

cloudera[®] CONNECT



SPLICE MACHINE

INDUSTRY

Databases

WEBSITE

www.splicemachine.com

COMPANY OVERVIEW

Based in San Francisco, Splice Machine provides the first hybrid, in-memory RDBMS powered by Hadoop and Spark. It is designed to scale real-time applications using commodity hardware and distributed in-memory technology without application rewrites.

PRODUCT OVERVIEW

The Splice Machine RDBMS is the first hybrid, in-memory RDBMS powered by Hadoop and Spark. Leveraging in-memory technology from Spark and scale-out capabilities from Hadoop, Splice Machine can replace Oracle[®] and MySQL[™] databases, while increasing performance by 10-20x at one-fourth the cost. With an innovative, hybrid architecture and advanced resource isolation, the Splice Machine RDBMS provides exceptional performance for simultaneous OLAP and OLTP workloads, enabling companies to unlock the insights in their Big Data to make decisions in the moment.

SOLUTION HIGHLIGHTS

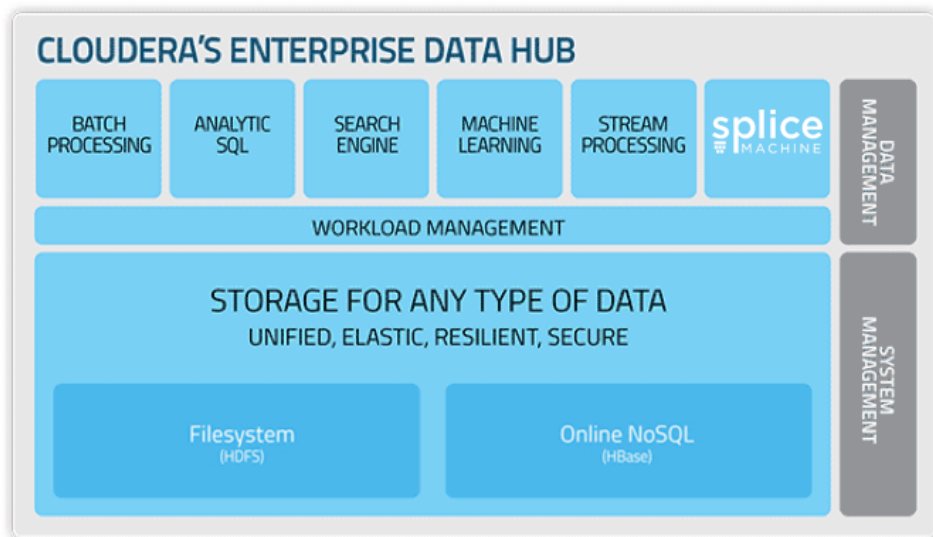
- ACID compliant transactions
- Standard ANSI-SQL
- Complex joins
- Secondary indexes
- Sub-queries
- Triggers
- User-defined functions (UDFs)
- Column-level security

Splice Machine and Cloudera Enable Customers to support simultaneous OLTP and OLAP workloads on Hadoop

Facing increased data growth and cost pressures, scale-out technology has become very popular as more businesses become frustrated with their costly scale-up RDBMSs. Leveraging Cloudera CDH, a hybrid, in-memory RDBMS is a natural choice to replace traditional relational databases, which struggle with cost or scaling issues.

The Best of All Worlds

Splice Machine, a technology partner with Cloudera, delivers an innovative hybrid of in-memory computing from Spark and disk-based storage from Hadoop. Like Oracle and MySQL, it is an operational database that can handle operational (OLTP) or analytical (OLAP) workloads, while scaling out cost-effectively from terabytes to petabytes on inexpensive commodity servers. Unlike in-memory-only databases, Splice Machine does not force companies to put all of their data in-memory, which can become prohibitively expensive as data volumes grow. Splice Machine uses Spark in-memory computation to materialize the intermediate results of long-running queries, but leverages the power of HBase to durably store and access data at scale. The partnership now allows businesses the best of all worlds: a standard SQL database, the proven scale-out of Hadoop, and the ability to leverage current staff, operations, and outstanding in-memory performance for OLAP queries with Spark.



