JISUMEN
极算门

ENTERPRISE AI INFRASTRUCTURE / PRODUCT BRIEF 2026

JISUMEN AI Compute Control Center

JISUMEN AI COMPUTE CONTROL CENTER

A platform for model operation and cluster maintenance in multi-node private compute environments

Manage model status, cluster resources, switching workflows, and operations logs together to support ongoing operation of large models in a private environment.

- 4× DGX SPARK
- MODEL SWITCHING
- GUIDED INITIALIZATION

Exceptional compute. Gateway to intelligence.

Dongguan JISUMEN Computer Technology Co., Ltd.



Product positioning PRODUCT POSITIONING

A unified, visible and controlled private compute operations interface

POSITIONING 02

Manage model status, cluster resources, runtime actions and fault information in one console.

01

Centralized runtime status

Show the active model, node health and resource status together, reducing fragmented information.

02

Controlled model operations

Switch using registered model configurations and validate key conditions.

03

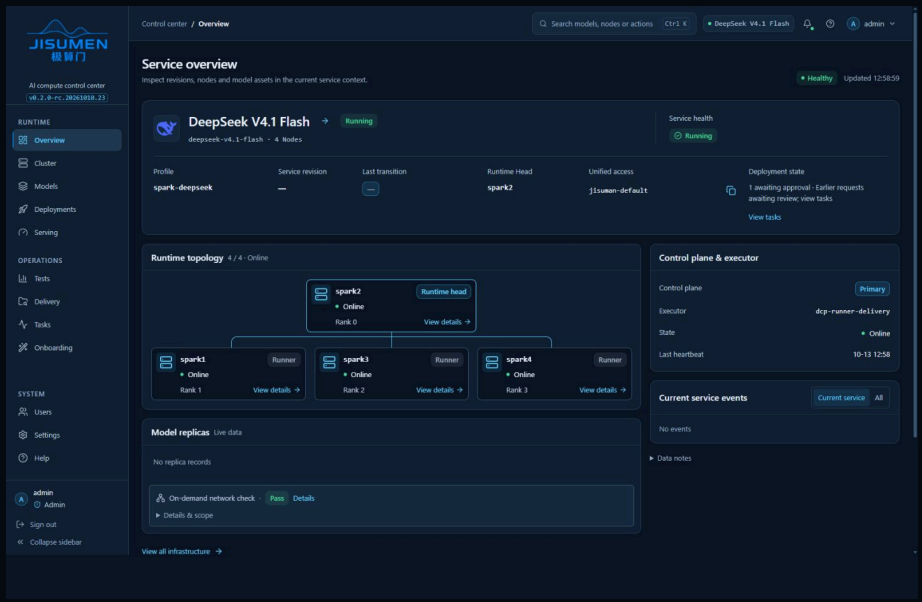
Unified issue tracking

Activity records, resource conflict warnings and raw logs provide evidence for diagnosis.

Version basis	4-node hardware demo and solution validation environment

OVERVIEW / CONTROL PLANE

LIVE INTERFACE / CURRENT BUILD



Runtime overview SYSTEM OVERVIEW

A unified view of model, cluster and engine status

OVERVIEW 03

View the active model, node health, resource use and recent activity together for consistent operational decisions.

