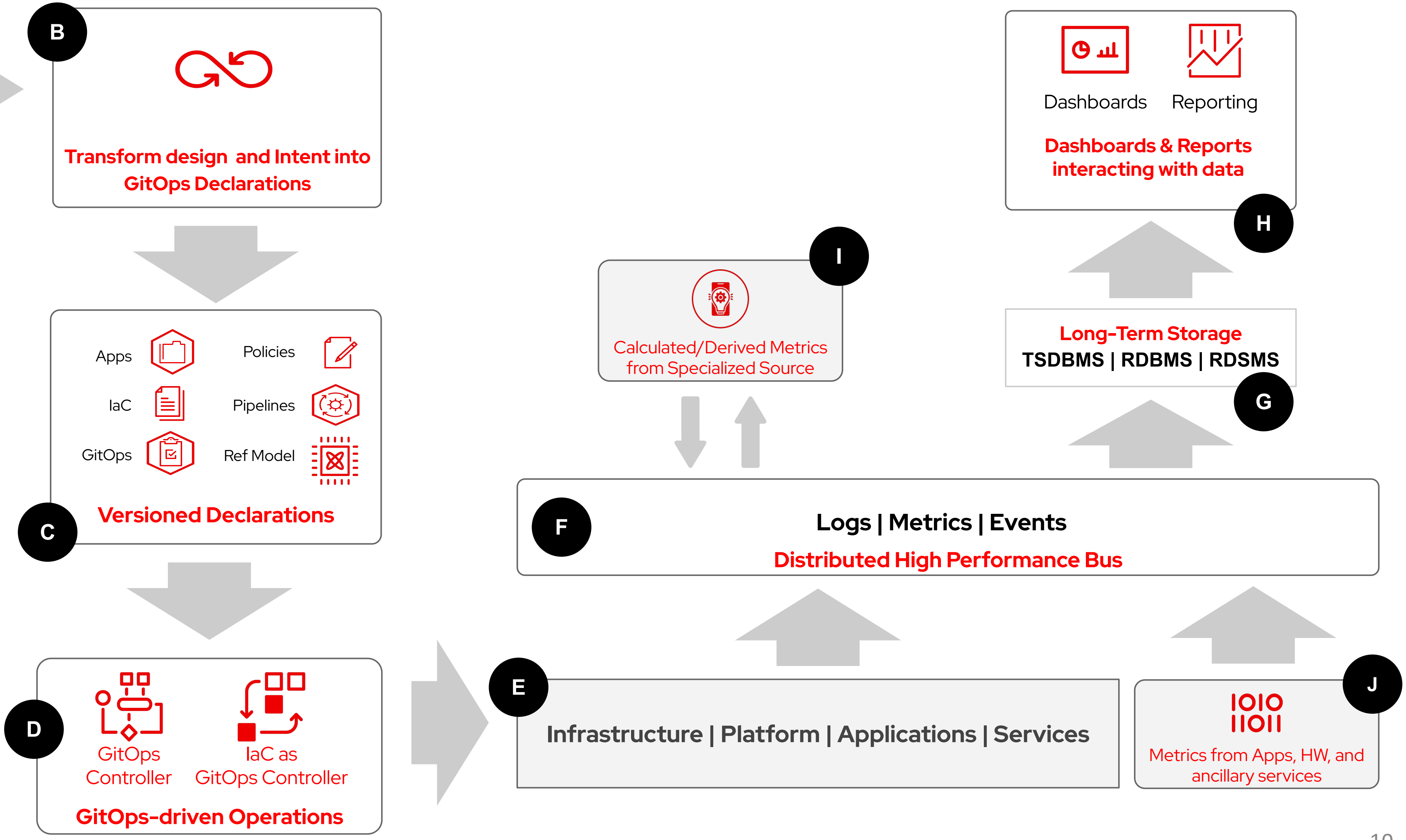
Step 1: Adopt GitOps Model



What do we gain?

- ❏ Declarative & Versioned
- ❏ Aggregation of localized metrics, logs, events
- ❏ Reproducible

