Solution Overview

Transforming Data Science With Accelerated Computing

NVIDIA RTX-powered AI workstations for data science workflows.

Challenges

Every industry is being transformed with AI to improve competitiveness and operational efficiency. In retail, chatbots are simplifying product inventory search for users. In the financial services industry, AI-based tools are enhancing fraud detection. In healthcare, AI is helping uncover new insights from genomics that are leading to better patient outcomes.

For AI to achieve its full potential in domain-specific use cases, it requires huge volumes of data that fit the end purpose to be trained on. To get that, data scientists preprocess datasets by cleaning the data, managing missing values, converting data types, and engineering specific features needed for the model. Data preprocessing is especially important in domain-specific areas, as specific and often proprietary data informs how intelligent the AI becomes in their area.

The good news is that we have more data than ever, with data volume expected to grow at a CAGR of 24.4% from 2023–2028.[1] The problem arising is that the need for efficient processing tools has never been more critical.

The explosive growth and increasingly complex demands for data are overwhelming traditional CPU-based data processing tools. Data scientists often work with massive datasets, organized as tables with millions of rows and columns. As dataset sizes grow, traditional data processing tools slow down significantly, reducing productivity and extending time-to-market for enterprises.

Accelerate Every Step of Your Data Science Workflow

To meet the demands of preprocessing data for AI, NVIDIA offers a comprehensive full-stack solution that combines accelerated data processing libraries with NVIDIA RTX™-powered AI workstations. RAPIDS™ includes zero-code change accelerators for Python-based libraries like pandas, NetworkX, and XGBoost, enabling developers to code without comprising on speed. **NVIDIA RTX-powered AI workstations** are designed to provide unmatched performance for data science tasks, leveraging the power of parallel processing in NVIDIA RTX GPUs. With RAPIDS on RTX-powered AI workstations, data scientists can use the tools they love in a local environment, making processing bottlenecks a thing of the past.

Key Challenges for Data Science

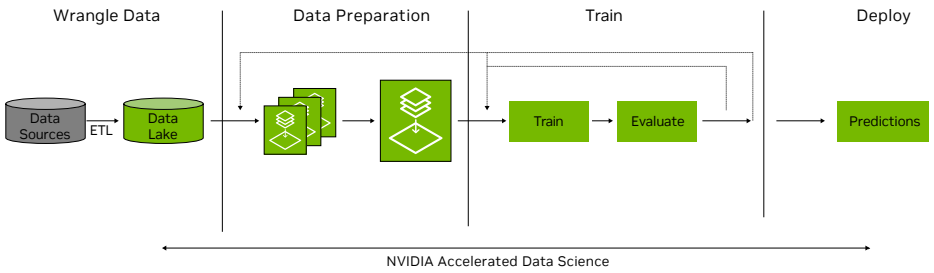
- **Data Preparation:** Data preparation is a complex, time-consuming process that takes most of a data scientists time.
- **Scaling:** Volume of data is growing at a rapid pace. Data scientists may resort to downsampling datasets to make large datasets more manageable, leading to suboptimal results.
- **Hardware:** Demand for accelerated AI hardware for data centers and cloud service providers (CSPs) is exceeding supply. Current desktop computing resources may not be suitable for data science workflows.

Benefits of RTX-Powered AI Workstations for Data Science

- Deliver unparalleled acceleration through physical workstations and RTX Virtual Workstations.
- Large GPU memory and up to four GPUs per system configuration enable data exploration, feature and model evaluation, and visualization.

1. Source: IDC, Worldwide IDC Global DataSphere Forecast, 2024–2028: AI Everywhere, But Upsurge in Data Will Take Time, Doc #US52076424, May 2024

End-to-End Accelerated Analytics with NVIDIA



- **Data Preparation:** Rapidly process massive datasets with zero code changes with NVIDIA RTX GPUs and libraries like RAPIDS cuDF and NetworkX. RAPIDS cuDF accelerates pandas operations by up to 100X performance gains with zero code changes for quick data preparation. This allows you to bypass the limitations of traditional tools and immediately benefit from enhanced processing speeds.
- **Exploratory Data Analysis:** Extract actionable insights with advanced analytics and visualization tools. RAPIDS and PyData libraries allow you to create expansive, detailed visualizations that are interactive and responsive to updates and changes. These visualizations enhance data understanding and support critical decision-making.
- **Model Training:** With GPU-accelerated XGBoost, model training times dramatically reduce from weeks to minutes. This acceleration enables faster iterations, improving operational efficiency and productivity.