01 Model operation
Active model and runtime status

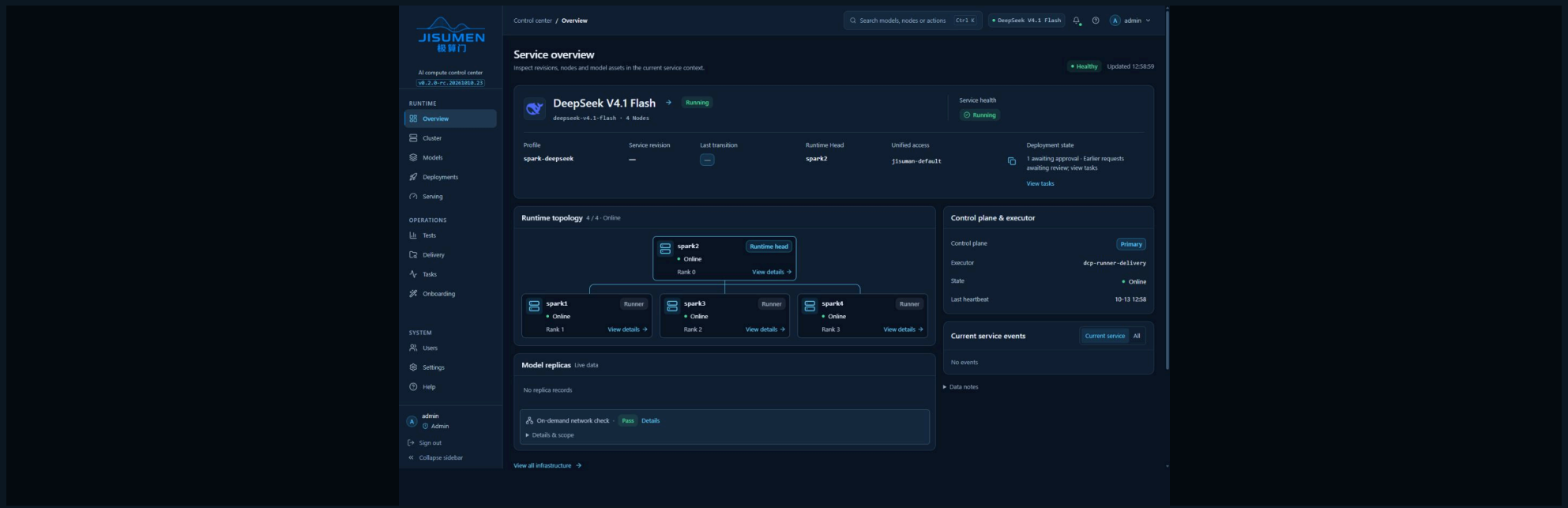
02 Cluster health
Health summary for four nodes

03 Node monitoring
Resource metrics and sampled charts

04 Activity records
Switching, checks and alerts

OVERVIEW / CURRENT CLUSTER STATE

LIVE INTERFACE / CURRENT BUILD



Data shows the demo cluster at the time of capture, and is not a fixed performance commitment.

Unified model management

MODEL ORCHESTRATION

Keep client access stable while controlling backend model switching

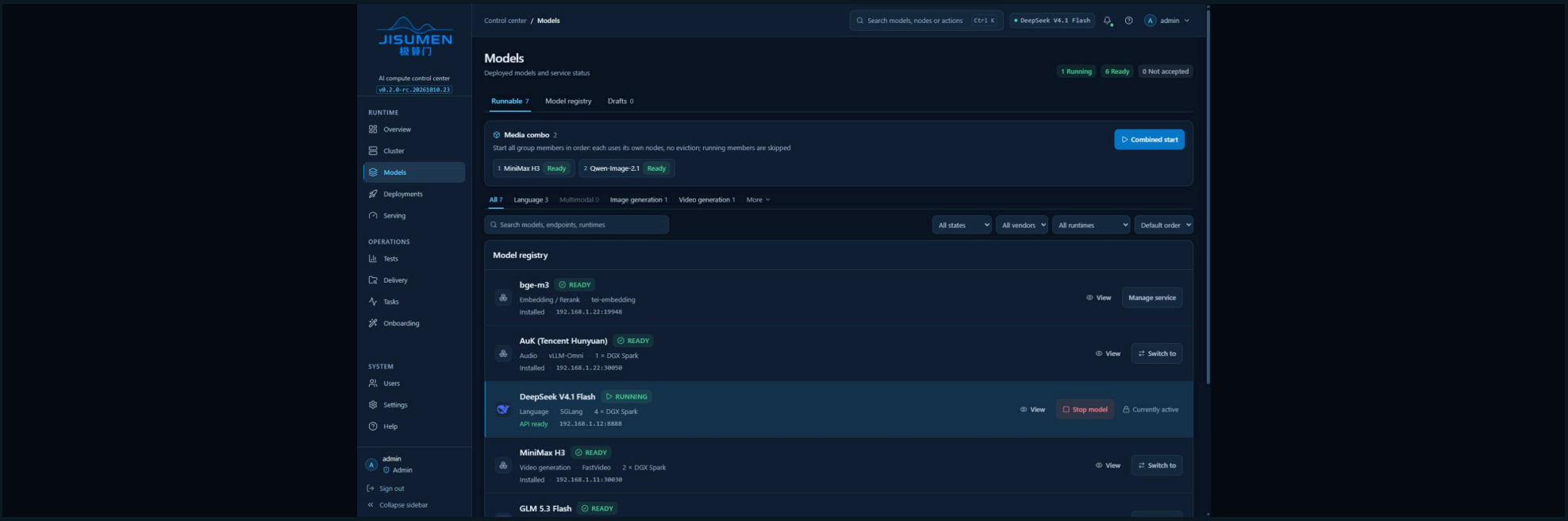
MODEL CONTROL 04

The demo has 7 runnable models covering language, image, video, audio and embedding / reranking.