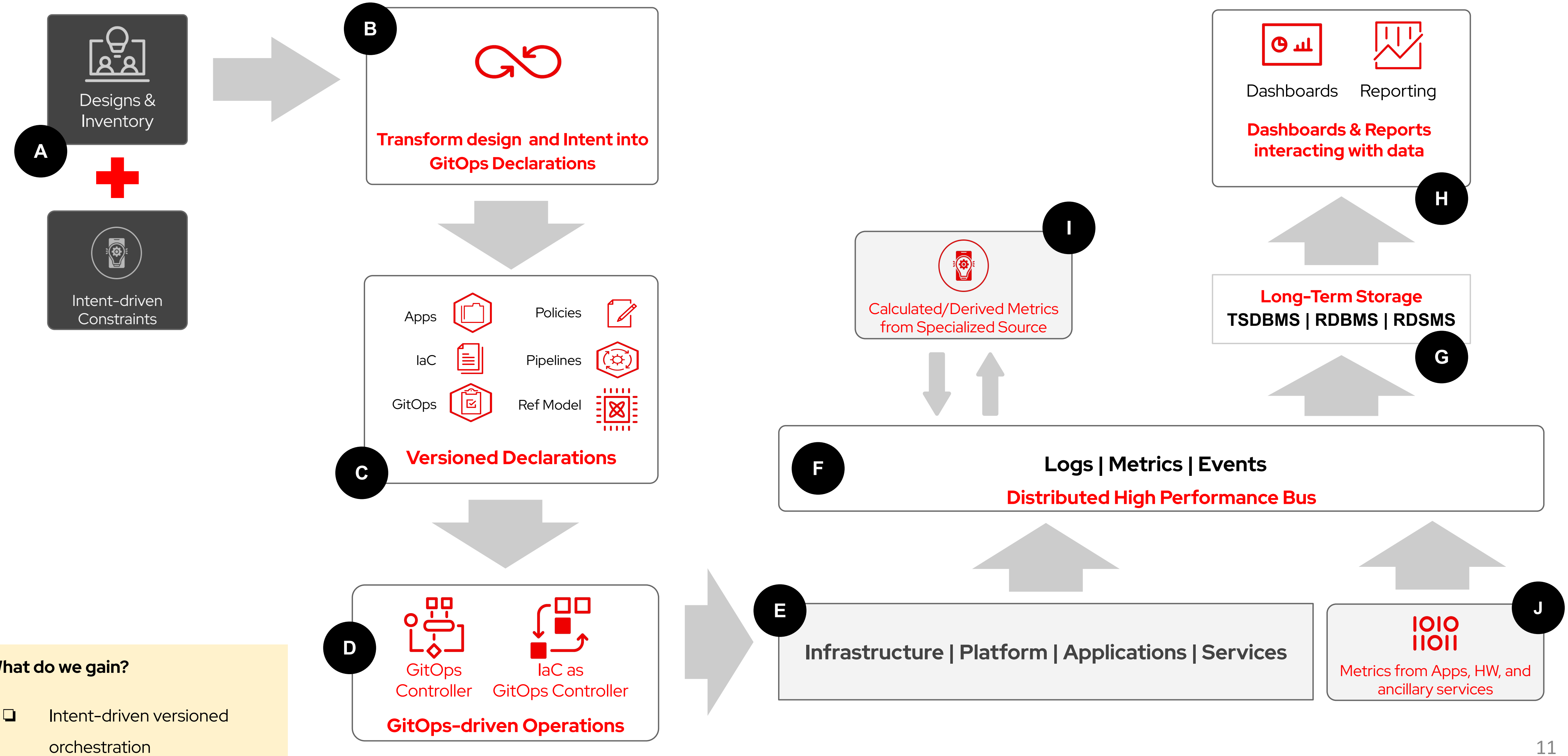
Step 2: Adopt Enhanced Visibility



What do we gain?

