NEWS RELEASE

Penguin Solutions Releases ICE ClusterWare Management Software 13.0 for Optimizing AI Infrastructure

2025-11-17

New cluster management software capabilities deliver sustained peak performance and network-isolated resource segmentation for AI and HPC applications

FREMONT, Calif.--(BUSINESS WIRE)-- **Penguin Solutions, Inc**. ("Penguin Solutions") (Nasdaq: **PENG**), a leading provider of high-performance computing and AI infrastructure solutions, today announced the release of **ICE ClusterWare™ software 13.0**. This latest version introduces powerful new capabilities that solve two critical challenges in production-scale AI and HPC: sustaining peak cluster performance and secure provisioning of a single cluster to diverse user groups. These new features enable organizations to maximize return on their AI infrastructure investments by safely sharing resources across more users while ensuring consistent, reliable performance.

Penguin Solutions releases ICE ClusterWare management software 13.0 with powerful new capabilities that solve two critical challenges in production-scale AI and HPC: sustaining peak cluster performance and secure provisioning of a single cluster to diverse user groups.

When an organization's AI deployments progress from isolated pilot projects to enterprise-wide production environments, operational demands on infrastructure intensify immediately. Penguin's ICE ClusterWare 13.0 addresses this with built-in anomaly detection and auto-remediation, along with network-isolated multi-tenancy—delivering the operational excellence required to support AI as a core business function.

"With the launch of our ICE ClusterWare software 13.0, we're delivering pivotal advancements to help organizations manage the growing complexity of modern AI and HPC environments," said Sharri Parsell, vice president software engineering for Penguin Solutions. "As AI continues to evolve from experimental pilots to enterprise-scale

deployments, organizations need robust, intelligent infrastructure that drives operational excellence and enables AI success across the enterprise.”

The patent-pending anomaly detection and auto-remediation technology ensures peak cluster performance and resource availability, continuously monitoring for hidden performance degradation that traditional diagnostic tools miss. Upon detection, the system automatically isolates underperforming nodes and initiates remediation in real time, ensuring that workloads are scheduled on validated, high performing nodes. This proactive approach reduces administrative burdens, prevents unplanned downtime, and maximizes the cluster’s usable capacity. As a result, this new capability significantly shortens model training by reducing restarts and loss of work.

The new optional network-isolated multi-tenancy feature enables organizations to securely and efficiently share high-value GPU clusters, creating dedicated subclusters to support different departments, projects, or GPU-as-a-Service (GPUaaS) customers. This capability provides isolated environments, giving tenants the autonomy to select their own workload manager, govern users, and run workloads with confidence that data and operations remain segregated and secure.

"The pace and quality of biomedical research are directly tied to the technology that supports it," said Assistant Dean for Information Technology Shailesh Shenoy, Albert Einstein College of Medicine. "AI and HPC are crucial to providing the computational power that biometrics, life science, and medical research require, but we also had to ensure that it is optimized for our specific use cases. Having a trusted partner in Penguin Solutions has enabled us to not only build out this infrastructure, but also helped ensure we can manage and optimize it to keep it running smoothly and at capacity, freeing our faculty and student researchers to continue their groundbreaking work without interruption."

Reducing the security and resource utilization conflicts that previously forced organizations to build separate clusters drastically improves time to value. This capability is essential for cloud service providers and hyperscalers providing GPUaaS, enterprises and research institutes delivering AI computing to internal business groups, and federal or government agencies that require the highest level of security and resource isolation.

General availability for ICE ClusterWare software 13.0 is scheduled for December 2, 2025. To learn more about ICE ClusterWare 13.0, visit **https://www.penguinsolutions.com/en-us/contact-us** and register for our upcoming webinar: **Navigating the AI Journey from Pilot to Production**, on December 17, 2025 at 10:00 a.m. PST.

About Penguin Solutions

The most exciting technological advancements are also the most challenging for companies to adopt. At Penguin Solutions, we support our customers in achieving their ambitions across our computing, memory, and LED lines of business. With our expert skills, experience, and partnerships, we turn our customers' most complex challenges into compelling opportunities.

For more information, visit **https://www.penguinsolutions.com**.

Media Contact

Maureen O'Leary, Director Corporate Communications
Penguin Solutions
602-330-6846
pr@penguinsolutions.com

Source: SMART Global Holdings, Inc.