Unified client entry	jisuman-default	Runnable models	7	Connection options	Cursor / Cherry Studio / Open WebUI

MODEL REGISTRY / CONTROLLED SWITCHING

LIVE INTERFACE / CURRENT BUILD



Controlled switching accounts for node resources. Models in a media group start in sequence on separate nodes without stopping one another.

Node onboarding GUIDED INITIALIZATION

A step-by-step wizard for node discovery, registration and checks

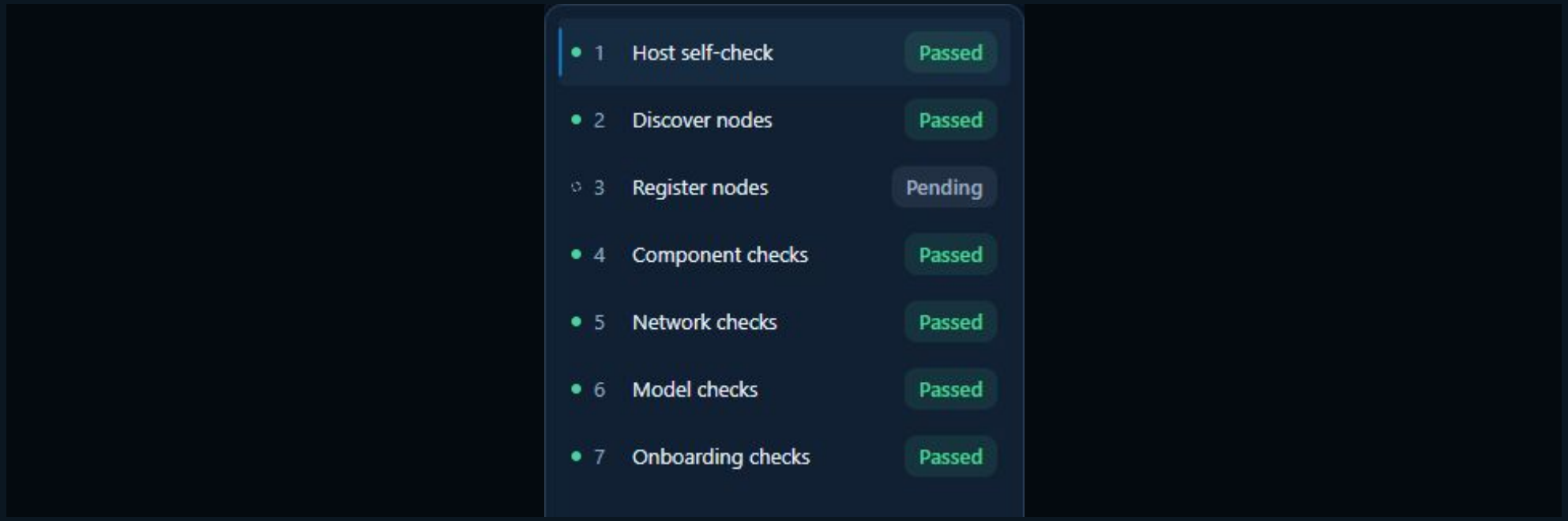
INITIALIZATION 05

Review control-host self-checks, node discovery and registration, followed by component, network, model, and onboarding checks.

01	02	03	04	05	06	07
Control-node checks	Discover nodes	Register nodes	Component checks	Network checks	Model checks	Onboarding checks

INITIALIZATION / CONTROL NODE CHECK

LIVE INTERFACE / CURRENT BUILD



The wizard discovers, registers, and checks nodes. Installation, deployment, and service delivery remain separate. The screenshot shows the workflow steps only.

