CAMBRIDGE NEXUS

CNEX Enterprise AI Factory Operational Readiness Overview

CAMBRIDGENEXUS

CNEX Enterprise AI Factory — Operational Readiness Overview

Executive Operations Blueprint · Enterprise Due Diligence & SLA Companion
CONFIDENTIAL — DRAFT

Enterprise AI Infrastructure Requires Enterprise Operations

AI infrastructure has moved beyond GPU procurement. As AI environments become mission-critical production platforms, enterprises increasingly evaluate providers on operational maturity, engineering scalability, resilience, and execution capability — not compute capacity alone. CambridgeNexus (CNEX) was founded to address these requirements. CNEX operates an Enterprise AI Factory Operating Platform that combines AI-driven operations, Site Reliability Engineering (SRE), strategic OEM engineering collaboration, standardized operational governance, and a global follow-the-sun delivery model. The objective is to function as an extension of a customer's own infrastructure organization across the AI infrastructure lifecycle — across CNEX-operated AI Factories, customer-owned data centers, and hybrid deployments — under a single, consistent operating standard.

Enterprise Design Objectives

Category	Operational Design Objective
Operations	24×7×365 Global AI Operations
Automation	AI-Driven Operations (Cortex™ & AboveCloud™)
Reliability	SRE & Infrastructure-as-Code
Engineering	Follow-the-Sun Delivery Model

Category	Operational Design Objective
Incident Response	SWARM™ Parallel Engineering Framework
Business Continuity	Multi-Layer DRBC
Customer Engagement	Extension of Customer Engineering Teams

Operational Design Roadmap

Operational Metric	2026 Target	2027 Target
GB300 Systems	200+	500+
Blackwell Ultra GPUs	15,000+	36,000+
Vera Rubin Systems	Planning	24
AI Factory Sites	5–10	15+
Enterprise Customers	20+	50+

Roadmap figures represent operational planning objectives, not contracted commitments.

Executive Key Takeaways

- Operational maturity, not compute access, is the basis of enterprise trust — CNEX is built around SRE discipline, automation, and governed incident response.
- SWARM™ replaces sequential vendor escalation with parallel engineering collaboration across CNEX, OEM partners, and customer teams.
- Cortex™ and AboveCloud™ provide the automation and observability layer required to operate fleets at scale without linear headcount growth.
- A multi-layer DRBC model — infrastructure, platform, geographic, operations, and

business continuity — is designed into the operating model rather than added after deployment.

Engineering Scalability & Operational Leverage

AI-Driven Operations: Cortex™ & AboveCloud™

As AI Factory environments expand from individual Super AI PODs to fleets of hundreds of GB300 systems, operational workload grows non-linearly with system count. CNEX integrates ProphetStor's Cortex™ and AboveCloud™ as the operational intelligence layer underneath every AI Factory: unified fleet monitoring collapses many independent dashboards into one operational view; policy-driven automation and infrastructure-as-code standardize configuration so consistency does not depend on any one engineer's manual process; and predictive analytics shift capacity planning from reactive to forward-looking. The combined effect is that engineering capacity scales sub-linearly with fleet size — the core requirement for enterprise-scale growth without proportional headcount expansion.

Operational Automation Matrix

Operational Function	Traditional Operations	CNEX + Cortex™ / AboveCloud™
Fleet Monitoring	Manual dashboards	Automated / Agentic-AI
Alert Correlation	Independent alerts	Automated / Agentic-AI
Provisioning	Manual configuration	Automated / Agentic-AI
Capacity Planning	Manual analysis	Automated / Agentic-AI
Configuration Management	Device-by-device	Automated / Agentic-AI
Health Validation	Periodic manual checks	Automated / Agentic-AI

Operational Function	Traditional Operations	CNEX + Cortex™ / AboveCloud™
Multi-Site Operations	Independent management	Automated / Agentic-AI

