SAS® Workload Management in Kubernetes

Hans-Joachim Edert
SAS Global Technology Practice EMEA
27.01.2022

SAS® Workload Management in Kubernetes

Agenda

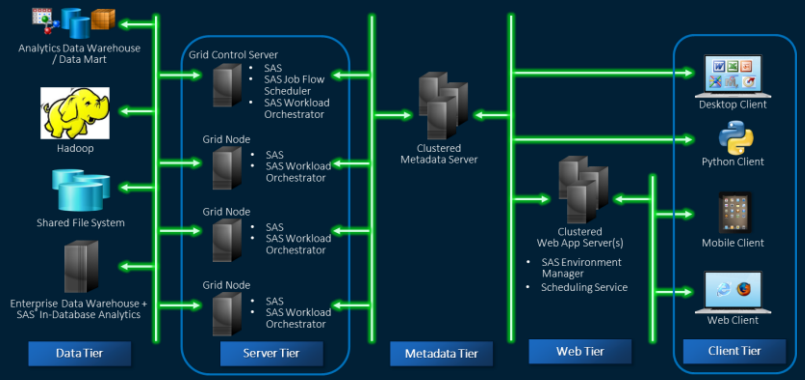
- SAS Workload Management on Kubernetes
  - Compared to SAS Grid (SAS 9.4)
  - Adding to Kubernetes facilities
  - Architecture
- Live Demo
- Roadmap
- Q & A

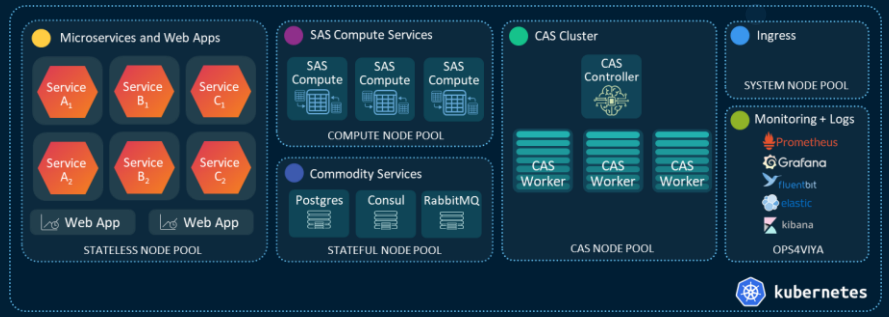
SAS Software editions

Version Overview

- SAS 9.4m8 ← SAS Grid
- SAS Viya 3.5
- SAS Viya 4 ← SAS WLM
  - Fully based on Kubernetes
  - Latest release
  - K8s distributions support
    - 11.2020 for Microsoft Azure (AKS)
    - 05.2021 for GCP and AWS (EKS)
    - 11.2021 for Red Hat OpenShift
    - 01.2022+ for other Kubernetes distributions (Open source, Tanzu …)

SAS Grid & SAS Workload Management





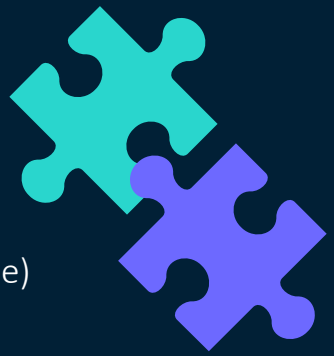
Different architecture, similar objectives & capabilities

