

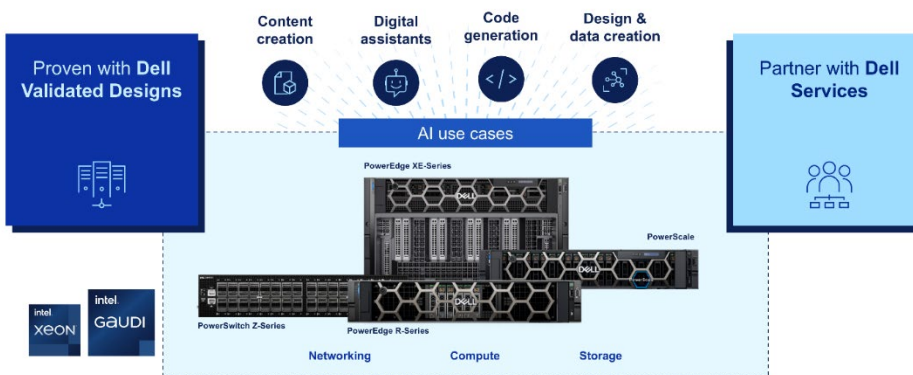
Dell Generative AI Solutions with Intel

In the realm of artificial intelligence and advanced computing, staying competitive requires robust, reliable, and high-performance solutions. To meet these needs, Dell Technologies introduced the Dell AI Factory - an approach to help accelerate AI innovation in organizations of all sizes with portfolio of products, solutions and services optimized for AI workloads. It brings together these AI-optimized technologies with an open ecosystem of partners, validated and integrated solutions, expert services, and best practices to help customers achieve AI outcomes faster.

The Dell AI Factory will soon include new Dell Generative AI Solutions with Intel. Powered by the Intel® Gaudi 3® AI accelerator, these innovative solutions deliver high performance at scale and the flexibility to choose the software, ethernet and implementation model that best meet your needs.

Joint solutions for GenAI workflows

Dell Generative AI Solutions with Intel combine the power of Intel Gaudi 3 AI accelerators with Dell servers, storage, networking, professional services, and open-source software to provide jointly engineered, tested and validated solutions that ensure a seamless deployment experience, all on your terms. Proven with Dell Validated Designs, these solutions can help you right-size your GenAI investment to support your targeted AI use cases and accelerate business outcomes. Learn more about how the Dell PowerEdge XE9680 and Intel Gaudi 3 accelerators come together to power GenAI workflows [here](#).



Professional services built for GenAI

Dell has an extensive portfolio of Professional Services to help accelerate business outcomes across every stage of your GenAI journey. Our expert services team can help develop, implement, and scale your AI solution, from strategy development to data preparation, platform deployment, and use case aligned model tuning, to operating at scale to ensure your organization is set up for success into the future.

Key benefits



Performance

High performance with a reduced TCO



Scale

Scale out architectures and ease of use combine to offer a simple path to growth

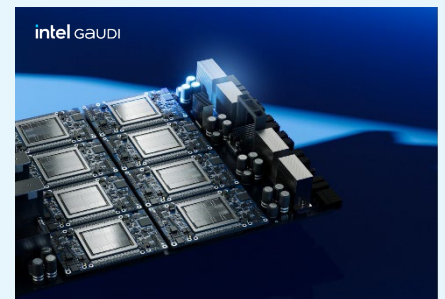


Flexibility

Expands choice of silicon, software, ethernet and implementation

Intel Gaudi 3 accelerator fast facts

- Universal Baseboard (HLB-325)
- 14.6 PFLOPS (FP8)
- >1TB HBM2e
- 29.6TB/S HBM Bandwidth
- 64 Matrix Multiplication engines
- 192 200GbE RDMA NICs
- 9.6TB/s bi-directional networking



Learn more about

[Intel Gaudi 3 accelerators](#)

Intel® Gaudi 3® AI accelerator

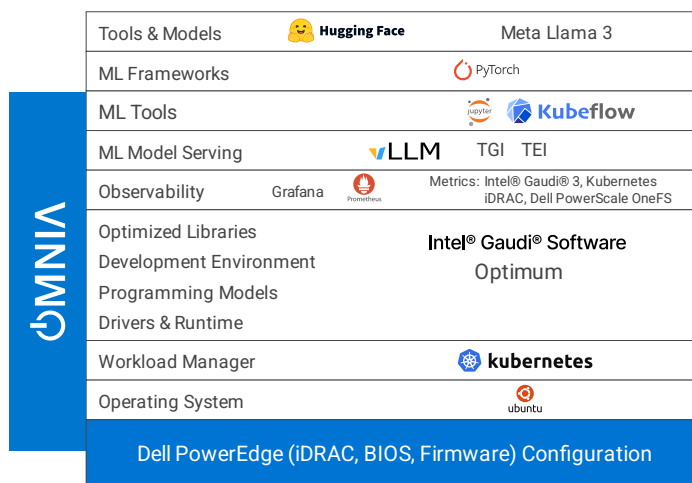
Intel Gaudi 3 accelerators offer transformative benefits for enterprises deploying generative AI solutions, combining exceptional performance, scalability, and ease of use, all with improved ROI and efficiency compared to competing solutions. Its advanced architecture supports demanding AI workloads, providing superior price-performance compared to competitors. The open standard Ethernet networking ensures seamless integration and flexibility with industry-standard networking solutions.

Integration with key AI frameworks and platforms like Pytorch and Hugging Face enhances developer productivity and simplifies model migrations, making it easier to harness a broad spectrum of AI applications. With Intel Gaudi 3 accelerators, organizations can achieve faster insights and maintain a competitive edge in the rapidly evolving AI landscape.

Fully validated open-source ecosystem

Dell GenAI Solutions with Intel feature a full stack of open-source software that has been tested, validated and benchmarked with Dell hardware and Intel Gaudi 3 accelerators. This ensures organizations can use the tools they need and avoid vendor lock-in without the guesswork.

Compute Node Software Stack



Ready to get started?

Sign up to evaluate Intel Gaudi accelerators on [Intel Developer Cloud](#)

Start testing your target use cases with real-world AI functions on [Intel AI Explorer](#) (JavaScript will need to be enabled to run this app)

About Dell Generative AI

Dell Technologies offers the world's broadest Generative AI solutions portfolio, from desktop to data center to cloud, all in one place.

You can accelerate your AI journey with Dell Technologies Services and a broad partner ecosystem that includes support for many open-source Generative AI models. Regardless of whether a team member is working with purpose-built Dell infrastructure to power organization-wide Generative AI applications, or simply prompting AI from a Dell laptop, they are working with the industry's broadest and most advanced portfolio of AI-enabled solutions.

Our approach is to bring AI to your data. This means meeting strict data sovereignty and compliance requirements while right-sizing AI infrastructure to reduce costs, support sustainability objectives, and keep you in control of the model and your data.

Next steps

Get started with a fee-waived Accelerator Workshop for Generative AI, which can help you gain consensus among your business and technical stakeholders on your solution and prioritized use cases.

Learn more

[Accelerator Workshop for Generative AI](#)



Learn more about Dell solutions for [Generative AI](#).



Contact a Dell Technologies Expert.



Join the conversation.