Using HBase as its storage layer, Splice Machine natively runs in the Hadoop stack, ensuring smooth integration to the ecosystem of Hadoop tools, such as Yarn, Hive, Spark, ZooKeeper, Flume, Storm, Kafka, and Sqoop. Using HBase co-processors, Splice Machine can efficiently execute complex SQL statements concurrently on parallel HBase nodes without the overhead of MapReduce.

Thanks to standards-based ODBC and JDBC drivers, Splice Machine provides seamless connectivity to BI tools, such as Tableau and MicroStrategy, and SQL tools, such as Toad and DbVisualizer.

Benefits of Cloudera

Stores and Analyzes Any Type of Data

- > Store and analyze huge volumes of structured and unstructured data that were previously impossible or impractical
- > No need to define a data model during ingest
- > Supports multiple, flexible schemas

Massively Scalable

- > Brings compute to the data, so no need for expensive data movement prior to analysis
- > Scales linearly on industry standard x86 hardware

Industry-Leading Management and Support

- > Centralized, end-to-end management through Cloudera Manager, supporting deployment, configuration, monitoring, and issue resolution
- > Makes handling even the largest enterprise clusters simple and efficient
- > Worldwide, dedicated team of Hadoop experts and project committers working for you

Benefits of Splice Machine

Leverage Existing SQL Tools

- > Power real-time operational applications without application rewrites
- > Eliminate retraining of SQL-based personnel

Cost-Effective Scaling

- > Scale-out on commodity servers using the proven auto-sharding of HBase
- > Increase performance by 10-20x at ¼ the cost vs. traditional databases

Real-Time Updates with Transactions

- > Full ACID transactions across rows and tables using snapshot isolation
- > State-of-the-art design delivers very high throughput

Fast In-Memory Performance

- > Accelerate OLAP queries for large-scale data
- > Resilient to node failures

Major Benefits

Companies using Splice Machine and Cloudera to become real-time, data-driven businesses can leapfrog their competition with these major benefits:

Cost-Effective Scaling and Performance with Commodity Hardware

Splice Machine leverages the proven auto-sharding of HBase to scale to dozens of petabytes with commodity servers. Parallelizing queries in a shared-nothing architecture. Splice Machine delivers compelling results **versus traditional RDBMSs like Oracle and IBM DB2**:

- 10-20x increase in query speeds
- 75% reduction in TCO

Advanced In-Memory Technology

Splice Machine embeds Apache Spark – a fast, open source engine for large-scale data processing – to accelerate OLAP queries. Spark has very efficient in-memory processing, can spill to disk (instead of dropping the query) and is unique in its resilience to node failures.

Real-Time Updates with Transactions

Database transactions ensure that real-time updates can be reliably executed without data loss or corruption. Transactions enable zero-downtime updates or ETL to data warehouses, as data can be updated while reports simultaneously see a consistent view of the data.

Reduced Risk with Safe Journey

Splice Machine has designed a Safe Journey program to significantly ease the effort and risk for companies migrating to a Splice Machine database. The Safe Journey program includes a proven methodology:

- Helps choose the right workloads to migrate
- Implements risk-mitigation best practices
- Includes commercial tools that automate most of the PL/SQL conversion process

About Cloudera

Cloudera delivers the modern data management and analytics platform built on Apache Hadoop and the latest open source technologies. The world’s leading organizations trust Cloudera to help solve their most challenging business problems with Cloudera Enterprise, the fastest, easiest and most secure data platform available for the modern world. Our customers efficiently capture, store, process and analyze vast amounts of data, empowering them to use advanced analytics to drive business decisions quickly, flexibly and at lower cost than has been possible before. To ensure our customers are successful, we offer comprehensive support, training and professional services. Learn more at <http://cloudera.com>.

About Splice Machine

The Splice Machine RDBMS is the first hybrid, in-memory RDBMS powered by Hadoop and Spark. Leveraging in-memory technology from Spark and scale-out capabilities from Hadoop, Splice Machine can replace Oracle® and MySQL™ databases, while increasing performance by 10-20x at one-fourth the cost. With an innovative, hybrid architecture and advanced resource isolation, the Splice Machine RDBMS provides exceptional performance for simultaneous OLAP and OLTP workloads, enabling companies to unlock the insights in their Big Data to make decisions in the moment.

