



# The First Hybrid, In-Memory RDBMS Powered by Hadoop and Spark





The content of this white paper, including the ideas and concepts contained within, are the property of Splice Machine, Inc. This document is considered proprietary and confidential, and should not be reproduced or reused in any way without the permission of Splice Machine.  
Published 11/15

# Table of Contents

Executive Summary .....	2
Decisions in the Moment .....	3
Inadequate Technology Options .....	4
The Advantages of SQL, the Scale Out of NoSQL, and the Performance of In-Memory ...	5
Proven Technologies .....	6
Comparison to Other Hadoop Technologies .....	8
Cross-Industry Use Cases .....	9
Applications .....	11
Key Features and Benefits of Splice Machine .....	15
Conclusion .....	21

# Executive Summary

## Splice Machine: The First Hybrid, In-Memory RDBMS Powered by Hadoop and Spark

Business is moving faster than ever before, driven by the proliferation of social media, cloud computing and mobile devices. To remain competitive, companies need to make decisions in the moment. Real-time decision-making demands technology that can manage the massive increases in both data volume and velocity that Big Data brings to the table.

Splice Machine builds upon proven open-source technologies like Hadoop®, HBase™, Apache Spark™ and Apache Derby™ to create an innovative, general-purpose database that provides the following benefits when compared to traditional RDBMSs such as Oracle, IBM DB2, or MySQL:

- **10-20x Faster** – leverages HBase, the distributed NoSQL DB, as well as in-memory cluster computing from Spark
- **¼ the Cost** – scales out on commodity hardware using Hadoop
- **ANSI SQL** - leverages existing SQL-based analysts, reports, and applications without rewrites
- **Distributed Transactions** – ensures reliable updates across multiple rows and tables, based on advanced research by Google
- **Flexible** – provides excellent performance for simultaneous OLTP and OLAP workloads
- **Elastic** – increases or decreases scale in just a few minutes

## Use Cases

The Splice Machine RDBMS is ideal for a broad range of use cases, including digital marketing, ETL acceleration, operational data lakes, data warehouse offloads, IoT applications, web, mobile, and social applications, and operational applications. It also enables a new class of data-driven applications that only Splice Machine can power in real-time.

# Decisions in the Moment

Being data-driven is almost cliché in business today, but the little known fact is that most business decisions are still made using stale data. Today's systems can only catch up to yesterday's data. What businesses need is to be able to make decisions in the moment. However, this is extraordinarily challenging because increasing data volume and velocity is overwhelming traditional databases.

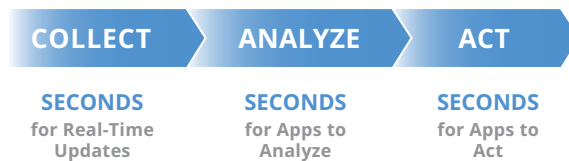
Data volume is growing 30-40% per year,<sup>1</sup> approximately an order of magnitude faster than IT budgets, which are growing at 3-4% per year.<sup>2</sup> Reaching this scale with traditional RDBMSs – the backbone of IT, requires you to scale up with increasingly larger, specialized servers from companies such as Oracle and IBM, but this is cost prohibitive.

Data velocity has accelerated greatly, driven by real-time data streams from social media, programmatic marketing, mobile devices, and the Internet of Things (IoT). For instance, in digital marketing, companies want to deliver real-time, personalized experiences across channels based on the “last 5 clicks” to drive greater engagement and conversion. This requires making decisions in the moment with real-time data, not after the fact with day-old data ETL'd (Extracted, Transformed, and Loaded) from another system. This also demands databases that can simultaneously handle OLTP and OLAP workloads to eliminate the ETL delay for analysis.

**fig. 1**  
By shortening the decision lifecycle, data-driven companies will create competitive advantage.

## REAL-TIME, DATA-DRIVEN BUSINESS

Real-Time Actions Based on Real-Time Data



## CURRENT BUSINESS

Delayed Actions Based on Out-of-Date Data



<sup>1</sup>Source: 2013 IBM Briefing Book

<sup>2</sup>Source: Gartner, Worldwide IT, Spending forecast, 3Q13 Update

Compared to SQL-on-Hadoop technologies, the Splice Machine RDBMS is unique as the only transactional ANSI SQL database leveraging both Hadoop and Spark.