- ❏ Location-specific metrics, logs, events from ancillary hardware and services

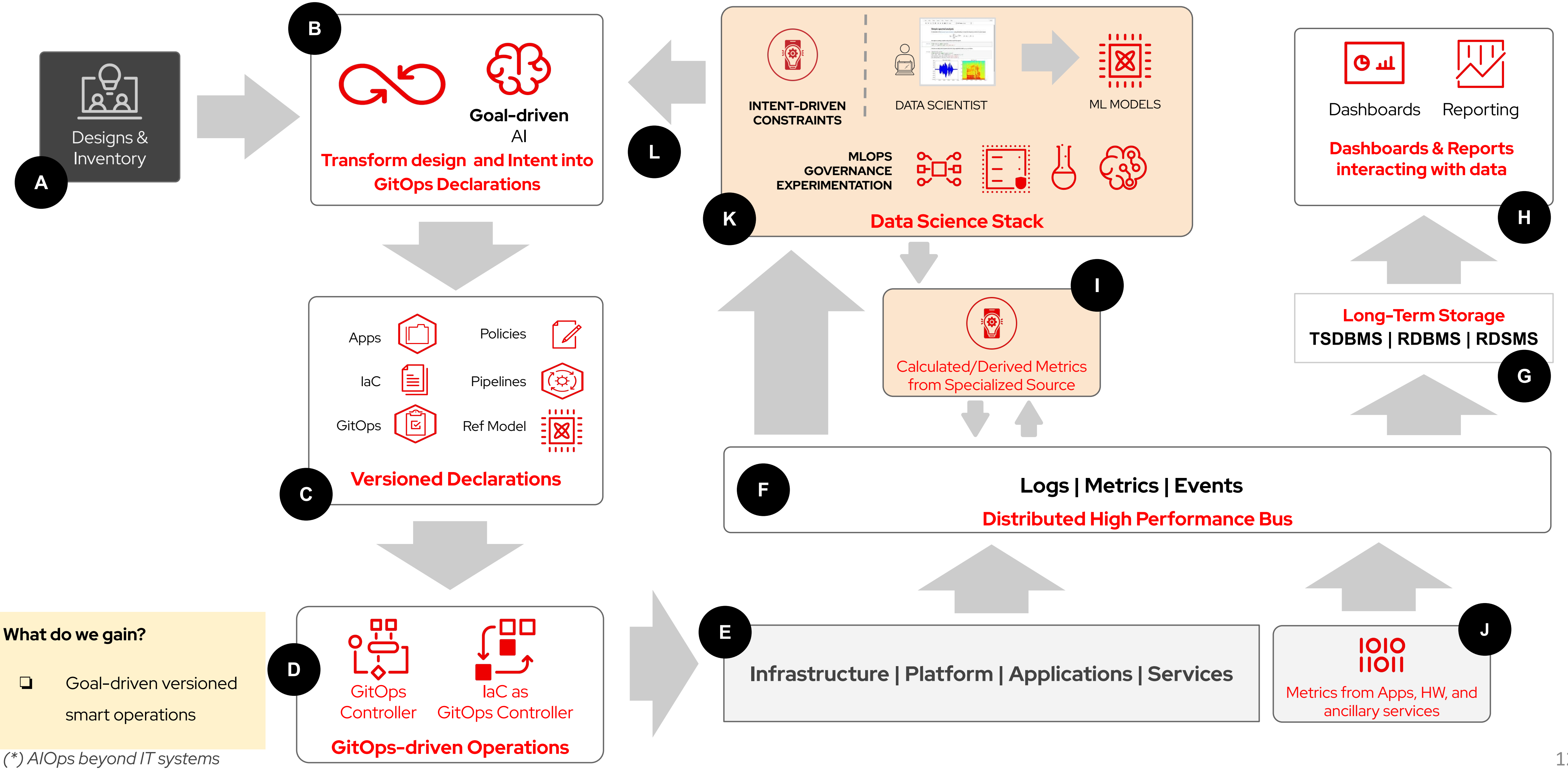
Step 3: Adopt Intent Driven Orchestration



What do we gain?

❏ Intent-driven versioned orchestration

Step 4: Adopt Goal-Driven AI (Intent + AIOps*)

