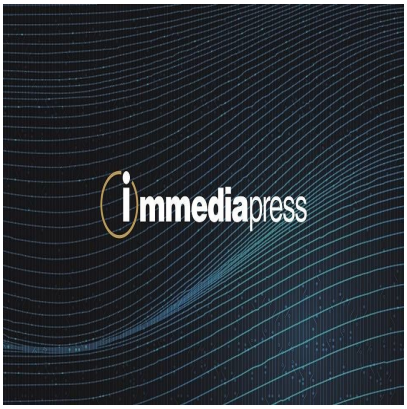

Frost & Sullivan White Paper Names Phancy Rise vGPU a Tier 1 Leading Platform

Descrizione

default watermark

COMUNICATO STAMPA â?? CONTENUTO PROMOZIONALE

HONG KONG SAR â??
Media OutReach Newswire
â?? 17 June 2026 â?? Frost & Sullivan, a globally renowned growth consulting firm, has released its â??2026 AI Infrastructure Orchestration Platform White Paperâ?•. The report recognizes Phancy Groupâ??s Rise vGPU as a Tier 1 Leading Platform, the highest maturity tier in heterogeneous GPU orchestration. Phancyâ??s ModelHub also achieved the highest Overall Score in the enterprise-grade model management platform evaluation. This marks a significant endorsement of Phancyâ??s technological capability in heterogeneous AI infrastructure.

According to the white paper, as large model applications scale rapidly, Chinaâ??s AI industry is facing structural challenges stemming from multi-chip coexistence. These include hardware heterogeneity, fragmented software stacks, persistently low GPU utilization (generally below 30%), and rising model adaptation complexity â?? all of which have become major bottlenecks for enterprise-scale AI deployment.

The report highlights a fundamental shift in AI infrastructure competitiveness â?? moving away from â??single-chip performanceâ?• toward â??cluster-scale system coordination.â?• At this critical juncture, Phancy has positioned itself as a leader in advanced orchestration through its full-stack AI infrastructure platform, offering a proven solution to heterogeneous compute challenges and helping drive Chinaâ??s AI industry from â??compute accumulationâ?• into a new era of â??compute orchestration.â?•

Phancy Rise vGPU: Tier 1 Leading Platform

In its assessment of mainstream AI infrastructure platforms, Frost & Sullivan defined Tier 1 criteria across three core dimensions: heterogeneous support, fine-grained control, and production-grade execution. Phancy Rise vGPU meets all three standards and has been recognized as a Tier 1 Leading Platform.

Rise vGPU transforms AI infrastructure from fragmented, low-efficiency device-level management to a unified software-defined control plane. Its key technology breakthroughs include:

â?? Comprehensive Heterogeneous Management: Unified onboarding and management across more than 10 mainstream GPU/NPU vendors, including NVIDIA, Ascend, Cambricon, Hygon, and others.

default watermark

â?? Ultra-Fine Resource Partitioning: Industry-leading sub-GPU level compute and MB-level memory granularity slicing.

â?? Significant Utilization Improvement: Through safe oversubscription and time/space multiplexing, GPU utilization is increased from industry averages below 30% to 70%-90%.

â?? Intelligent Precision Scheduling: Multi-dimensional scheduling algorithms based on priority, topology, load, and resource awareness to achieve optimal compute allocation.

â?? Production-Grade SLA Assurance: The Deterministic Execution Layer delivers committed and auditable SLA guarantees for critical inference workloads.

â?? Full Lifecycle Operability: Comprehensive monitoring, metering, and cost allocation capabilities that turn GPU resources into truly operable digital assets.

Model Hub: Highest Overall Score in Model Management Platform Evaluation

Beyond compute orchestration, the report underscores the strategic importance of enterprise-grade model management platforms. As a powerful complement to Rise vGPU, Phancy ModelHub enables enterprises to build a complete full-stack AI infrastructure â?? from compute to models and from resource scheduling to business delivery.

The white paper notes that Phancy ModelHub delivers leading performance in key areas such as Model & Chip Compatibility, Execution Stability & Performance, and Model-GPU Coordination & Scheduling, achieving the highest Overall Score. Through its unified model management and execution platform,

ModelHub creates a seamless closed-loop process covering model onboarding, deployment optimization, inference services, and version governance â?? significantly lowering the barrier to model deployment and accelerating AI innovation.

Dr. Dai Wenyuan, Founder & CEO of Phancy, said: â??The Frost & Sullivan white paper accurately captures the inflection point in AI infrastructure development. The recognition of Rise vGPU as a Tier 1 Leading Platform and ModelHubâ??s top Overall Score provide important authoritative validation of Phancyâ??s technology strategy and product strength. As a full-stack AI cloud service platform, Phancy believes the next wave of competitiveness in the AI industry will come from systematic improvements in compute orchestration efficiency. We will continue to focus on heterogeneous compute unified scheduling and model ecosystem operations, working closely with customers and industry partners to advance Chinaâ??s AI industry from â??compute accumulationâ?? to a true â??compute orchestrationâ?? era.â?•

default watermark

About Phancy Group

Phancy Group (6682.HK) is a leading full-stack AI cloud services platform, providing comprehensive solutions for the AI 2.0 era. Our offerings include Rise vGPU, ModelHub and SageAIOS, delivering efficient and scalable AI infrastructure with end-to-end capabilities. We provide a complete solution from heterogeneous compute resource management and optimization to the deployment of intelligent agent models. These solutions empower digital transformation across a wide range of industries, supporting our vision of building a large-scale and efficient â??Token Factory.â?•

Guided by the mission of â??AI for Everyoneâ?• and positioned as the â??Navigator of AI,â?• Phancy Group is committed to becoming a global leader in Artificial General Intelligence.

Contatti:
Immediapress
comunicati@immediapress.it

Media Contact:Suki LeungPR Directorsukileung@4paradigm.com

COMUNICATO STAMPA â?? CONTENUTO PROMOZIONALE

ResponsabilitÃ editoriale di Immediapress

â??

immediapress

Categoria

1. Comunicati

Tag

1. ImmediaPress

Data di creazione
Giugno 17, 2026
Autore
redazione

default watermark