## Inadequate Technology Options

Current technology solutions cannot support real-time, data-driven businesses that may need to process terabytes or even petabytes of data to make decisions in the moment.

### Traditional Commercial RDBMSs

Traditional RDBMSs (Relational Database Management Systems) such as Oracle® or IBM DB2® can support real-time updates, but they require expensive specialized hardware to “scale up”. Costing up to millions of dollars per system, these specialized systems become cost prohibitive quickly.

### Traditional Open-Source RDBMSs

Traditional open source databases such as MySQL™ and PostgreSQL® are unable to scale beyond a few terabytes without manual sharding. Manual sharding requires a partial rewrite of every application and becomes a maintenance nightmare to periodically rebalance shards. For most companies, manual sharding is not a viable option.

### In-Memory-Only Databases

New in-memory-only databases provide very fast performance, but all of the data must be stored in-memory, which can become very expensive beyond a few terabytes. To solve these cost issues while still maintaining their speed advantages, in-memory databases need to have spill-to-disk and handle node failures without dropping queries.

### Hadoop

New Big Data technologies such as Hadoop and HBase provide cost-effective and proven scalability from terabytes to petabytes with inexpensive commodity servers. They also provide very limited SQL support, the de facto data language in most organizations. The lack of full ANSI SQL support is a major barrier to Hadoop adoption, as companies cannot leverage existing investments in SQL-trained resources as well as SQL-based applications and Business Intelligence (BI) tools. These factors have combined to prevent companies from getting value out of their current Hadoop deployments beyond being a massive data store.

# Splice Machine: The Advantages of SQL, the Scale Out of NoSQL, and the Performance of In-Memory

Splice Machine delivers a unique hybrid of state-of-the-art technologies:

- **Distributed, Transactional SQL** – Support existing applications without rewrites or retraining of existing SQL-based analysts
- **Scale-Out** – Cost-effectively scale out on commodity hardware with proven technology from Hadoop and Spark
- **In-Memory** – Experience outstanding performance for OLAP queries with in-memory technology from Spark

## Cost-Effective, Hybrid Architecture

The Splice Machine RDBMS is an innovative hybrid of in-memory computing from Spark and disk-based storage from Hadoop. 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.

## State-of-the-Art Optimizer

Splice Machine started with Apache Derby, an ANSI SQL Java database, and replaced its storage layer with HBase/Hadoop and Spark. The Splice Machine optimizer automatically evaluates each query and sends it to the right computation engine: OLTP queries (i.e., small read/writes, range queries) go to HBase/Hadoop, OLAP queries (i.e., large joins or aggregations) go to Spark.

## Mixed OLTP/OLAP Workloads

The Splice Machine RDBMS is a high-performance database designed to simultaneously handle OLTP and OLAP queries. With separate processes and resource management for Hadoop and Spark, Splice Machine can ensure that OLAP queries do not interfere with OLTP queries.

Questions to consider when selecting your next database:

Can it support **simultaneous OLTP and OLAP queries?**

Does it **scale-out on commodity hardware?**

Does it have **in-memory technology?**

Can it **support ACID transactions?**

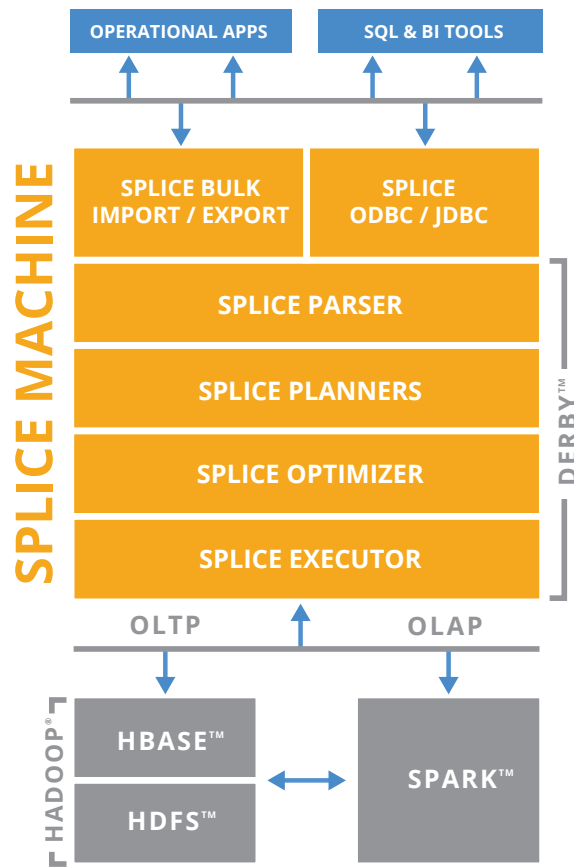
# Proven Technologies

Splice Machine marries three proven technology stacks: Apache Derby, HBase/Hadoop and Spark.