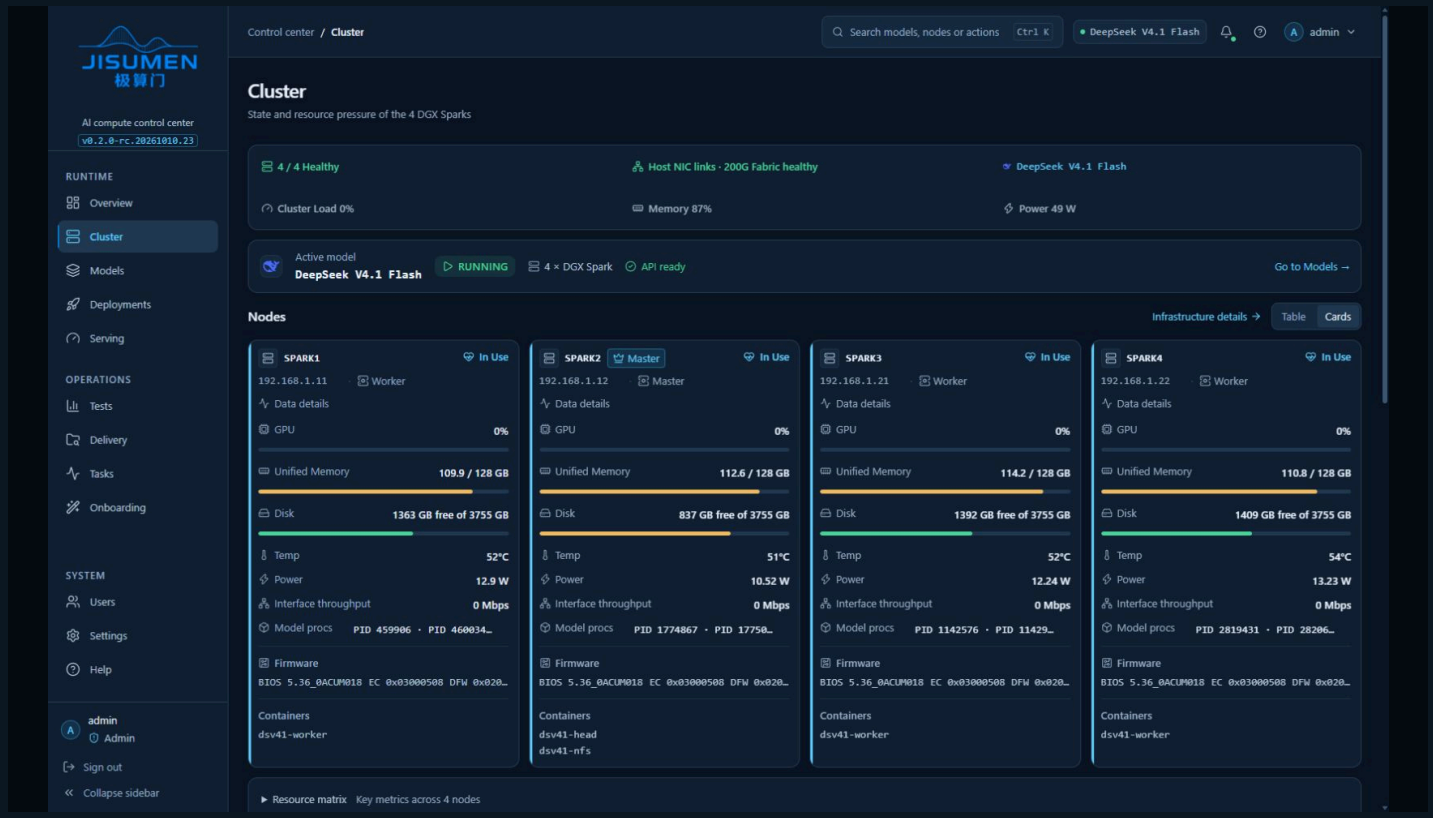
Cluster monitoring CLUSTER OBSERVABILITY

Centralized multi-node status

View node roles, resource metrics, model processes and containers together.

CLUSTER 06

CLUSTER / NODE TELEMETRY LIVE INTERFACE / CURRENT BUILD



NODE OPERATIONS

01

Node role identification

Show Master and Worker roles, node addresses and current usage.

02

Resource monitoring

Show GPU, unified memory, temperature, power and network metrics.

03

Processes and containers

Show model process PIDs and container names for quick diagnosis.

