

ACCELERATE YOUR CLOUDERA DATA PLATFORM WORKLOADS WITH NVIDIA-CERTIFIED SYSTEMS



Easily deploy end-to-end data science pipelines on Cloudera Data Platform running on NVIDIA accelerated infrastructure to improve your data-driven operations.

CLOUDERA DATA PLATFORM

- Accelerate your data science toolchain with minimal code changes and deliver hassle-free scaling on NVIDIA-Certified systems
- Increase data science productivity with visibility, security, and governance of the complete data science lifecycle

PROOF POINTS

- GPU-computing runs Apache Spark jobs in CDP 10x faster than legacy Hadoop CPU-based infrastructure and 5x faster over modern CPU-based servers
- Including both the cost of accelerated hardware and Cloudera's software premium for GPUs, achieve better than 3x cost-relative benefit

NVIDIA-CERTIFIED SYSTEMS

- Cost-effective NVIDIA infrastructure empowers IT teams to an accelerated CDP solution for intuitive, self-service ML—now and into the future.
- Accelerated CDP Starter Solutions available to enable enterprises to confidently deploy scalable hardware and software solutions that securely and optimally run accelerated workloads.

Learn more at [nvidia.com](https://www.nvidia.com)

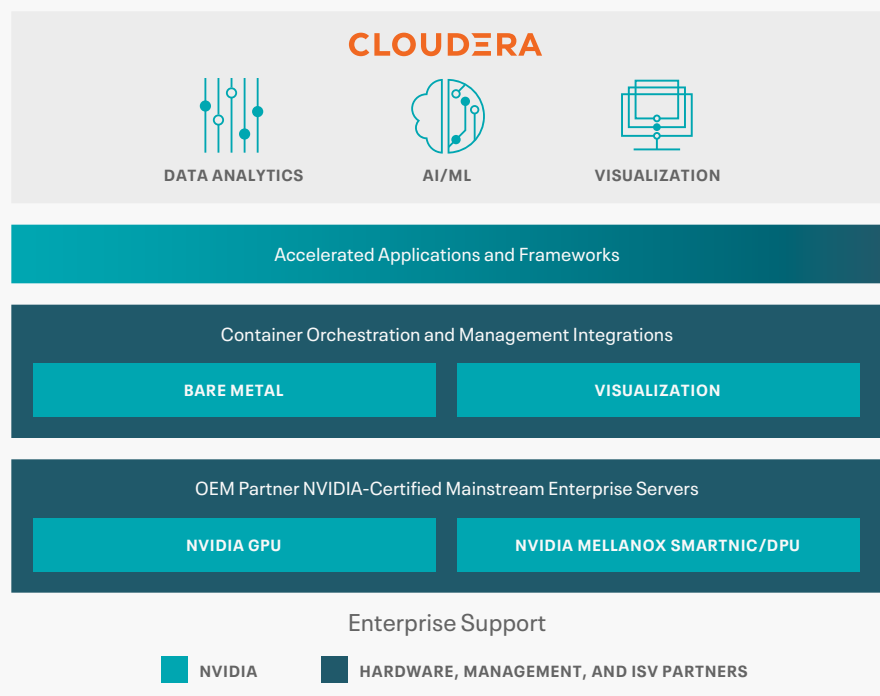
Addressing the Challenges of Modern Data-Driven Enterprises

Data science and machine learning has become the world's largest compute segment. Modest improvements in the accuracy of predictive machine learning models can translate into billions to the bottom line. The training of predictive models is at the core of data science. In fact, the majority of IT budgets for data science are spent on building machine learning models. To build the best models, data scientists need to train, evaluate, and retrain with many iterations. Today, these iterations take days, limiting how many can occur before deploying to production and impacting the quality of the final result.

These operations are bottlenecked by the serial nature of CPU-only computing, which is compounded when scaling out for large processes. With NVIDIA's Enterprise Computing platform (EGX), enterprises can easily leverage parallel GPU computing to remove bottlenecks and quickly improve performance. This increased performance significantly improves time to insight and the return on investment for data-driven enterprises.

