# Building a Multi-Tenant Kubernetes Platform on Bare Metal with CloudWorks and vCluster

taikun

Solution Overview

This solution brief outlines how enterprises and service providers can deliver a secure, scalable, and efficient multi-tenant Kubernetes platform on bare metal infrastructure using Taikun CloudWorks and vCluster. The architecture enables seamless tenant isolation, self-service application delivery, GPU acceleration for AI/ML, and deep automation and observability, empowering both platform teams and end users.

Architecture at a Glance
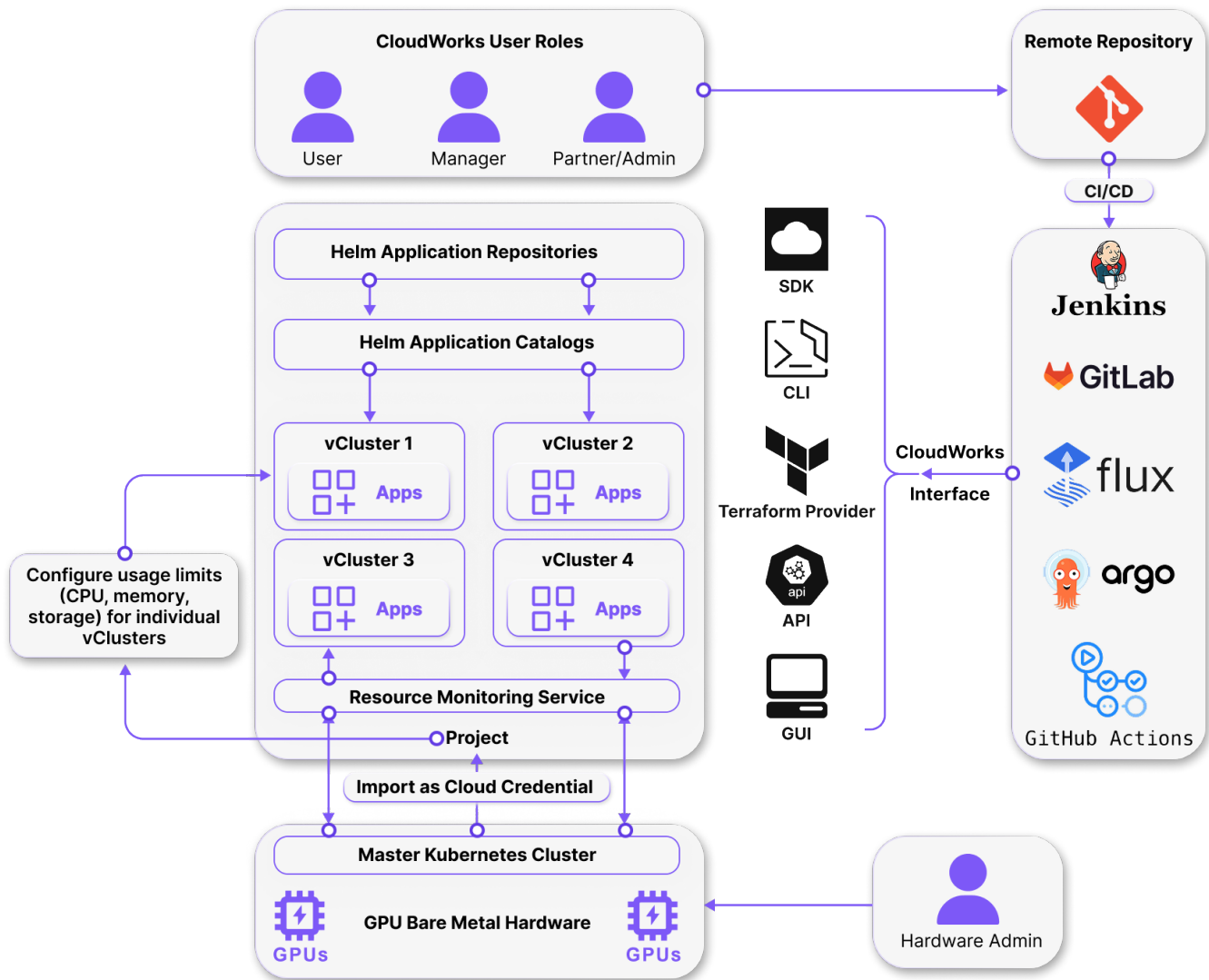
Key Solution Components

1. Bare Metal Parent Cluster

- **Master Kubernetes Cluster on Bare Metal:** Serves as the foundational parent (host) cluster, providing direct access to hardware resources—including GPUs—for maximum performance and efficiency.