## Apache Derby

With over 15 years of development, Apache Derby is a Java-based SQL database. Splice Machine chose Derby because it is a full-featured ANSI SQL database, lightweight (<3 MB), and easy to embed into the HBase/Hadoop stack.

fig. 2  
The Splice Machine  
Solution Stack





**Splice Machine is unique in its ability to enable a new breed of real-time, analytical applications that can collect, analyze, and react to data in real-time to transform industries.**

## HBase and Hadoop

HBase and Hadoop have become the leading platforms for distributed computing. HBase is based on Google's Big Table technology and uses the Hadoop Distributed File System (HDFS) for reliable and available persistent storage. HBase provides auto-sharding and failover technology for scaling database tables across multiple servers.

It also enables real-time, incremental writes on top of the immutable Hadoop file system. HBase and Hadoop are the only technologies proven to scale to dozens of petabytes on commodity servers and are used by companies such as Facebook, Twitter, Adobe and Salesforce.com. Splice Machine chose HBase and Hadoop because of their proven auto-sharding, replication, and failover technology.

## Apache Spark

Apache Spark has emerged as the leading in-memory cluster computing engine. It has very efficient in-memory processing, with spill-to-disk so that queries that do not fit in memory, do not fail. Unlike many other in-memory technologies, Spark is resilient to node failures. In other words, it can retry work on other active nodes in light of a failure so that a computation does not fail.

Splice Machine integrates these technology stacks by replacing the storage engine in Apache Derby with HBase and its query executor with HBase and Spark. Splice Machine retains the Apache Derby parser, but it redesigned the planner, optimizer, and executor to leverage the distributed HBase computation engine.

Each physical HBase node has two separate processes and memory spaces: the HBase Region Server with the Splice Machine parser, planner, and optimizer embedded; and the Spark computation engine. The Splice Machine executor compiles its planned computation in bytecode that is serialized to each distributed HBase region (i.e., data shard) via HBase co-processors or to each Spark worker, depending on which computation engine is selected for the statement. This enables massive parallelization by pushing the computation down to each distributed data shard.

## Compatible With Any Hadoop Distribution

Splice Machine can be used with any standard Hadoop distribution that has both HBase and Spark. Supported Hadoop distributions include Cloudera, MapR and Hortonworks.

# Comparison to Other Hadoop® Technologies

The SQL-on-Hadoop space has become crowded with many technologies. Unlike Splice Machine, these technologies are designed for analytics only and cannot support high volumes of real-time updates. This means that updates require extensive and time-consuming ETL (Extract, Transformation, and Load) processes that delay reports and often run only on a daily basis. Most of the technologies also provide limited and non-standard dialects of SQL.

Compared to SQL-on-Hadoop technologies, the Splice Machine RDBMS is unique as the only transactional ANSI SQL database leveraging both Hadoop and Spark. Transactions ensure that real-time updates can be reliably made without data loss or corruption. It is also the only general-purpose SQL database on Hadoop that can simultaneously support both operational and analytical workloads.

**Cost-effective  
scaling with  
commodity  
software**

## Cross-Industry Use Cases

General-purpose databases such as MySQL and Oracle are used across a range of operational (OLTP) and analytical (OLAP) applications. As a general-purpose database that can scale-out with commodity servers, Splice Machine can be used anywhere that MySQL or Oracle can be used, but it really excels whenever the database size exceeds a few terabytes.

### Unlock the Value in Hadoop

Hadoop has become a standard when it comes to storage of massive amounts of data, but it is a file system, not a database. With Splice Machine, companies can use Hadoop to power operational applications and analytics on real-time data.

### NoSQL Falls Short

Companies requiring cost-effective scaling on commodity hardware have historically turned to NoSQL solutions such as MongoDB®, Cassandra™, or HBase. However, the scalability comes at a great cost — little or no SQL language support, as well as lack of transactions, joins, aggregations, and secondary indices. Application developers can quickly discover these omissions and are then forced to recreate that functionality for each application – a costly, time-consuming, and error-prone effort. Splice Machine enables these developers to enjoy the best of both worlds — standard SQL AND scale-out on commodity hardware.