More possibilities on Viya 4

Some SAS Grid features to be supported in future

SAS Workload Management for Viya

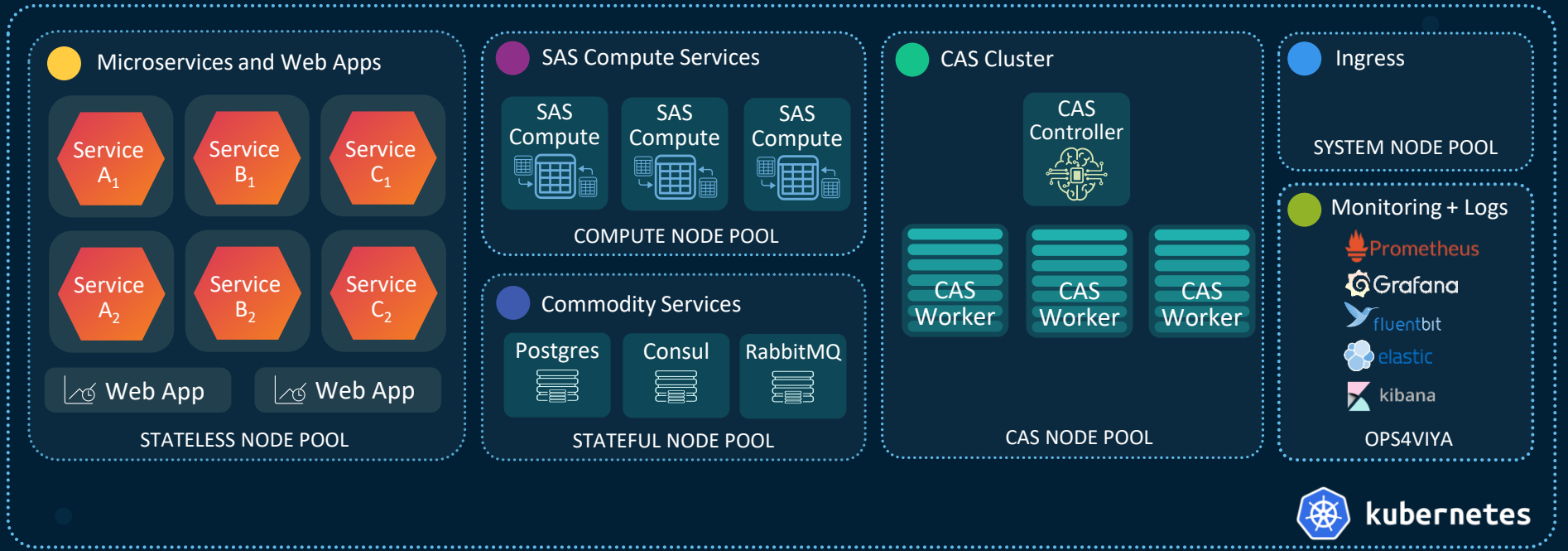
Workload Management for Kubernetes – complementary capabilities



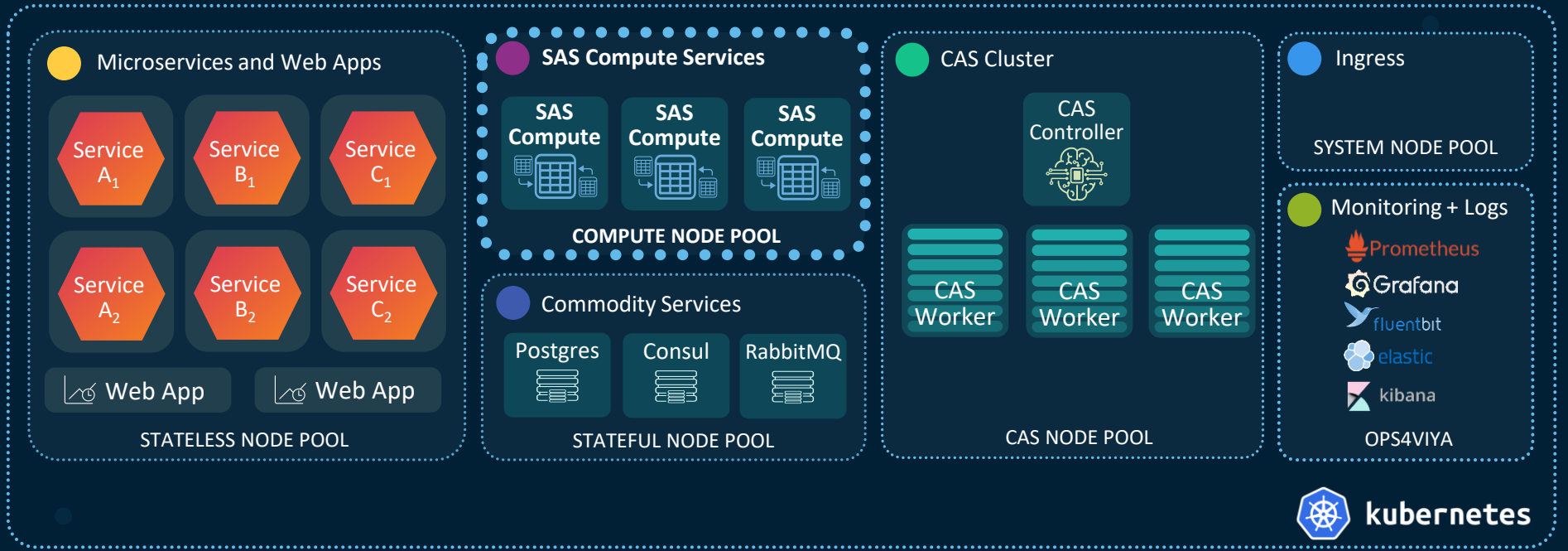
- Kubernetes
  - Strong on the infrastructure layer
    - Scaling, ensuring high availability
  - Rather weak on supporting higher-level workload management
    - API only knows about „Job“ (one time launch) and „CronJob“ (time-based schedule)
    - Can‘t chain pod execution, no flow logic („if then“)
    - Can only restart pod as a whole