200G FABRIC

Dual rails

Environment consistency checks

PRE-FLIGHT VALIDATION

Validate node dependencies before starting a model

VALIDATION 07

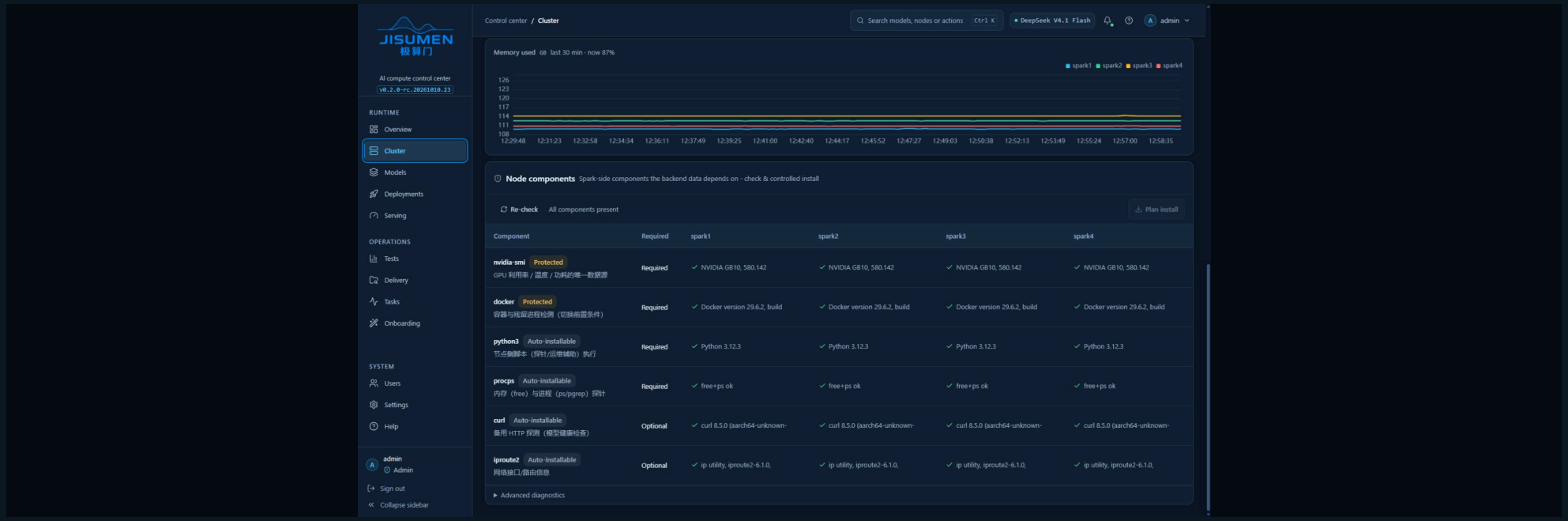
Check each node's components for high-frequency data collection and operation, improving deployment consistency.

DEPENDENCY MATRIX

- nvidia-smi
- docker
- python3
- procps
- curl
- iproute2

CLUSTER / COMPONENT CHECK

LIVE INTERFACE / CURRENT BUILD



The system distinguishes protected, automatically installable, and required/optional components for pre-deployment validation.

