Big Data Demands Big Performance

Data analytics is the life blood of today’s business success. While there are many different options for building big data analytics solutions, open-source Apache™ Hadoop® is one solution that many businesses consider. To analyze data and bring insights to customers, more businesses are looking for faster, more efficient analytics solutions; more businesses are moving to Apache Spark™ to meet this requirement.

Spark takes advantage of high-performance DRAM and flash to bring faster time to insights for your business. Micron understands the demands put on these resources by these solutions, and we have used our expertise to define a high-performance Micron accelerated Hadoop analytics solution using Spark to move your business analytics to the next level.

Built around core industry-standard x86 servers using Micron DDR4 DRAM and Micron’s industry-first, QLC-based SATA SSDs, along with NVMe for optimized write performance, our all-flash Hadoop/Spark reference architecture can offer from 30% to 52% improved performance over legacy hard disk-based deployments.¹

Micron’s 5210 ION SSD is the industry’s first quad-level cell (QLC) flash SSD, ideal for low-cost, high-capacity read-centric workloads; exactly what is needed for cost-effective, high-performance big data and analytics solutions.

Micron’s 9300 family of NVMe TLC SSDs are the perfect solution when microseconds count. Offering optimized read and write throughput,² the 9300 provides the performance required for caching, high-performance transaction processing and AI/ML/DL requirements.

---

¹ Subset of HiBench benchmark suite optimized to use Apache Spark. Additional HiBench benchmarks not compatible with Spark at the time of testing.
² Offered on Micron 9300 PRO 7.68TB and 15.36TB and 9300 MAX 6.4TB and 12.8TB SSDs; throughput of 3.5 GB/s at 128KB block size for sequential read and write operations.
Apache for Data Analytics

The Apache open source projects for big data analytics provide unique value for your solutions:

Apache HDFS and YARN

- **Scalable**: Apache HDFS provides a scalable distributed filesystem designed to support thousands of nodes.
- **Fault Tolerant**: HDFS provides a robust set of data protection and management services.
- **Cost-Effective**: HDFS and YARN are open source, providing a low-cost entry point to big data capacity and performance.

Apache Spark

- **Performance**: Spark is a high-performance data analytics engine that efficiently executes both batch and streaming data sets.
- **Flexible**: Spark provides the flexibility data analytics programmers and data scientists need to be successful.

Micron Accelerated Storage Solutions Deliver

**Optimized**: Micron leverages its deep technical knowledge about advanced NAND and DRAM technology and SSD design to maximize solution performance.

**Trusted**: Micron is a trusted component provider to many OEM and cloud providers so that you can trust Micron for your solutions.

**Simple**: Micron Accelerated Solutions provide prescriptive guidance that provide you with all the information you need to be successful. Built upon “x86 architecture,” Micron Accelerated Solutions focus on the Micron value and performance, allowing you to use the server platform of your choice.

**Excellence**: Micron solutions engineers have years of experience working with the most important workloads and software demanded by highly successful businesses.