- SAS Workload Management
  - Relies on stable infrastructure layer, profits from scalability and high availability capabilities
  - Adds advanced workload management features to Kubernetes, e.g. queues, prioritization, preempt/restart, optimization (AI/ML)

SAS Viya deployed in Kubernetes

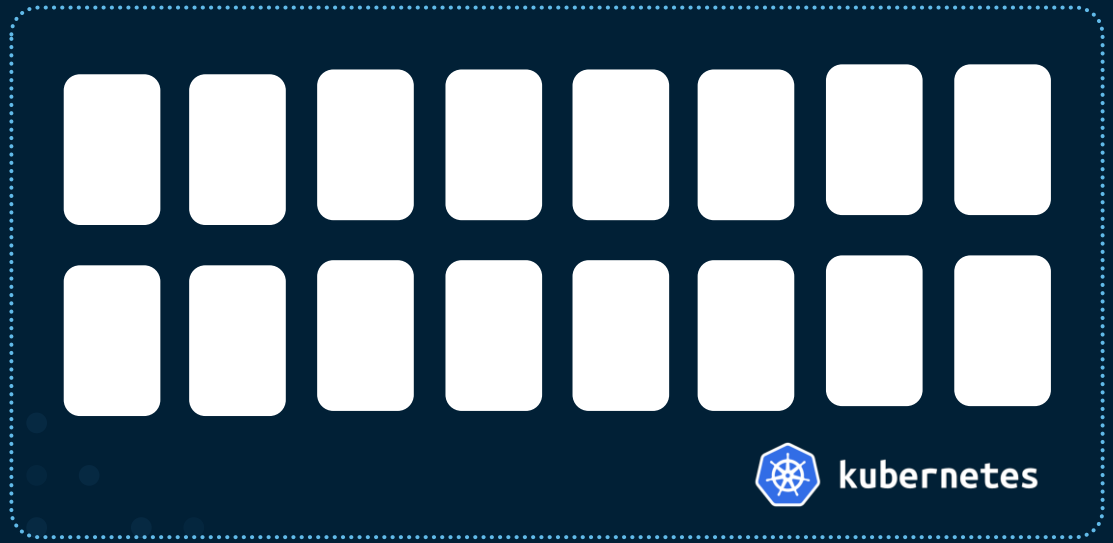