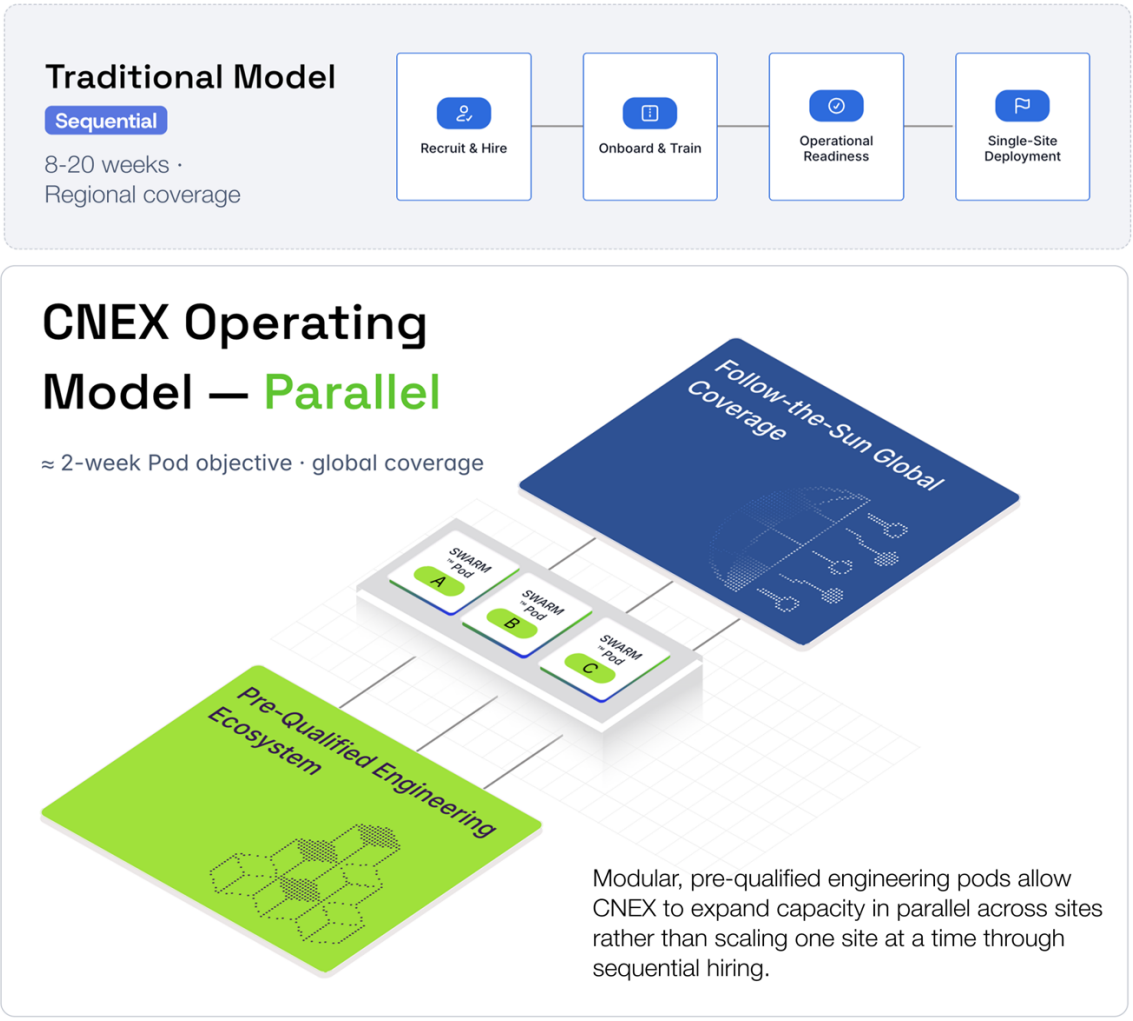
Engineering Capacity Model

AI Factory Scale	Core CNEX	SWARM™ Pods (6 Eng./Pod)	Global Engineering Ecosystem*	Enterprise Delivery Capacity
1–25 GB300	20	2	60+	80+
26–100 GB300	45	6	150+	230+
101–200 GB300	70	12	300+	440+
201–500 GB300 + 24 VR200	120	24	500+	760+

**Global Engineering Ecosystem represents collaborative engineering capacity available through strategic partnerships and is not equivalent to CNEX payroll headcount.*

