Press release

# Oracle Unveils Next-Generation Oracle Cloud Infrastructure Zettascale10 Cluster for AI

*Largest AI supercomputer in the cloud delivers 10X the amount of zettaFLOPS of peak performance*

*Built on Oracle Acceleron RoCE networking architecture with NVIDIA AI infrastructure, OCI Zettascale10 will provide multi-gigawatt AI workload capacity and scale*

**Oracle AI World, Las Vegas—Oct 14, 2025 –** Oracle today announced Oracle Cloud Infrastructure (OCI) Zettascale10, the largest AI supercomputer in the cloud. OCI Zettascale10 connects hundreds of thousands of NVIDIA GPUs across multiple data centers to form multi-gigawatt clusters that deliver up to an unprecedented 16 zettaFLOPS of peak performance. OCI Zettascale10 is the fabric underpinning the flagship supercluster built in collaboration with OpenAI in Abilene, Texas, as part of Stargate. Built on next-generation Oracle Acceleron RoCE networking architecture, OCI Zettascale10 is powered by NVIDIA AI infrastructure that delivers breakthrough scale, extremely low GPU-GPU latency across the cluster, industry-leading price-performance, improved cluster utilization, and the reliability required for large scale AI workloads.

OCI Zettascale10 is a powerful evolution of the first Zettascale cloud computing cluster, which was introduced in September 2024. OCI Zettascale10 clusters are housed in large gigawatt data center campuses that are hyper-optimized for density within a two-kilometer radius to offer the best GPU-GPU latency for large scale AI training workloads. This architecture is being deployed with OpenAI at the Stargate site in Abilene.

“With OCI Zettascale10, we’re fusing OCI’s groundbreaking Oracle Acceleron RoCE network architecture with next-generation NVIDIA AI infrastructure to deliver multi-gigawatt AI capacity at unmatched scale,” said Mahesh Thiagarajan, executive vice president, Oracle Cloud Infrastructure. “Customers can build, train, and deploy their largest AI models into production using less power per unit of performance and achieving high reliability. In addition, customers will have the freedom to operate across Oracle’s distributed cloud with strong data and AI sovereignty controls.”

“OCI Zettascale10 network and cluster fabric was developed and deployed first at the flagship Stargate site in Abilene, Texas – our joint supercluster with Oracle,” said Peter Hoeschele, vice president, Infrastructure and Industrial Compute, OpenAI. “The highly scalable custom RoCE design maximizes fabric-wide performance at gigawatt scale while keeping most of the power focused on compute. We’re excited to keep scaling Abilene and the broader Stargate program together.”

OCI plans to offer multi-gigawatt deployments of OCI Zettascale10 to customers. Initially, OCI Zettascale10 clusters will target deployments of up to 800,000 NVIDIA GPUs delivering predictable performance and strong cost efficiency, with high GPU-to-GPU bandwidth enabled by Oracle Acceleron’s ultra-low-latency RoCEv2 networking.

“Oracle and NVIDIA are bringing together OCI’s distributed cloud and our full-stack AI infrastructure to deliver AI at extraordinary scale,” said Ian Buck, vice president of Hyperscale, NVIDIA. “Featuring NVIDIA full-stack AI infrastructure, OCI Zettascale10 provides the compute fabric needed to advance state-of-the-art AI research and help organizations everywhere move from experimentation to industrialized AI.”

**Oracle Acceleron RoCE networking delivers scale, reliability, and efficiency for AI on OCI Zettascale10**

Oracle Acceleron RoCE networking architecture is a critical innovation for customers to build, train, and inference AI workloads in the cloud, while taking full advantage of OCI Zettascale10’s power and capabilities. It uses the switching capability built into modern GPU NICs (network interface cards), allowing them to connect to multiple switches simultaneously, with each on a separate and isolated network plane. This approach dramatically increases the network’s overall scale and reliability by shifting traffic to other network planes when one has a problem, avoiding costly stalls and restarts. Key features of Oracle Acceleron RoCE networking that help customers with their critical AI workloads, include:

- **Wide, shallow, resilient fabric:** Helps customers deploy larger AI clusters faster at lower total cost by using the GPU NIC as a mini-switch and connecting to multiple physically and logically isolated planes. This boosts scale while reducing network tiers, cost, and power.
- **Higher reliability:** Helps customers maintain the stability of AI jobs by eliminating data sharing across planes. This shifts traffic away from unstable or congested planes, which keeps training jobs running and avoids costly checkpoint restarts.
- **Consistent performance:** Provides customers with more uniform GPU-to-GPU latency by removing a tier versus traditional three-tier designs, improving predictability for large-scale AI training and inference.
- **Power-efficient optics:** Supports customer workloads with Linear Pluggable Optics (LPO) and Linear Receiver Optics (LRO) to cut network and cooling costs without sacrificing 400G/800G throughput. This allows customers to devote more of their power budget to compute.
- **Operational flexibility:** Helps customers reduce downtime and speed up feature rollouts through plane-level maintenance and independent network operating system updates.

OCI is now taking orders for OCI Zettascale10, which will be available in the second half of next calendar year, with up to 800,000 NVIDIA AI infrastructure GPU platforms.

###

### Additional Resources

- Watch the Oracle AI World keynote with Mahesh Thiagarajan
- Learn more about OCI AI infrastructure

### Contact Info

**Chris Kanaracus**

Oracle PR

chris.kanaracus@oracle.com

+1.207.256.0556

### About Oracle

Oracle offers integrated suites of applications plus secure, autonomous infrastructure in the Oracle Cloud. For more information about Oracle (NYSE: ORCL), please visit us at www.oracle.com.

### About Oracle AI World

Oracle AI World is where customers and partners discover the latest product and technology innovations, see how AI is being applied across industries, and connect with experts and peers. Attendees will gain practical tips and insights to drive immediate impact within their organizations and explore how Oracle is helping unlock the full potential of cloud and AI. Join the event to see new capabilities in action and hear from thought leaders and industry movers. Register now at oracle.com/ai-world or follow the news and conversation at oracle.com/news and linkedin.com/company/oracle.

### Future Product Disclaimer

The preceding is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.

### Forward-Looking Statements Disclaimer

Statements in this article relating to Oracle's future plans, expectations, beliefs, and intentions are "forward-looking statements" and are subject to material risks and uncertainties. Many factors could affect Oracle's current expectations and actual results, and could cause actual results to differ materially. A discussion of such factors and other risks that affect Oracle's business is contained in Oracle's Securities and Exchange Commission (SEC) filings, including Oracle's most recent reports on Form 10-K and Form 10-Q under the heading "Risk Factors." These filings are available on the SEC's website or on Oracle's website at oracle.com/investor. All information in this article is

current as of October 14, 2025 and Oracle undertakes no duty to update any statement in light of new information or future events.

**Trademarks**