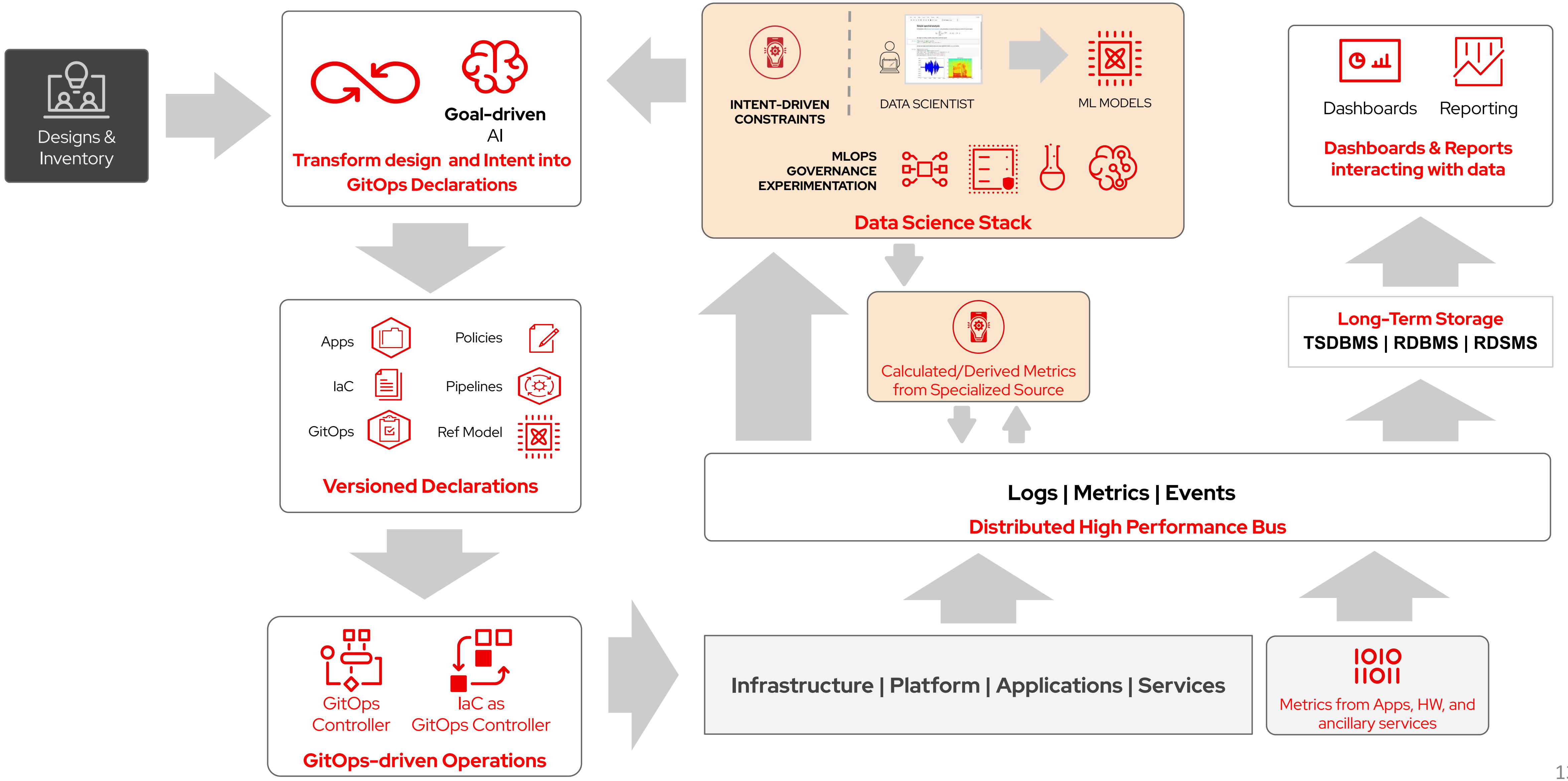


What do we gain?

- ❏ Goal-driven versioned smart operations

(*) AIOps beyond IT systems

The Complete Stack



Example of Use Cases for Stack

- Goals Aware Workload Management
  - Maximum CO2 emission per-node or cluster
  - **Energy aware** service-to-service **path selection**
  - **Predictive in-time** workload **provisioning**
- **Domain-aware autonomous orchestration**
  - Autonomous operations of services to **reduce or optimize power** consumption, or **CO2 emissions**
    - **Auto-tiering of services** (cloud, virtualization, bare-metal) based on multi-dimensional factors (financial models, energy sources, geo location, time, bandwidth, latency, QoE)
    - **Smart maintenance** operations (auto-scheduling and execution of maintenance windows)

This is today's POC!!!

CO2 & Energy Aware Scheduling POC

CO_2 Emission Aware Placement & Scheduling

KubeCon CloudNativeCon North America 2022

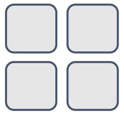
1 Watt = 1W = 1 J/s (1 Joule p/sec)
1 kW = 1000 J/s

CO_2 Emissions per pounds per kWh

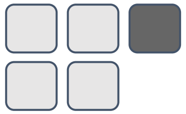
Node (w3)	Node (w2)	Node (w4)
Coal (2.23 pounds)	**Petroleum** (2.13 pounds)	**Natural Gas** (0.91 pounds)

Goal max 11 pounds per kWh (per Node)

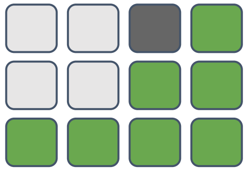
= 1 kWh Pod



4 Pods
8.92 pounds



5 Pods
10.65 pounds



12 Pods
10.92 pounds

(*) US Energy Information Administration How much carbon dioxide is produced per kilowatt hour of U.S. electricity generation? (https://www.eia.gov/tools/faqs/faq.php?id=74&t=11)

