

VESSL AI Unveils Blueprint for 300MW AI Factory at '2026 OCP Korea Tech Day'

SEOUL, SOUTH KOREA, August 21, 2026 /EINPresswire.com/ -- CEO Jaeman An announced today that VESSL [AI](#) will participate in the '2026 OCP Korea Tech Day' held at COEX in Seoul on August 21st.

The company will showcase its operational technology for 'VESSL [Cloud](#)' — a multi-datacenter-based [GPU](#) cloud service — along with its datacenter expansion strategy. Leveraging these foundations, VESSL AI plans to accelerate its 'AI Factory

Service,' offering an end-to-end framework that integrates computing resources, development and operational environments, and infrastructure management, all tailored to client-specific AI workloads.

To address the growing demand for compute, VESSL AI also outline a concrete capacity roadmap for its AI datacenters. The company aims to bring a total of 100 MW (megawatts) of AI datacenter capacity online by next year (50 MW domestically and 50 MW internationally), with plans to expand to 300 MW by 2028. This multi-datacenter model, spanning multiple Korean regions and overseas locations, is designed to flexibly accommodate distributed AI training and inference workloads.

In addition, VESSL AI will share its procurement and deployment plans for new GPUs, including NVIDIA's next-generation architecture, while offering pre-reservation consultations for enterprise and institutional clients to secure next-gen GPU allocations.

VESSL AI leadership will also participate in two presentation sessions during the event:

□ Jaeman An (CEO) will present in the Datacenter Track on "From Datacenters to AI Factories." He will highlight shifting token demand driven by the evolution of AI agents and compare GPU clouds to "oil refineries" — explaining how physical inputs (power, silicon, and capital) combined with software operational efficiency dictate the ultimate yield per token.

□ Jihwan Chun (CTO) will present in the AI Solution Track on "GPU Infrastructure Reliability



Engineering in the Era of Agentic AI." He will discuss the infrastructure safeguards that VESSL AI has developed while operating H100, H200, and B300 clusters across multiple Cloud Service Providers (CSPs) in Korea, including GPU monitoring, and a multi-cluster control plane with an automated recovery system optimized to meet inference SLAs (service level agreements).

"With token throughput skyrocketing due to the spread of AI agents, infrastructure competitiveness is shifting from 'how many GPUs you possess' to 'how many tokens you can extract from those GPUs,'" said CEO Jaeman An. "Through this event, we look forward to sharing the technical expertise VESSL AI has accumulated through multi-datacenter operations, along with our capacity roadmap to support it."

VESSL AI

PR

+82 10-9274-5094

[email us here](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/935619605>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2026 Newsmatics Inc. All Right Reserved.