SAS Workload Management – focused on SAS Compute

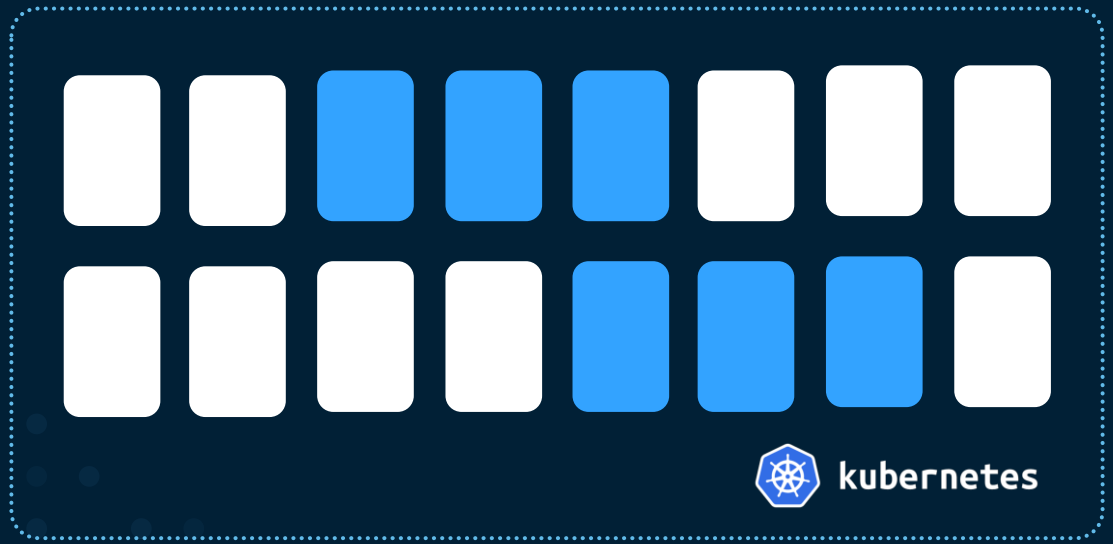


Demo Flow

- This is a Kubernetes cluster with 16 nodes

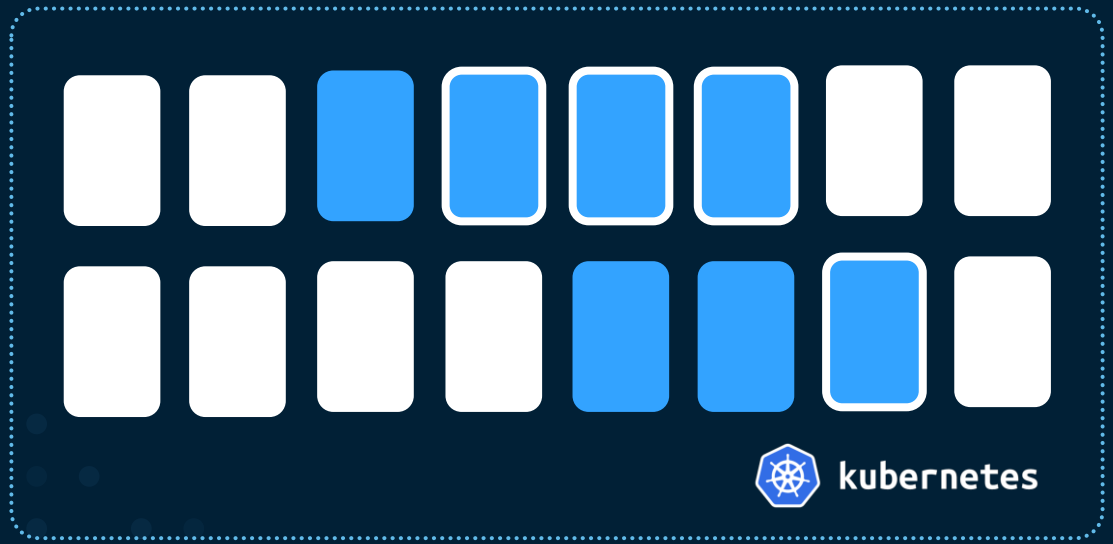


Demo Flow



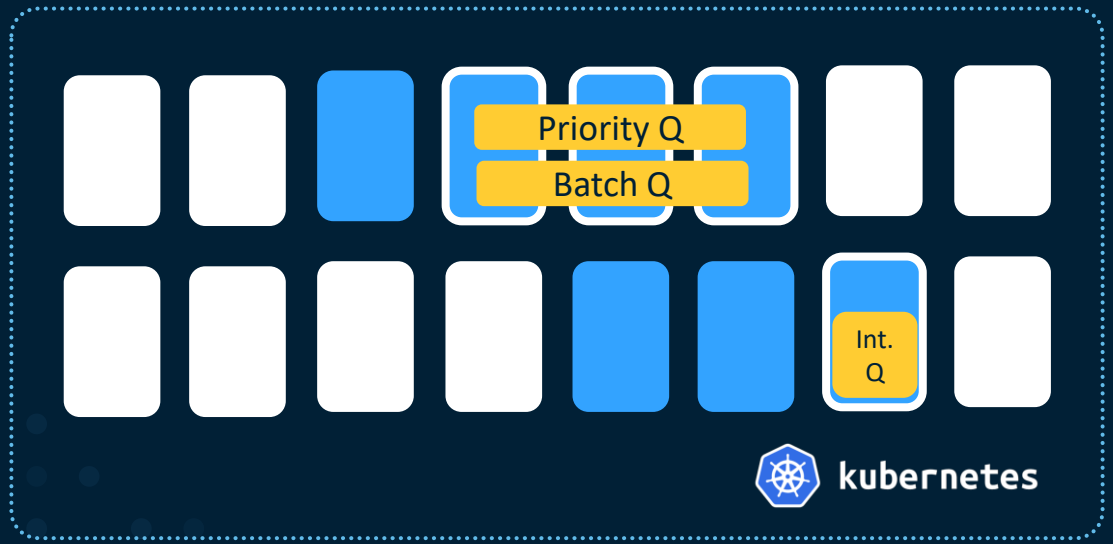
- This is a Kubernetes cluster with 16 nodes