PoC High-Level Architecture

KubeCon | CloudNativeCon North America 2022

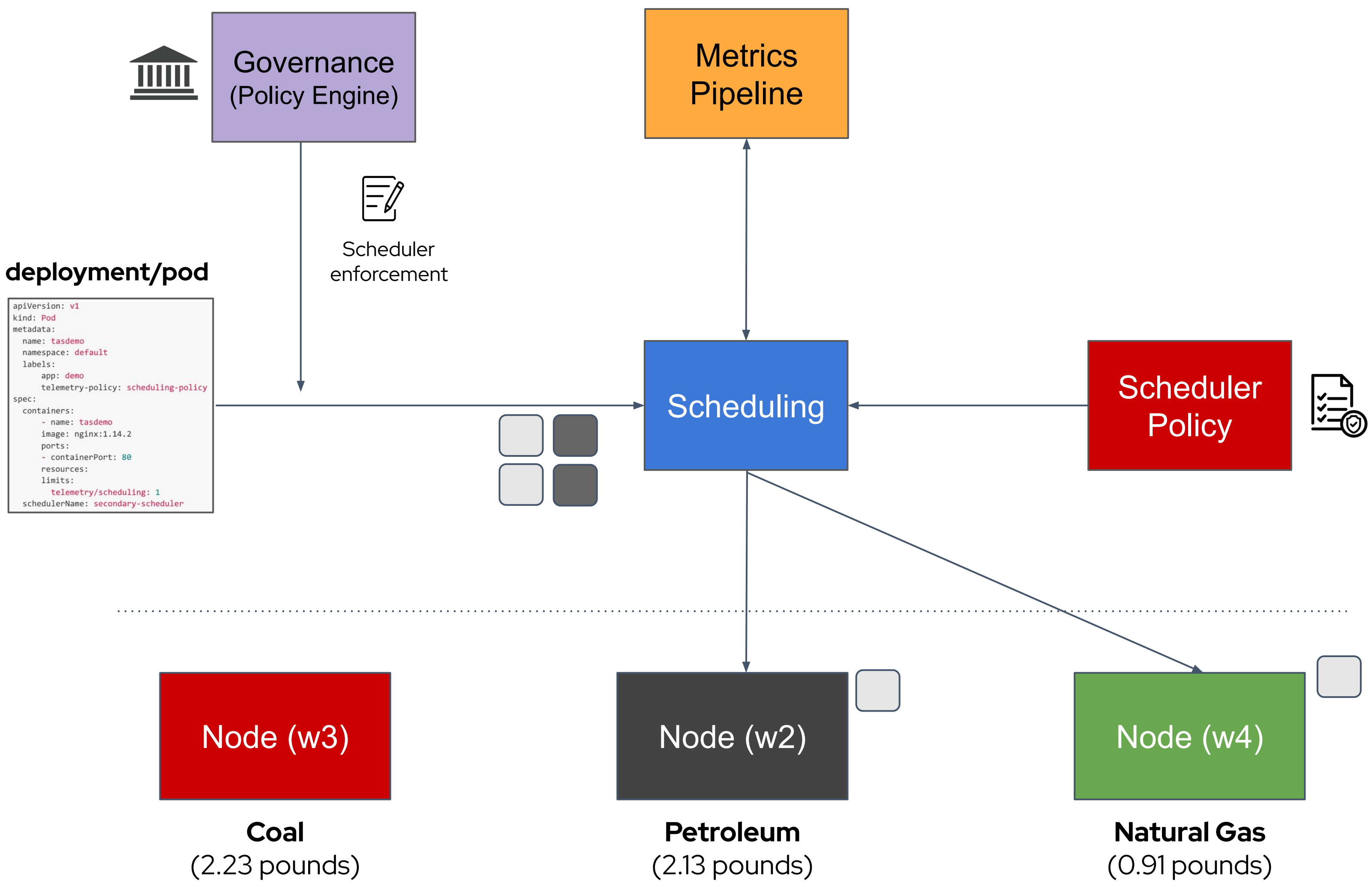
COMPONENTS

Governance
Policy engine to enforce scheduler usage for emission aware placement.

Scheduler Policy
Define strategies and rules for scheduling, descheduling and “dont schedule”. Policies are consumed by the Scheduling component.

Scheduling
Policy driven scheduling. Consumes policies, filters nodes in violation and makes placement/deschedule decision.

Metrics Pipeline
Provides metrics to be used in the policies and consumed by the scheduler.