### “Big Iron” Too Expensive

Traditional RDBMSs such as Oracle and IBM DB2 typically require specialized server hardware to scale beyond a few terabytes. Each of these servers can cost millions of dollars. Even small spikes in volume can require massive spending to scale to requirements. For most companies, the cost becomes prohibitive to handle terabytes, let alone petabytes, of data. Splice Machine enables highly cost-effective scaling out on commodity servers for companies currently trapped with spiraling costs from “Big Iron”.

### MySQL Can't Scale

Open source databases such as MySQL and PostgreSQL can no longer scale once the data volume exceeds a few terabytes. Manual sharding is the only option to handle increased data volumes. It requires a partial rewrite of every application as well as the rebalancing of shards. Splice Machine can help those companies scale beyond a few terabytes with the proven auto-sharding capability of HBase.

## Data Warehouse Offload

Teradata or Netezza data warehouses employ very expensive, specialized hardware. For many companies, they recognize that up to 50% of the workloads (e.g., operational reports) on those systems can be offloaded to less expensive, scale-out systems. With Spark in-memory technology, the Splice Machine RDBMS can deliver similar performance at a dramatically lower cost.

## “Future Proof” New Apps

It can often be difficult to anticipate future data requirements for new applications. However, as businesses become more data intensive, it is likely that many applications will exceed a few terabytes of data in their lifetime. No company wants to build an application and then endure a costly database migration in a few years because data growth was beyond expectations. Splice Machine provides a “future-proof” database platform that can scale cost-effectively from gigabytes to petabytes for new applications.