Accelerate Data Science and Machine Learning Workflows

Running data science workloads on Cloudera Data Platform accelerated on NVIDIA computing greatly improves time to value by enabling data scientists to collaborate in a single unified platform that is all inclusive for powering any AI use case. In addition, accelerated Apache Spark 3.0 workloads now run seamlessly on CDP, empowering enterprises to data science their needs—from research to production—with a secure, scalable, and open platform. Data science teams can leverage purpose-built tooling for agile experimentation, data analytics and machine learning 10x faster and at lower cost with GPU parallelization and acceleration.



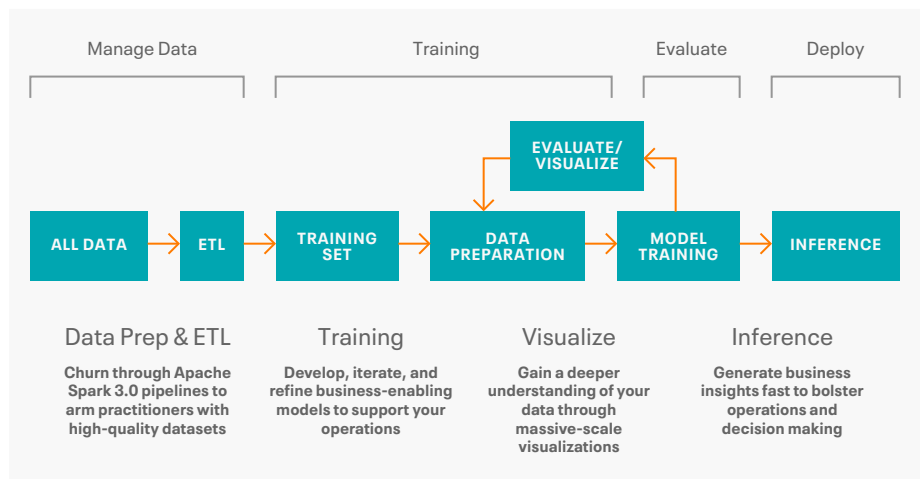
About Cloudera

At Cloudera, we believe that data can make what is impossible today, possible tomorrow. We empower people to transform complex data into clear and actionable insights. Cloudera delivers an enterprise data cloud for any data, anywhere, from the Edge to AI. Powered by the relentless innovation of the open source community, Cloudera advances digital transformation for the world's largest enterprises.

Learn more at cloudera.com


Delivering Lightning-Fast Performance on Big Data

The NVIDIA EGX platform accelerates the end-to-end data science life cycle, whether your organization needs to reduce processing time of ETL pipelines or accelerate terabyte-scale machine learning workflows. With RAPIDS, data practitioners can accelerate pipelines on NVIDIA GPUs, reducing data operations like data loading, processing, and training from days to minutes. RAPIDS abstracts the complexities of accelerated data science by building upon popular Python and Java libraries, enabling users to see benefits immediately.




Deploy High-Performance CDP in Your Data Center Today

The NVIDIA-Certified Systems on CDP dramatically boosts the performance of your end-to-end analytics workflows, speeding up the time to value while reducing overall cost. NVIDIA-Certified servers are available from leading OEM server vendors and come preconfigured, tested, and validated to work for a variety of AI-related workloads—including Apache Spark. Ready to start your AI journey? Accelerated CDP Starter Solutions are the easiest path to deploy GPU-accelerated CDP. Complete reference architectures from Cloudera, NVIDIA, and our OEM partners are in development and will soon be available.




NVIDIA-Certified Systems

NVIDIA-Certified Systems create the essential platform for the modernization of enterprise data centers



Accelerated CDP Starter Solutions

Deployment made simple with Accelerated CDP Starter Solutions made available from any leading server OEM



Reference Architectures

Effortlessly scale from a single server to multi-node, multi-GPU clusters with a consistent, intuitive architecture