Metrics Pipeline

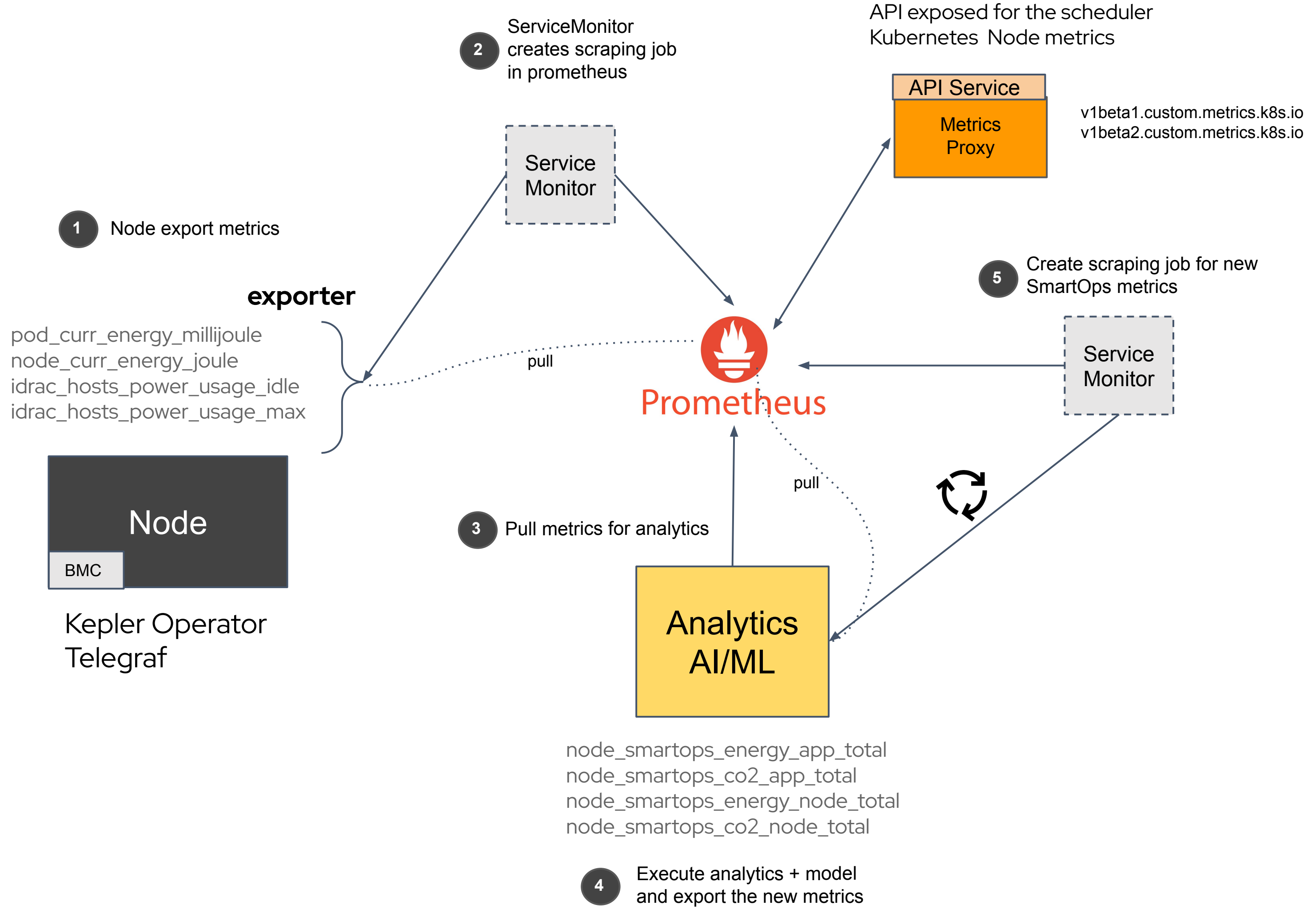
METRICS PIPELINE

KEPLER (K8s Efficient Power Level Exporter)
Uses eBPF to probe energy related system stats and export as Prometheus metrics.

Telegraf
Server agent for collecting and reporting metrics. Uses custom SNMP plugin to pull power usage metrics from iDRAC BMC.

Analytics AI/ML Agent
Custom analytics agent. Pulls metrics from Prometheus. Add penalties based on node type. Feeds AI/ML model and export SmartOPS metrics.

Metrics Proxy
Convert prometheus metrics into Kubernetes Node Metrics to be consumed by the scheduler.