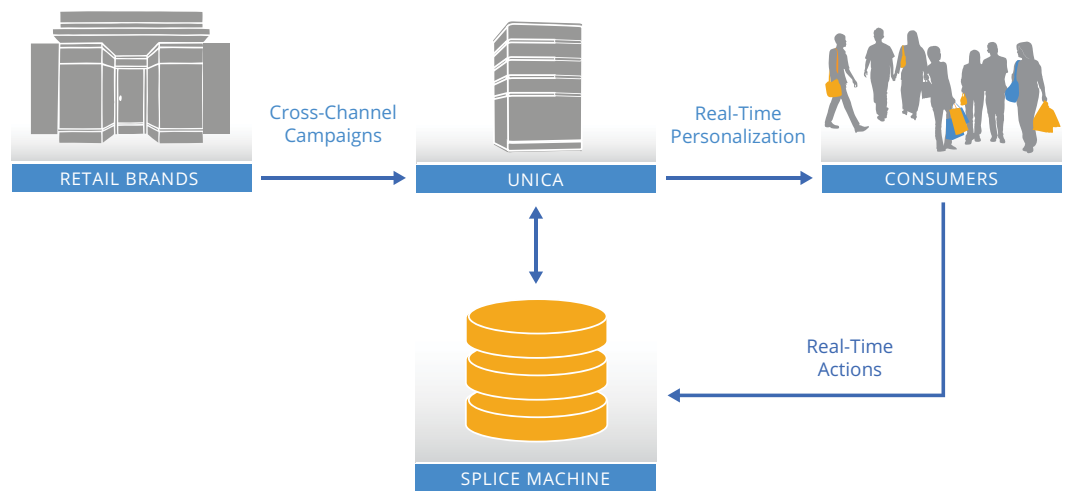
# Applications

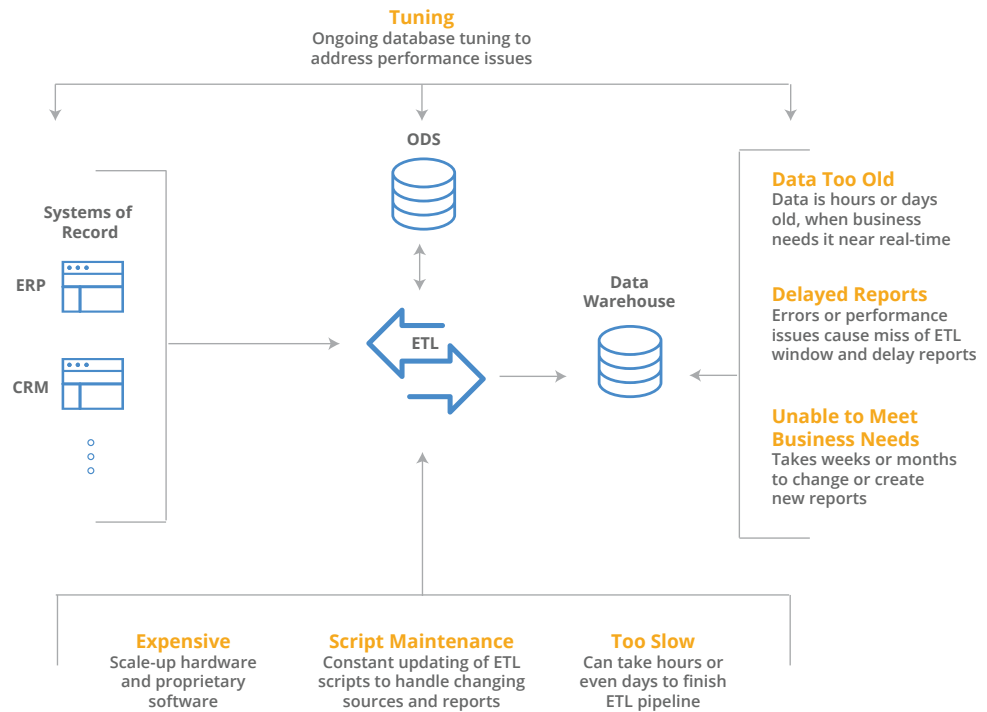
General-purpose databases like Splice Machine can be used in almost every industry across many different real-time applications. However, Splice Machine is unique in its ability to enable a new breed of real-time, analytical applications that can collect, analyze, AND react to data in real-time to transform industries. The rest of this section highlights a few examples of real-time applications that Splice Machine can power.

## Digital Marketing

Consumer marketing optimizes interactions with millions of consumers across websites, mobile devices, emails and ads. The “holy grail” is real-time personalization, which shows “the right message, to the right person, at the right time”. Traditional solutions employ analytical models based only on yesterday’s data. Splice Machine enables companies to make decisions in the moment with real-time data, not after the fact with day-old data ETL’d (Extracted, Transformed, and Loaded) from another system.

**fig. 3**  
Splice Machine’s ability to process real-time updates enables retail campaigns that provide true real-time personalization.





**fig. 4**  
ETL issues create inertia and hidden costs for businesses

## ETL Acceleration

The ETL process is a hidden burden within most IT departments, raising costs and slowing down decision-making. With Big Data, ETL pipelines have become an even bigger bottleneck for applications and analysts who rely on the data flowing through them. By replacing overwhelmed operational data stores (ODS) using RDBMSs like Oracle with the Splice Machine RDBMS, companies can drive ETL lag down from days and hours to minutes and seconds. The Splice Machine RDBMS combines the best of all worlds – the transactional integrity of an RDBMS, the proven scale out of Hadoop and the in-memory performance of Spark – to provide an easier way to scale out the capacity of the ETL pipeline.

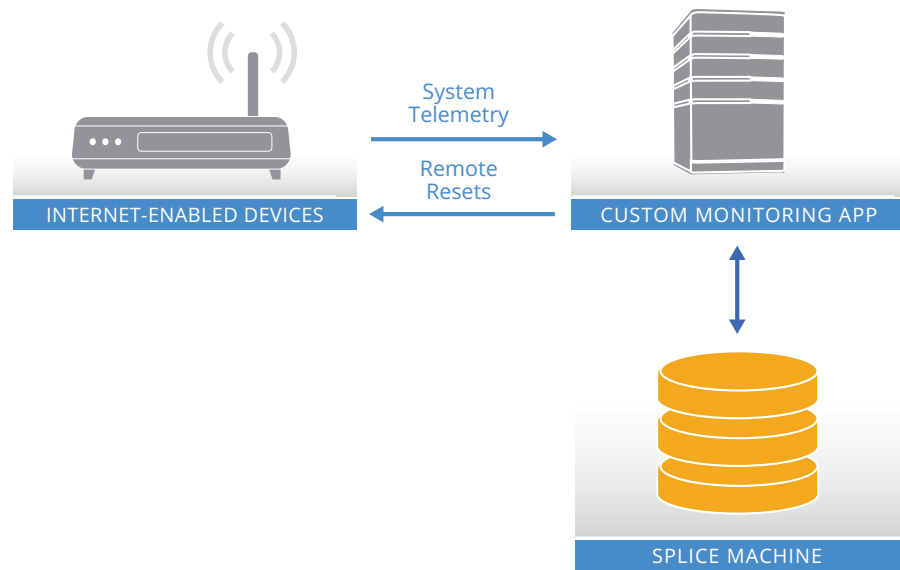
## Operational Data Lake

An operational data lake (ODL) is a modern replacement for older Operational Data Stores (ODSs). An ODL enables offloading real-time reporting and analytics from more expensive OLTP and data warehouse systems, as well as performing aggregation and transformation for ETL processes.