- You've been granted 6 of them for deploying SAS Viya

Demo Flow



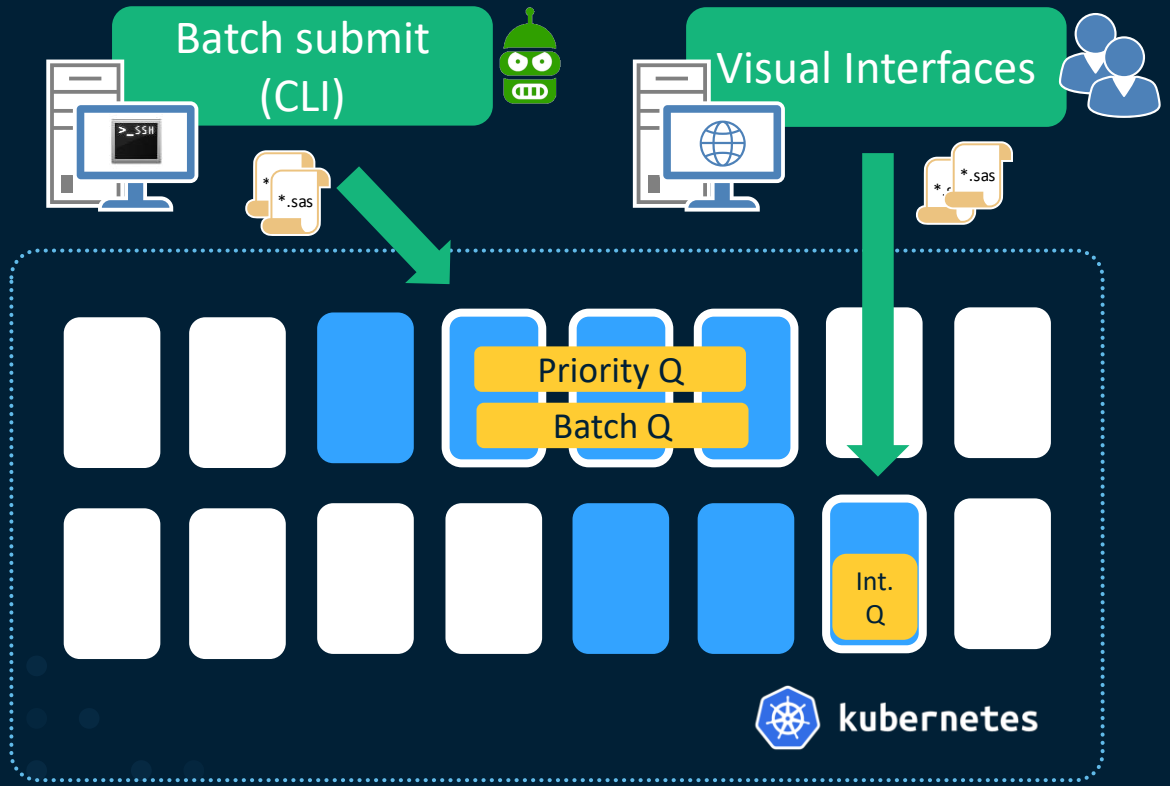
- This is a Kubernetes cluster with 16 nodes
- You've been granted 6 of them for deploying SAS Viya
- You (as the SAS admin) have decided to use 4 of them for running SAS programs (= the SAS compute nodepool)