Governance / Scheduling

GOVERNANCE / SCHEDULING

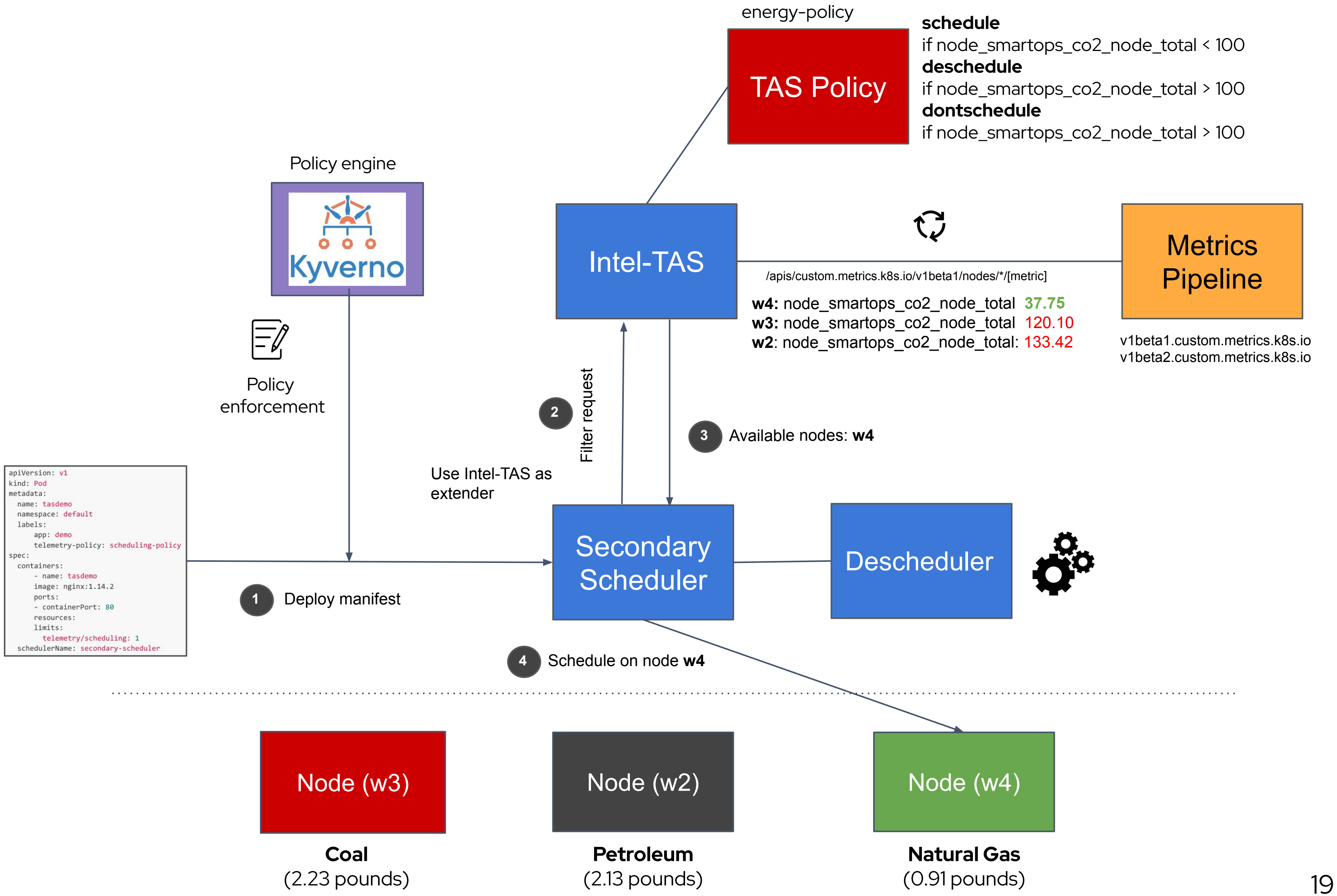
Kyverno
Policy engine. Used to enforce usage of secondary-scheduler and the energy-policy by injecting additional specs on pod manifest

Secondary Scheduler Operator
Customized scheduler image using scheduler plugin KubeSchedulerConfiguration

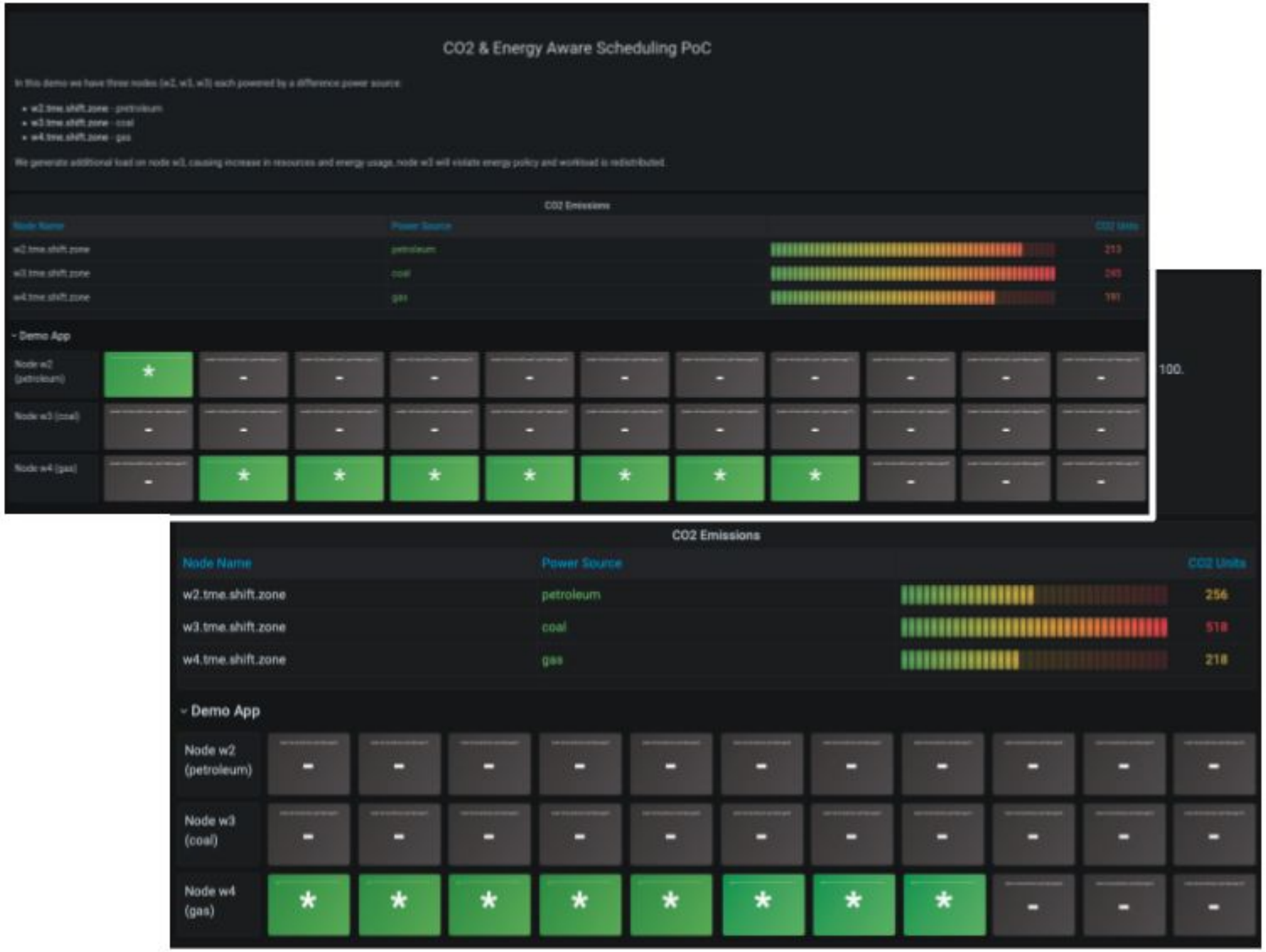
Descheduler Operator
Deschedule based on affinity/anti-affinity rules