CAMBRIDGENEXUS

Engineering Expansion: Traditional vs. CNEX Operating Model



Activity	Traditional Enterprise	CNEX Operating Model
Recruit & Hire	8–20 weeks	Pre-qualified engineering ecosystem
Operational Readiness	4–8 weeks	~2-week SWARM™ Pod objective
Parallel Expansion	Limited	Supported across multiple sites

Activity	Traditional Enterprise	CNEX Operating Model
Global Coverage	Regional	Follow-the-Sun

Business Benefits

- Engineering capacity scales with fleet size without proportional headcount growth, improving unit economics at each expansion tier.
- Standardized playbooks and automation reduce operational variance across sites, supporting consistent SLA performance.
- Follow-the-Sun coverage reduces dependency on any single geography or shift for incident response and engineering throughput.

Global AI Operations & SWARM™

SWARM™ — Scalable Workload Assurance, Resilience & Management

SWARM™ is CNEX's operational framework for deployment, production operations, and incident response — not a support team, but the structure through which engineering capacity is organized and activated. Rather than relying on sequential vendor escalation, SWARM™ enables parallel collaboration across CNEX, strategic OEM partners, customer engineering teams, and the NVIDIA ecosystem under a single, unified Incident Command structure.

CAMBRIDGE NEXUS

Parallel activation replaces sequential escalation - all workstreams engage simultaneously under unified Incident Command.

CAMBRIDGE NEXUS

SWARM™ Components

Component	Primary Responsibilities
CNEX AI Operations	Incident command, monitoring, SLA governance
CNEX SRE	Reliability, automation, production stability
ProphetStor	Follow-the-Sun engineering, Cortex™, AboveCloud™, RCA
ASUS ISBG	AI Factory architecture, NVIDIA reference architecture guidance, networking, liquid cooling, infrastructure integration — through strategic engineering collaboration providing access to 250+ infrastructure engineering professionals
ASRock Rack	Platform engineering, firmware, diagnostics, system integration
Customer Engineering	Workload validation and production coordination
NVIDIA Ecosystem	Advanced platform collaboration when appropriate

Strategic-collaboration relationships (ASUS ISBG, ProphetStor) are described as engineering partnerships; they do not constitute dedicated or exclusive contractual headcount unless specified in a customer-specific agreement.

CAMBRIDGENEXUS

Operational Readiness & Enterprise Differentiation

Disaster Recovery & Business Continuity (DRBC)

Enterprise Service Objectives*

Metric	Operational Design Objective
Operations Coverage	24×7×365
MTTD	< 5 minutes
MTTA	< 10 minutes

Metric	Operational Design Objective
Priority 1 Response	< 15 minutes
SWARM™ Activation	< 15 minutes
Executive Notification	< 30 minutes
Engineering Escalation	< 30 minutes
Initial RCA	< 24 hours
Formal RCA	≤ 5 business days
Engineering Pod Activation	~ 2 weeks
Infrastructure-as-Code Coverage	100%
Automated Monitoring Coverage	> 95%
Change Success Objective	> 98%
Platform Availability Objective	99.95%

**Illustrative operational design objectives. Final commitments are defined in customer-specific SLAs.*

Operating Model Comparison

The comparison below reflects differences in operating model design — not a claim about any specific competitor's capabilities.

Design Area	Conventional Model	CNEX Operating Model
Engineering Expansion	Sequential hiring	SWARM™ Pods + engineering ecosystem
Incident Management	Sequential escalation	Parallel engineering collaboration
Operations Coverage	Regional	Global Follow-the-Sun
AI Operations	Limited automation	Cortex™ + AboveCloud™
OEM Engagement	Warranty / RMA focused	Collaborative engineering model
Customer Integration	Service provider	Extension of customer engineering
Operational Governance	Site-based	Unified AI Operations Command Center
Business Continuity	Infrastructure focused	Multi-layer DRBC + Incident Command

Executive Closing

CNEX's operating model is built around a single premise: enterprise AI infrastructure is an operations problem as much as a procurement problem. Fleet-scale automation through Cortex™ and AboveCloud™, parallel engineering activation through SWARM™, a layered DRBC framework, and a Follow-the-Sun delivery model are designed to work together as one operating system — so that engineering capacity, incident response, and business continuity scale consistently as AI Factory footprints grow.