Resource conflict protection

WORKLOAD PROTECTION

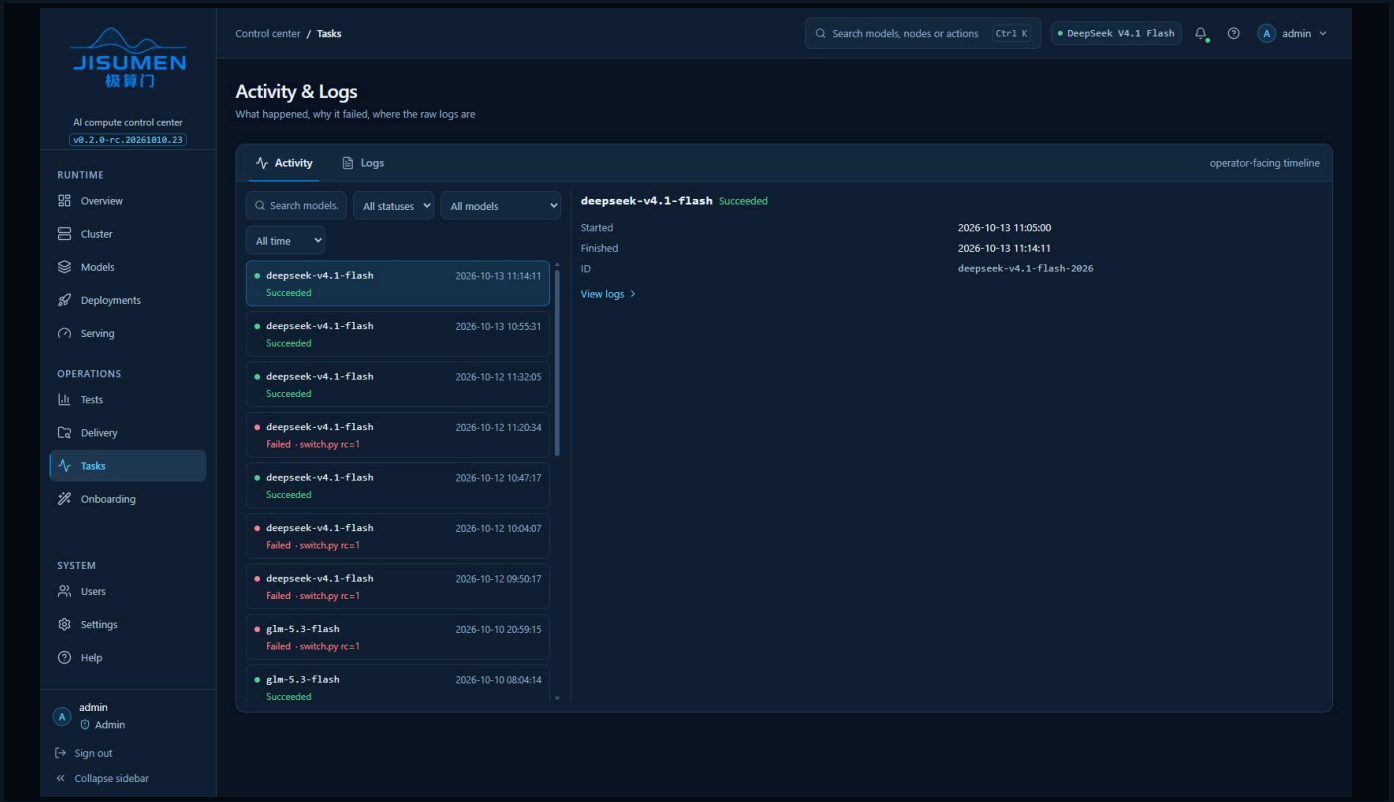
Detect external workloads to reduce resource contention

GUARD RAILS 08

If model processes outside the platform are detected, the console warns of risks and blocks direct startup of a new model.

ACTIVITY / OPERATION HISTORY

LIVE INTERFACE / CURRENT BUILD



CONTROL POLICY

BLOCK

Block startup until resources are released

Detect external model processes and residual workloads to prevent old and new models from occupying cluster resources together.

Operation records

- Switch history and results
- Failure / Running status
- Risk and failure diagnosis