Demo Flow

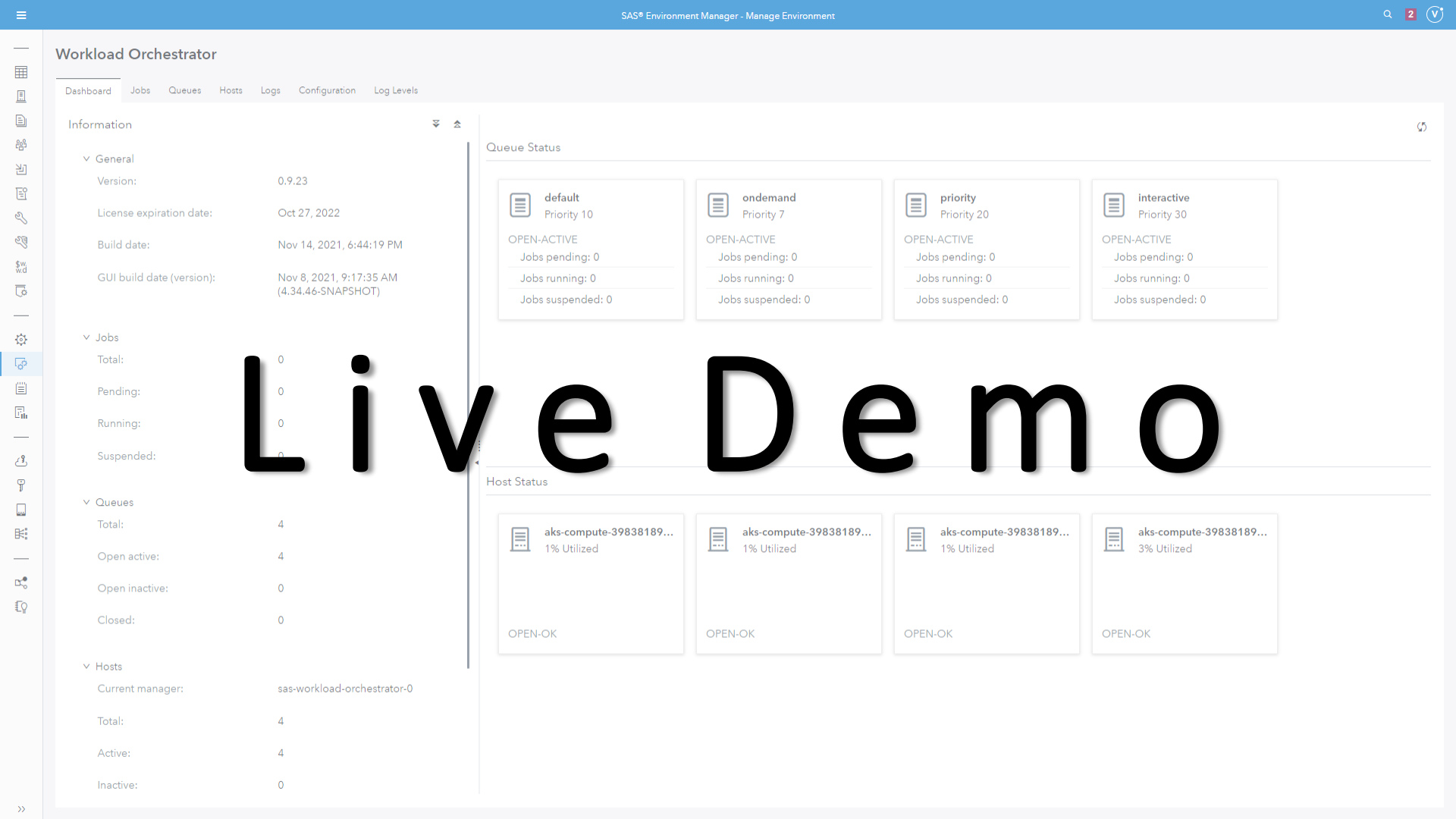


- This is a Kubernetes cluster with 16 nodes
- You've been granted 6 of them for deploying SAS Viya
- You (as the SAS admin) have decided to use 4 of them for running SAS programs (= the SAS compute nodepool)
- You set up queues according to your expected workload and assign them to nodes
  - E.g. decide on the ratio batch vs. interactive
  - E.g. set up "priority lanes"

Demo Flow



- This is a Kubernetes cluster with 16 nodes
- You've been granted 6 of them for deploying SAS Viya
- You (as the SAS admin) have decided to use 4 of them for running SAS programs (= the SAS compute nodepool)
- You set up queues according to your expected workload and assign them to nodes
  - E.g. decide on the ratio batch vs. interactive
  - E.g. set up "priority lanes"
- Depending on the source, jobs will land in a specific queue on a specific set of nodes
  - With preemption used if needed


SAS® Environment Manager - Manage Environment
Workload Orchestrator
Dashboard
Jobs
Queues
Hosts
Logs
Configuration
Log Levels
Information
General
Version: 0.9.23
License expiration date: Oct 27, 2022
Build date: Nov 14, 2021, 6:44:19 PM
GUI build date (version): Nov 8, 2021, 9:17:35 AM (4.34.46-SNAPSHOT)
Jobs
Total: 0
Pending: 0
Running: 0
Suspended: 0
Queues
Total: 4
Open active: 4
Open inactive: 0
Closed: 0
Hosts
Current manager: sas-workload-orchestrator-0