Compared to traditional ODSs using scale-out technologies, an ODL provides a 20x price/performance advantage with scale-out technology. It can also handle semi-structured and unstructured data as part of a larger Hadoop-based data lake.

Many companies find that replacing an ODS with an operational data lake represents an excellent first project in Big Data.

fig. 5  
Splice Machine powers system telemetry applications that transform the way devices are serviced.



## Internet Of Things

Telemetry is a stream of data from devices used to detect conditions that require action in real-time. These conditions can be system failures, resource optimizations, fraudulent actions, marketing or service opportunities, or intrusions. System-intensive industries such as telecoms, traditional utilities, internet service providers (ISPs), and television cable/satellite companies use real-time telemetry to proactively detect failures, isolate them, and then attempt remote resets to prevent service calls and dispatch of on-site technicians.

Cross-industry examples include server monitoring to detect any degradation of website performance followed up with offers to the affected users to increase customer retention. Cyberthreat security applications monitor firewalls to correlate real-time activity with historic firewall logs to determine if there is a real threat. Traditional systems are unable to ingest and analyze the sheer volume of system telemetry data. Splice Machine has the scalability to ingest, analyze, and react to massive volumes of telemetry data.

**Splice Machine has the scalability to ingest, analyze, and react to massive volumes of telemetry data.**

**Splice Machine provides real-time analytics on a Big Data scale.**

## Precision Medicine

As healthcare costs continue to rise faster than inflation, countries such as the United States have started to shift away from activity-based payments to outcome-based payments. Under the Affordable Care Act, hospitals in the United States will not be paid if a patient is readmitted in less than 30 days. This provides a powerful incentive to improve the long-term effectiveness of treatments.

The medical industry has begun to focus on precision medicine, which uses genomic, clinical, and diagnostic data to tailor a treatment for a particular individual. Genomic data, which can be massive, can help predict outcomes and possible complications, as well as tailor medications for complex diseases such as cancer. Clinical data includes Electronic Health Records (EHRs) to ensure that care is coordinated across multiple treatments and medical providers. Diagnostic data, especially from real-time device data both at the hospital and at home, can trigger intervention before complications reach a tipping point and require hospital readmission.

Traditional systems are unable to handle the massive volume of data required for precision medicine. Splice Machine can support precision medicine by handling the huge stores of genomic data, providing the structure for EHRs, and ingesting large quantities of real-time device data.

## Operational Applications

Traditional RDBMSs typically power real-time operational systems (OLTP), including enterprise software as well as web, mobile, and social applications. These systems require a high volume of real-time reads and writes with transactional integrity. With ACID (Atomicity, Consistency, Isolation, Durability) transactions, Splice Machine can power these applications, while simultaneously providing cost-effective scaling with commodity hardware.

## Operational Analytics

Operational analytical systems (OLAP) on traditional RDBMSs have traditionally supported real-time, ad-hoc queries on transactional and aggregated data. However as that transactional data gets augmented with an explosion of real-time web, social, mobile and machine-generated data, traditional systems cannot scale.

Specialized appliances such as Teradata and Netezza can scale to handle massive data sets, but they are expensive and require extensive ETL. Splice Machine provides the best of both worlds: real-time operational analytics on a Big Data scale.



# Key Features and Benefits of Splice Machine

## Standard ANSI SQL to Leverage Existing SQL Tools & Skill Sets

Leveraging the proven SQL processing of Apache Derby, Splice Machine provides a true RDBMS on Hadoop. Splice Machine is SQL-99 compliant and includes functionality such as:

- Transactions
- Joins
- Secondary indexes
- Aggregations
- Sub-queries
- Triggers
- Constraints
- User-defined functions (UDFs)
- Column-level security
- Stored procedures
- Views
- Virtual tables
- Window functions

Standard SQL enables companies to leverage existing SQL-trained resources to develop and maintain applications on Splice Machine. Existing applications can migrate to Splice Machine with minimal effort, thanks to ODBC and JDBC drivers that provide seamless connectivity to BI tools such as IBM Cognos®, SAP Business Objects®, Tableau® and MicroStrategy®, ETL tools like Informatica® and Ab Initio®, statistical tools such as SAS® and R®, and SQL tools such as Toad® and DbVisualizer™.