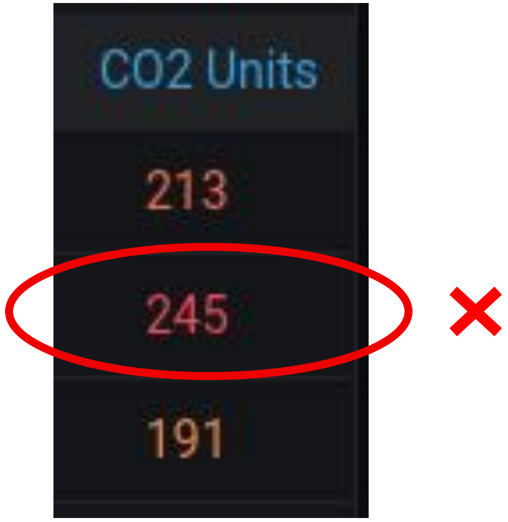
Intel-TAS (Telemetry-Aware Scheduler)
Extender implementation for K8S scheduler that consumes platform metrics and makes intelligent scheduling/descheduling decisions based on defined policies

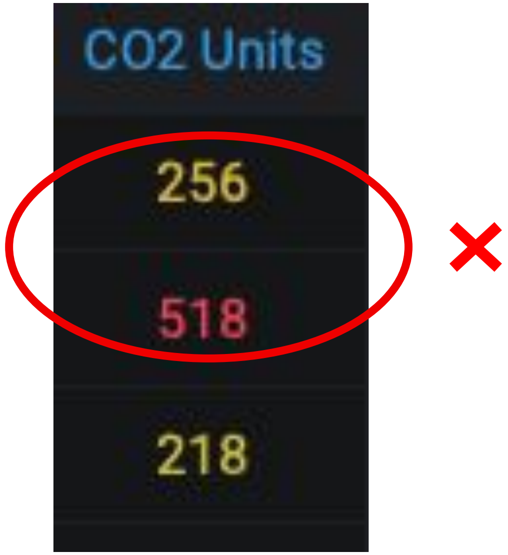
TAS Policy
Policy to define rules for scheduling or de-scheduling decisions based on metrics.



POC w/Goal of CO2 < 220 units per Node







Referenced Projects

- **Kubernetes-based Efficient Power Level Exporter (Kepler)**
  - https://sustainable-computing.io
- **Kubernetes Event-driven Autoscaling (KEDA)**
  - https://keda.sh
  - **Carbon Aware Scaling POC** (KEDA + Kepler)
    - https://www.youtube.com/playlist?list=PLhiQtL6gPrYrdpfhCoxD9ydT6TqNQBA4A
- **Intel Telemetry Aware Scheduling (TAS)**
  - https://github.com/intel/platform-aware-scheduling
  - **Intel Intent Driven Orchestration Planner** (for future integrations)
    - https://github.com/intel/Intent-Driven-Orchestration
- **Open DataHub (ODH)** (a Data Science & ML Platform)
  - https://opendatahub.io
  - ***Kubeflow**, Kafka, **Prometheus, Grafana, ArgoCD**, Tekton, Knative, **Jupyter Notebooks**, Elyra, **sklearn**, Seldon, Tensorflow, Airflow, Apache Superset, Hue, and many others.*

Care about

green computing?

Let's work together @

CNCF **ENVIRONMENTAL SUSTAINABILITY** TAG

Please scan the QR Code above to leave feedback on this session