



KubeCon



CloudNativeCon

Europe 2026

#KubeCon #CloudNativeCon

Advancing Kubernetes AI Conformance: Current State and Roadmap

Janet Kuo, Staff Software Engineer, Google

Yuan Tang, Senior Principal Software Engineer, Red Hat

Mario Fahlandt, Customer Delivery Architect, Kubermatic

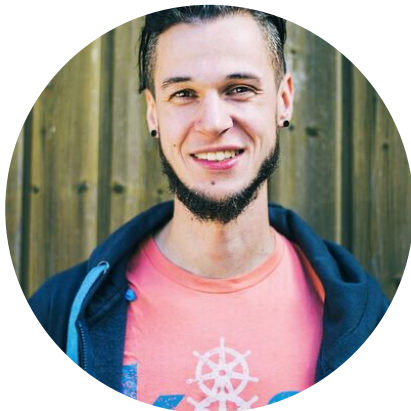




Janet
Kuo



Yuan
Tang



Mario Fahlandt



Rita
Zhang



KubeCon



CloudNativeCon

Europe 2026

Intro



How It Started



KubeCon



CloudNativeCon

Europe 2026

- AI introduces new requirements: GPU-intensive, synchronously scheduled workloads, specialized hardware, networking security, etc.
- Without standardized conformance, organizations suffered from significant fragmentation.
- This fragmentation resulted in major difficulties when migrating AI models and created vendor lock-in.

Kubernetes AI Conformance



KubeCon



CloudNativeCon

Europe 2026

- Kubernetes Conformance isn't enough
 - Some parts are outside of Kubernetes core
 - Faster iterations
- WG AI Conformance was officially established in 2025 to solve this
 - Sponsored by SIG Architecture
- Original AI Conformance proposal (SHOULD & MUST requirements)
- Additional requirements for *platforms*, on top of Kubernetes Conformance
 - It doesn't limit how end users want to run AI on Kubernetes
 - It require platforms to support capabilities / APIs most use cases need

- Define a standardized set of capabilities, APIs, and configurations for running AI/ML workloads on Kubernetes.
- Aims to simplify AI/ML operations on Kubernetes, accelerate adoption, guarantee interoperability and portability for AI workloads, and enable ecosystem growth on an industry-standard foundation.
- “Write once, run anywhere”, for AI workloads too!

The Current State



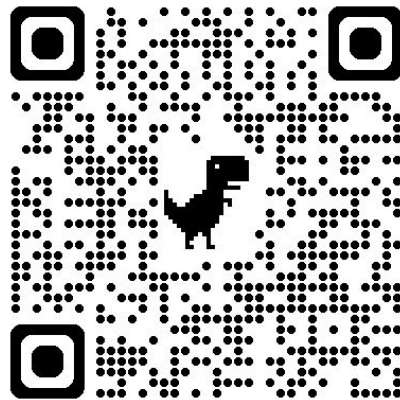
KubeCon



CloudNativeCon

Europe 2026

- Working Group is temporary in nature, and can't own code
- Now a SIG Architecture Subproject
- A permanent home for the program's long-term maintenance and evolution (same path as Kubernetes Conformance)
- <https://github.com/kubernetes-sigs/ai-conformance>





KubeCon



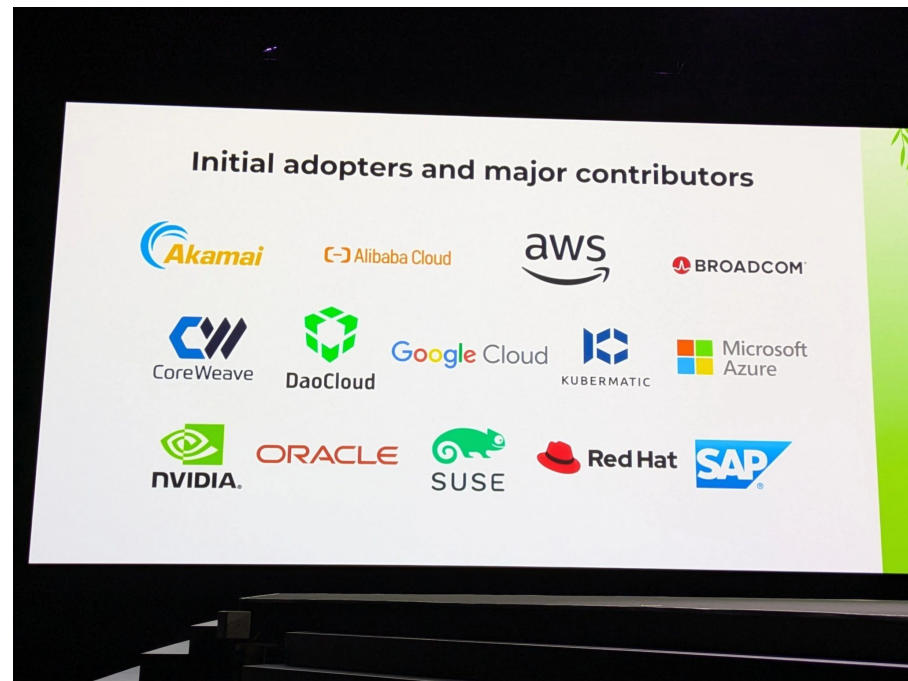
CloudNativeCon

Europe 2026

What have we accomplished?