Total: 4
Active: 4
Inactive: 0
Queue Status
default
Priority 10
OPEN-ACTIVE
Jobs pending: 0
Jobs running: 0
Jobs suspended: 0
ondemand
Priority 7
OPEN-ACTIVE
Jobs pending: 0
Jobs running: 0
Jobs suspended: 0
priority
Priority 20
OPEN-ACTIVE
Jobs pending: 0
Jobs running: 0
Jobs suspended: 0
interactive
Priority 30
OPEN-ACTIVE
Jobs pending: 0
Jobs running: 0
Jobs suspended: 0
Live Demo
Host Status
aks-compute-39838189...
1% Utilized
OPEN-OK
aks-compute-39838189...
1% Utilized
OPEN-OK
aks-compute-39838189...
1% Utilized
OPEN-OK
aks-compute-39838189...
3% Utilized
OPEN-OK

ROADMAP HIGHLIGHTS

- Migration from SAS 9 Grid Manager
- CAS on-demand
- Optimization of cloud resources
  - Cost thresholds
  - Performance SLAs
  - Static profiling
  - Historic profiling

Roadmap subject to change, consult product management and official roadmap for more information

Summary

SAS Workload Management on the Kubernetes platform

- Building on-top of Kubernetes capabilities
- Manages all types of compute services
  - Interactive sessions (launched by applications (e.g. SAS Studio))
  - SAS/Connect
  - Batch sessions (launched using the batch CLI)
- Adds advanced workload management features to Kubernetes
  - Queues
  - Prioritization (preempt/restart)
- Roadmap for 2022 focus points
  - Migration support
  - Adding CAS sessions
  - Optimization using analytics (AI/ML)

Thank you for your time!