



# NVIDIA RAPIDS Accelerator for Apache Spark

Faster results and lower cost for data analytics and machine learning in energy.



Energy companies and utilities are looking to advance net-zero emissions goals, modernize power grids, and deliver reliable, lower-cost energy. These enterprises can realize significant ROI and cost savings using data analytics and machine learning. Many enterprises use Apache Spark for key operations such as ingesting raw data into data lakes, improving business process analytics, loading data into data warehouses, and adding data preprocessing at the start of machine learning pipelines. However, growing workloads are being constrained by slow, CPU-based infrastructure—and slow processing costs time, money, and energy, resulting in a larger carbon footprint.

## The Solution: Accelerate Apache Spark

The **NVIDIA RAPIDS™ Accelerator for Apache Spark** takes advantage of NVIDIA GPUs to accelerate Apache Spark workloads without code changes. It operates as a plug-in to popular Apache Spark platforms. The RAPIDS Accelerator speeds up selected Spark operations while allowing other operations to continue running on the CPU. As a result, processing time can be accelerated up to 5X or more, allowing the same work to be completed with 4X less infrastructure cost.

## Who Needs the RAPIDS Accelerator?

As global energy demand increases, Spark is enabling fast and accurate demand forecasting, real-time supply management, and optimal energy generation, transmission, and distribution—including renewables such as solar and wind.

Heavy industrial operators can use Spark for predictive maintenance of turbines, generators, transformers, and other industrial equipment that minimizes downtime and reduces operational cost. Power-constrained data center operators can also accelerate AI and HPC workloads with Spark to reduce power usage, heating and cooling costs, and carbon emissions. Spark lets energy traders analyze large datasets related to markets, pricing, supply-demand dynamics, and risk factors to improve energy trading and hedging strategies.

And at the grid-edge, energy companies can use Spark to monitor and protect substations, power lines, smart meters, and other critical infrastructure for enhanced grid resiliency and lower-cost, reliable power.

## 5X

### Faster Execution Time

- > Move data in and out of data lakes more quickly
- > Take advantage of faster analytics
- > Accelerate AI pipelines

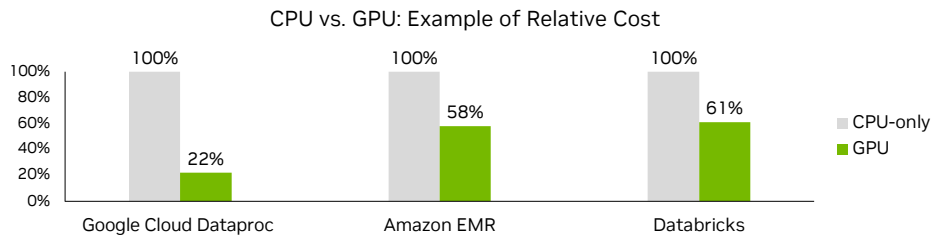
## 4X

### Lower Costs

- > Save on cloud usage costs
- > Reduce power consumption and carbon footprint

## Quickly Analyze Workloads for Cost and Performance

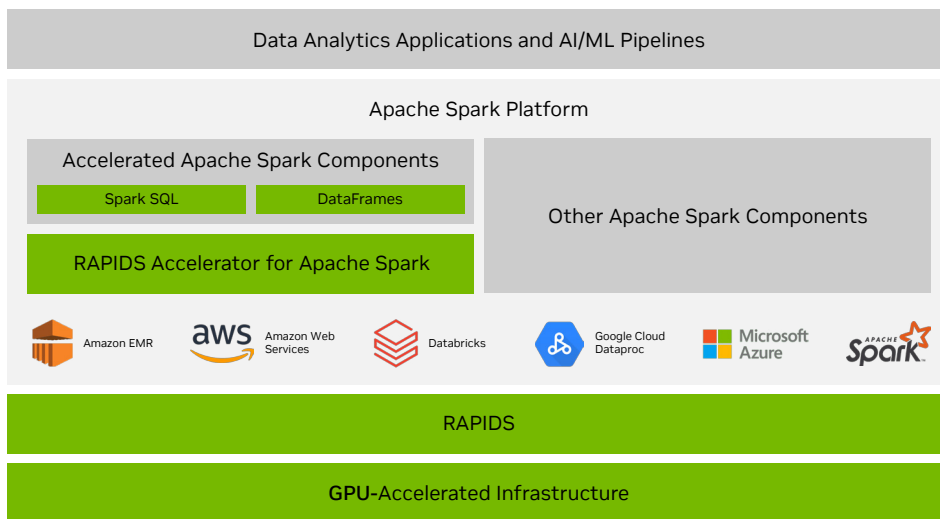
The RAPIDS Accelerator includes the **Accelerated Spark Analysis Tool**, which analyzes workloads and provides savings estimates. This makes it easy to determine which workloads would benefit most from GPU acceleration.



Based on NDS benchmark.  
Does not include NVIDIA AI Enterprise license cost.

## How Does It Work?

The RAPIDS Accelerator supports data analytics workflows for GPU-accelerated infrastructure. By operating as a plug-in to Apache Spark platforms, it accelerates workflows without requiring code changes. It can work with Spark standalone, YARN clusters, and Kubernetes clusters. Additionally, advanced features such as Adaptive Query Execution for SQL and Dynamic Partitioning Pruning ensure consistent accelerated performance.



## NVIDIA AI Enterprise Software Platform

The RAPIDS Accelerator is available with **NVIDIA AI Enterprise**, an end-to-end, secure, cloud-native AI software platform optimized to accelerate enterprises to the leading edge of AI. NVIDIA AI Enterprise delivers validation and integration for NVIDIA AI open-source software, access to AI solution workflows to speed time to production, certifications to deploy AI everywhere, and enterprise-grade support, security, and API stability to mitigate the potential risks of open-source software. NVIDIA AI Enterprise is optimized and tested for production deployments on all major Apache Spark platforms.

## NVIDIA Enterprise Support

NVIDIA AI Enterprise includes **NVIDIA Enterprise Business Standard Support**.

With direct access to NVIDIA experts to support implementation and optimization, defined service-level agreements, and priority security notifications and fixes, this service ensures that projects stay on track.

## Ready to Get Started?

To learn more about GPU-Accelerated Apache Spark, visit:

[nvidia.com/spark-tool](https://nvidia.com/spark-tool)

© 2023 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, and RAPIDS 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. 2866514. AUG23