High-Performance AI Workstations for Data Science

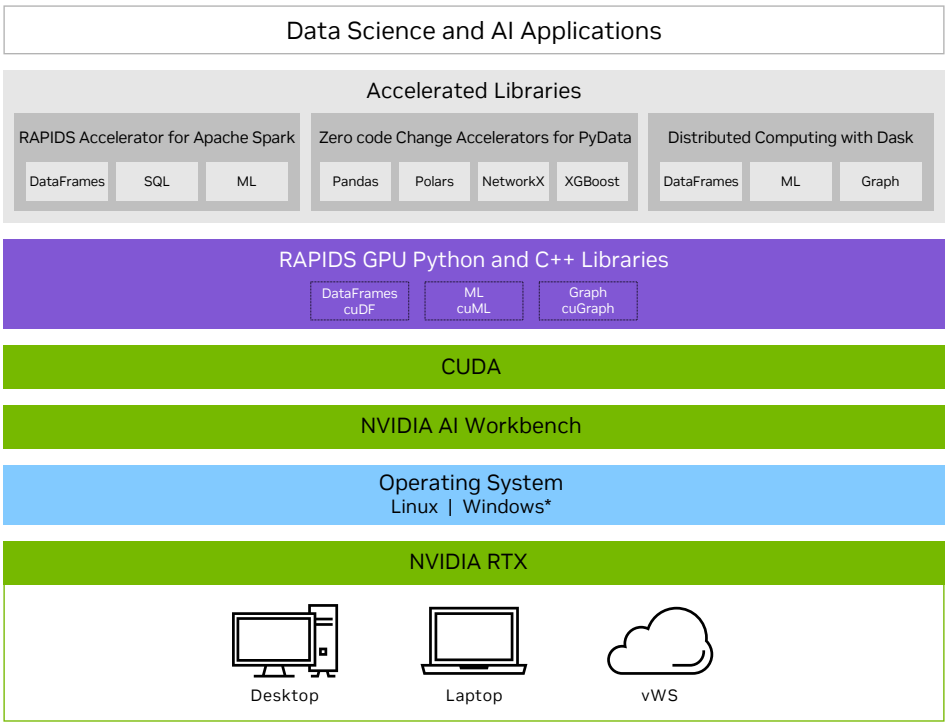
NVIDIA RTX-powered AI workstations are built to drive high-performance data science. By combining the latest NVIDIA RTX GPUs with a robust software stack—including NVIDIA CUDA-X™ and RAPIDS data processing and machine learning libraries—these workstations deliver accelerated workflows for data preparation, model training, and visualization.

AI workstations are optimized for fast, local processing with minimal latency, offering a seamless development experience. With **NVIDIA AI Workbench**, they streamline project setup, development, and collaboration across various environments from desktops to cloud platforms.

- Provides additional AI computing resources to augment data center and cloud instances for development and R&D tasks.
- AI workstations are optimized to leverage NVIDIA CUDA-X libraries, including RAPIDS. RAPIDS is an open-source suite of GPU-accelerated data science and AI libraries with APIs that match the most popular open-source data tools.
- Enterprise-grade hardware maximizes uptime with enterprise-level performance, reliability, and support.