Operations records OPERATIONS TRACE

Activity tracking and raw log analysis

LOGS 09

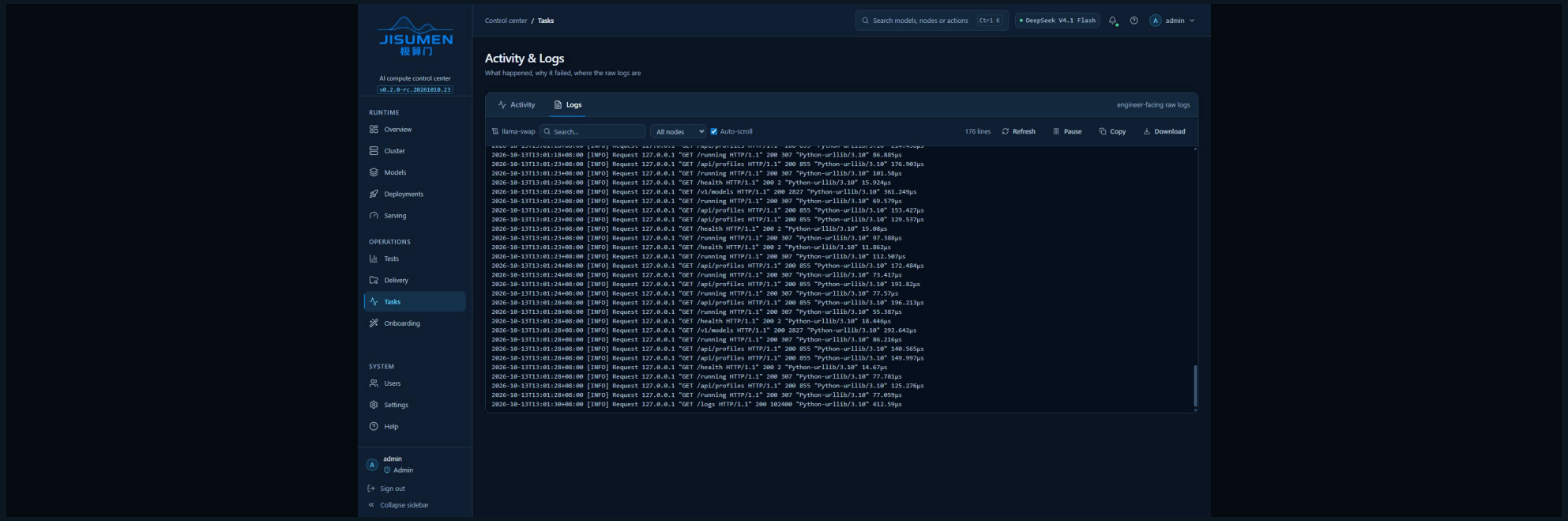
Activity status for users; raw logs and search tools for operations staff.

AVAILABLE OPERATIONS

- Search
- Filter by node
- Auto-scroll
- Refresh
- Pause
- Copy
- Download

ACTIVITY / RAW LOG STREAM

LIVE INTERFACE / CURRENT BUILD



Access control IDENTITY & ACCESS

Users, roles and permissions

ACCESS 10

Define Admin, Operator, Approver and Viewer roles, with account security and clear audit boundaries.

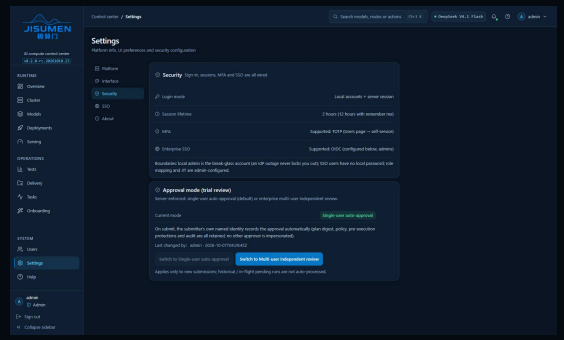
ADMIN	OPERATOR	VIEWER	
Management and operations	Runtime operations	Read-only access	Local accounts + server sessions / TOTP MFA and OIDC SSO supported

USER MANAGEMENT / ROLE ASSIGNMENT LIVE INTERFACE / CURRENT BUILD

Role & permission matrix Available actions by role

Role	View status	Switch model	Add model	Approve requests	Manage users	View audit
Admin you	✓	✓	✓	✓	✓	✓
Operator	✓	✓	✓	–	–	✓
Approver	✓	–	–	✓	–	–
Viewer	✓	–	–	–	–	–

SECURITY SETTINGS



Delivery value DELIVERY VALUE

A private compute management system built for sustained operation

DELIVERY 11

Visibility, controlled operations, issue tracking and permission boundaries support ongoing operations after model deployment.

CURRENT CAPABILITY SCOPE

Current feature scope

01 4-node cluster health and resource overview

02 Initialization wizard, step status and execution boundaries

03 Model registration, status and controlled switching

04 Unified client entry guidance

05 Node component checks and dependency matrix

06 External workload detection and startup blocking

07 Activity timeline, raw logs, roles and permissions

VERSION BOUNDARY

Version and claim boundaries

This brochure uses actual interfaces from the current demo environment with four DGX Spark systems. Features not yet integrated are not presented as available, and the demo is not described as a fully productized release.

Intended users

For individual developers, technical teams and small or medium enterprises that run large models locally and value model switching, cluster checks, safe operation and efficient troubleshooting.

JISUMEN / DGX Spark private compute deployment and operations

On-site deployment in Shenzhen and the Pearl River Delta | Remote operations | Solutions designed for models and concurrency

Exceptional compute.
Gateway to intelligence.