- **GPU Node Integration:** GPU hardware is available for scheduling AI/ML workloads directly from tenant vClusters, supporting high-performance and data-intensive use cases.

2. CloudWorks Import and Multi-Tenancy

- **Import as Cloud Credential:** The bare metal Kubernetes cluster is imported into CloudWorks as a reusable cloud credential. This enables it to serve as the host for multiple, isolated vClusters, each mapped to a different project, tenant, or team.

- **Multi-Project Flexibility:** The same parent cluster can be imported and assigned to multiple projects, maximizing utilization and supporting organizational boundaries.

3. vCluster Provisioning and Resource Governance

- **vClusters per Tenant:** Each tenant, project, or team receives a dedicated vCluster within the parent cluster. These vClusters are fully isolated, with their own Kubernetes API, RBAC, and resource quotas.

- **Resource Usage Limits:** Platform teams can configure per-vCluster and namespace-level quotas for CPU, memory, and storage, preventing noisy neighbor issues and ensuring predictable multi-tenant operations.

- **Integrated Monitoring Service:** Continuous, centralized monitoring of both the parent cluster and all vClusters, with real-time dashboards for resource usage, health, and persistent volumes.

4. Application Delivery and Self-Service

- **Helm Application Repositories & Catalogs:** Platform engineers curate application catalogs from both public and private Helm repositories. These catalogs are bound to vClusters, empowering tenants to deploy approved applications through a self-service portal.

- **Self-Service App Deployment:** Tenants can deploy, update, and manage applications from curated catalogs via the CloudWorks GUI, CLI, or API—accelerating development and reducing operational bottlenecks.

5. Automation and CI/CD

- **CI/CD Integration:** Native support for remote repositories (GitHub, GitLab) and CI/CD tools (Jenkins, Flux, Argo, GitHub Actions) enables automated application delivery and GitOps workflows.

- **Automation Interfaces:** Full automation is available through SDKs, APIs, and Terraform providers, supporting Infrastructure-as-Code and advanced DevOps use cases.

6. Observability and Operations

- **Role-Based Access Control (RBAC):** Granular user roles (User, Manager, Partner/Admin) ensure secure delegation and operational boundaries across the platform.

- **LiveOps and Day-2 Operations:** Users can interact with resources, scale applications, and troubleshoot—all from the CloudWorks dashboard, with full visibility into YAML, logs, and resource status.

End-to-End Tenant Workflow

1. **Import Bare Metal Cluster:** Register the bare metal Kubernetes cluster as a cloud credential in CloudWorks, making it available as the parent for vCluster creation.

2. **Provision vClusters:** For each tenant, create a vCluster within the parent cluster. Assign quotas, RBAC, and bind relevant Helm catalogs.

3. **App Delivery:** Tenants deploy applications from curated Helm catalogs into their vClusters, using the CloudWorks UI, CLI, API or Terraform.

4. **GPU-Enabled AI/ML:** Tenants schedule AI/ML workloads to GPU nodes within their vCluster, leveraging direct hardware access for maximum performance.

5. **Automation & CI/CD:** Developers and DevOps teams integrate their CI/CD pipelines, enabling automated deployment and GitOps workflows.

6. **Monitoring & Day-2 Operations:** Platform teams and tenants monitor usage, scale workloads, and manage resources through real-time dashboards and LiveOps features.

Key Benefits

Feature	Value
Performance & Efficiency	Bare metal and GPU passthrough deliver maximum speed and resource utilization
Secure Multi-Tenancy	vClusters provide strong workload isolation and governance
Self-Service App Delivery	Helm catalogs empower tenants to deploy and manage apps independently
AI/ML Ready	GPU integration and catalog-driven AI/ML tooling accelerate innovation
Unified Management	CloudWorks delivers a single interface for all operations, monitoring, and billing
Automation & GitOps	Full CI/CD and Infrastructure-as-Code support for platform and tenant teams
Cost Optimization	Higher resource density, reduced licensing, and efficient GPU sharing

Conclusion

CloudWorks together with vCluster(s) enable organizations to maximize the value of bare metal infrastructure by offering secure, efficient, and fully automated multi-tenant Kubernetes as a platform. Tenants gain self-service vClusters, curated app delivery, GPU-powered AI/ML, and seamless CI/CD integration, while the platform team retains centralized control, observability, and policy enforcement—delivering a true cloud-native experience on-premises.