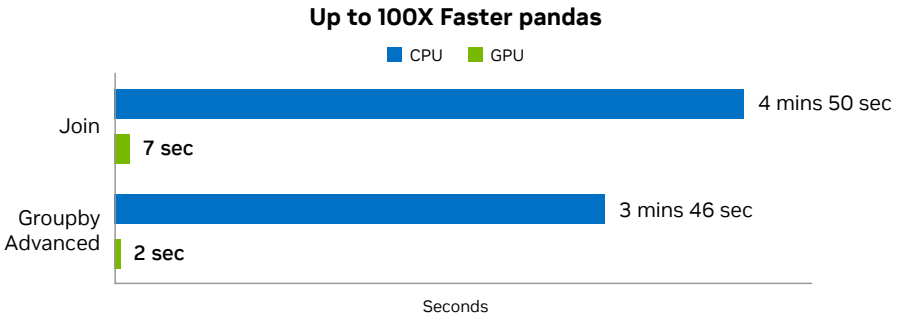
100X Acceleration

- **Zero Code Change Acceleration:** Users can enable GPU acceleration by simply loading the cudf.pandas extension in a Jupyter Notebook or using a command-line flag in a Python script.
- **Unified CPU/GPU Workflows:** Operations are first attempted on the GPU. If the GPU cannot handle the operation, it falls back to the CPU, ensuring compatibility and robustness.
- **Third-Party Library Compatibility:** cudf.pandas is compatible with most third-party libraries that operate on pandas objects, allowing for seamless integration into existing workflows.
- **Scalable Development Environment:** With AI workstations, pandas workloads can be accelerated from laptop to cloud across teams, enabling seamless collaboration.

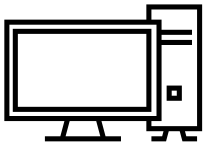


*requires WSL2

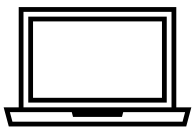
For data scientists using Python, NVIDIA's cuDF provides a zero-code change solution that transforms pandas operations, offering up to 100X faster performance. NVIDIA accelerated end-to-end data analytics toolchain maximizes data scientist productivity compared to CPU-based solutions.



Standard DuckDB Data Benchmark (5GB)
HW: NVIDIA Grace Hopper, CPU: Intel® Xeon® Platinum 8480C
SW: pandas v2.2, RAPIDS cuDF 23.10



Desktop	Good	Better	Best
CPU	w3 or 4th Gen Intel Xeon Silver or AMD Ryzen Threadripper Pro	Intel Xeon w5 Silver or AMD Ryzen Threadripper Pro	Intel Xeon w5 Gold or AMD Ryzen Threadripper Pro
System Memory	128GB ECC DDR5	256GB ECC DDR5	1TB ECC DDR5
Storage	1TB boot + 2TB SSD, NVMe	1TB boot + 2-4TB SSD, NVMe RAID[3]	2TB boot + 2TB SSD, NVMe RAID[3]
NIC	10GbE NIC	NVIDIA ConnectX-6Dx (256GbE)	NVIDIA ConnectX-6Dx (256GbE)
OS	Ubuntu/RHEL/SUSE[1]	Ubuntu/RHEL/SUSE[1]	Ubuntu/RHEL/SUSE[1]
GPU	NVIDIA RTX 6000 Ada Generation or NVIDIA A800 40GB Active[2]	2x NVIDIA RTX 6000 Ada Generation or 2x NVIDIA A800 40GB Active[2]	4x NVIDIA RTX 6000 Ada Generation or 3x NVIDIA A800 40GB Active[2]



Laptop	Best
CPU	Intel i7 or i9
System Memory	32GB DDR5
Storage	1TB NVMe
OS	Ubuntu/RHEL/SUSE[1]
GPU	NVIDIA RTX 5000 Ada Generation Laptop GPU

1. Linux: See NVIDIA AI Enterprise documentation for exact release support.
2. A800: Display out options include A1000 and RTX 4000 Ada Generation see mix-and-match support matrix for more details.
3. RAID: for large datasets that need fast/reliable storage.

Enterprise-Ready Solutions

Available from leading OEM manufacturers, NVIDIA RTX-powered AI workstations are designed and built for demanding enterprise deployments. Powered by the latest generation of workstation CPUs and available with NVIDIA ConnectX® high-performance networking solutions, NVIDIA AI workstations are ready to tackle demanding AI development workflows. The latest generation of OEM desktop, mobile, and virtual workstations are available now.

With a full stack of enterprise-level deployment, support, and optimization tools, NVIDIA AI workstations easily fit into existing IT infrastructure, providing drop-in solutions for AI training, development, and inferencing on the desktop or in the data center. The NVIDIA GPU architecture scales from cloud to data center, desktop, laptop, and embedded device, supporting the same software stack across devices, which enables users to move workloads seamlessly between them.

Ready to Get Started?

To learn more about NVIDIA RTX-powered AI Workstations, visit: **www.nvidia.com/ai-workstations**

Contact sales at: **nvidia.com/en-us/contact/sales**

© 2024 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, ConnectX, NVIDIA CUDA-X, Hopper, RAPIDS, and NVIDIA RTX are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. Other company and product names may be trademarks of the respective owners with which they are associated. 3412930. NOV24

NVIDIA