Program Launch at KubeCon North America 2025



<https://www.cncf.io/announcements/2025/11/11/cncf-launches-certified-kubernetes-ai-conformance-program-to-standardize-ai-workloads-on-kubernetes/>

Certified Kubernetes - AI Platform in CNCF Landscape

Platform	Certified Kubernetes - AI Platform									
	 DaoCloud	 Giant Swarm	 KUBERMATIC	 Palette	 RANCHER	 Red Hat OpenShift	 vmware by Broadcom	 aws	 Azure Kubernetes Service (AKS)	 BAIDU AI CLOUD
	 CoreWeave	 Google Cloud	 linode	 ORACLE	 OVHcloud	 联通云 Unicom Cloud	 Gardener	 SIDERO	 Alibaba Cloud	

<https://landscape.cncf.io/?group=certified-partners-and-providers>

<https://github.com/cncf/landscape>

Conformance Checklists



KubeCon

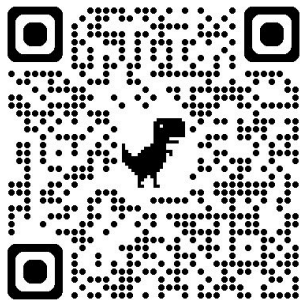


CloudNativeCon

Europe 2026

Defined and released conformance checklists for multiple Kubernetes versions, providing a clear baseline for AI-ready clusters:

- [AIConformance-1.33.yaml](#)
- [AIConformance-1.34.yaml](#)
- [AIConformance-1.35.yaml](#)



<https://github.com/cncf/k8s-ai-conformance>

metadata:

```
kubernetesVersion: v1.35
platformName: "[Platform Name]"
platformVersion: "[Platform Version]"
vendorName: "[Vendor Name]"
```

...

spec:

accelerators:

- id: dra_support
description: "Support DRA APIs ..."
level: MUST
status: ""
evidence: []
notes: ""
- id: driver_runtime_management
description: "Provide a verifiable mechanism"
level: SHOULD

...

Requirement Definition and Graduation Process



KubeCon



CloudNativeCon

Europe 2026

- Transitioned requirement tracking from documents to a [structured GitHub Project board](#).
- Defined the [Kubernetes AI Conformance Requirement \(KAR\)](#) process, mirroring the KEP process, to manage the lifecycle of requirements (SHOULD/MUST).

KAR-NNNN: Your short, descriptive title

Description

Motivation

Graduation Criteria

SHOULD

- ☐ Describe how users can test it for self-attestation with scripts, documentation, etc
- ☐ Starting v1.37, new SHOULDs must include proposed automated tests in the automated tests section below

MUST

- ☐ Starting v1.37, new MUSTs must include automated tests that have been added to the AI conformance test suite
- ☐ Demonstrate at least two real-world usage of SHOULD before graduating to MUST
- ☐ Kubernetes core APIs must be GA

Test Plan

How We Might Test It

Automated Tests

Implementation History

🔗 Related KARs



KubeCon



CloudNativeCon

Europe 2026

What's next?



What's next?



KubeCon



CloudNativeCon

Europe 2026

1.36: Inference

- [KAR-0010: High-Performance Pod-to-Pod Communication](#)
- [KAR-0011: Advanced Inference Ingress](#)
- [KAR-0041: Disaggregated Inference Support](#)

1.37: agents?

Automate tests

```
go test -v ./test [-run <TestName>]
```

```
[-kubeconfig=<path/to/kubeconfig>] [-accelerator-type=<type>]
```

We need more conformance tests for specific areas, and more standardization.



KubeCon



CloudNativeCon

Europe 2026

Conformance programs in a nutshell

Why? & Who?



KubeCon



CloudNativeCon

Europe 2026

Why?

Improves compatibility among solutions built upon the project

Guarantees that users get the expected functionality and a consistent interface, even when moving from one vendor offering to another.

Creates a defined standard for what it means to be a "conformant" or "certified" solution in your projects ecosystem.

Who?

The Project

The Governing Board of the CNCF

Project Team of CNCF

If it is overarching the TOC



KubeCon



CloudNativeCon

Europe 2026

How to get involved?

How to get involved



KubeCon



CloudNativeCon

Europe 2026

- AI Conformance Biweekly Meeting
 - Thursdays at 7 PM CET | 10 AM PDT
- Communications:
 - Slack: #ai-conformance on K8s Slack
 - Mailing list: sig-architecture mailing list
- <https://github.com/kubernetes/community/tree/main/sig-architecture>





KubeCon



CloudNativeCon

Europe 2026