## Query Acceleration with 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, but it can spill to disk (instead of dropping the query) if the query processing exceeds available memory.

More importantly, Spark is unique in its resilience to node failures, which can happen in a commodity cluster. Other in-memory technologies will drop all queries associated with a failed node, while Spark uses ancestry to regenerate its in-memory Resilient Distributed Datasets (RDDs) on another node and therefore does not require redundant storage.

## Cost-Effective Scaling with Commodity Hardware

Splice Machine leverages the proven auto-sharding of HBase to scale with commodity servers. At companies such as Facebook, HBase has scaled to dozens of petabytes using inexpensive commodity hardware.

HBase horizontally partitions or splits each table into smaller chunks or shards that are distributed across multiple servers. For performance reasons, it automatically splits shards that become too large. This auto-sharding occurs automatically without impacting applications built on Splice Machine.

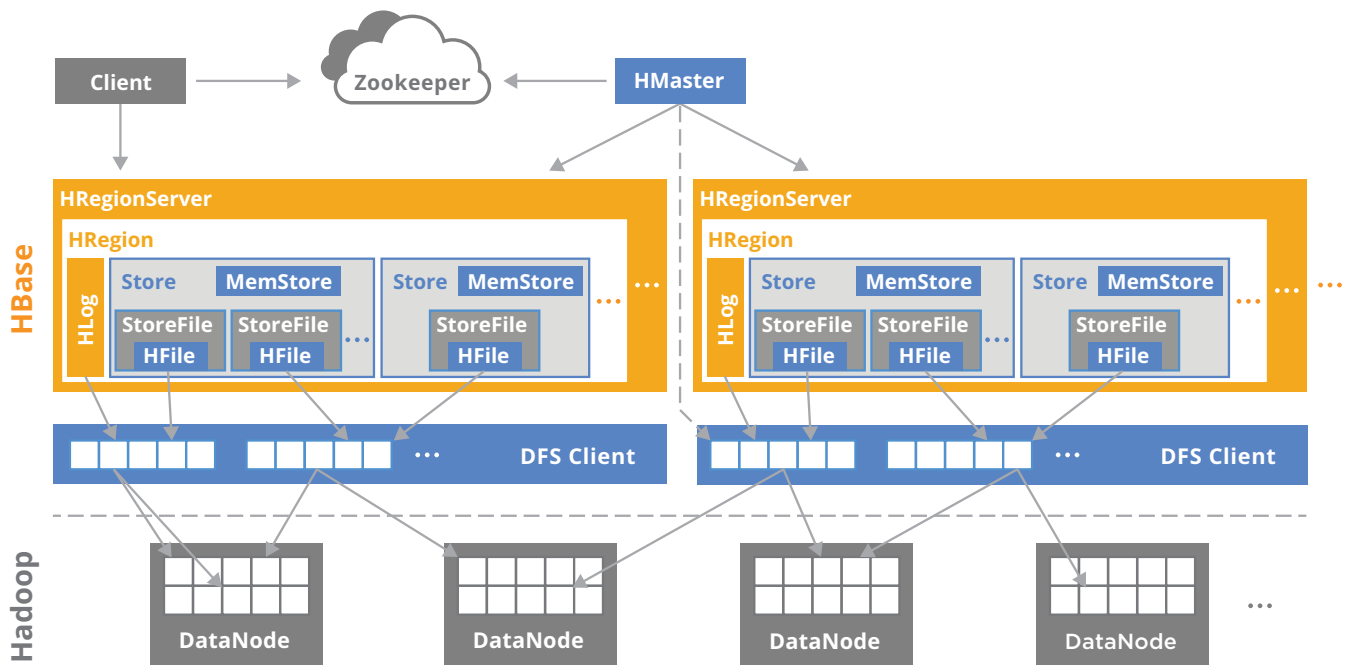


fig. 6  
HBase architecture

## Unprecedented Support for Simultaneous OLTP & OLAP Workloads

The hybrid in-memory architecture of the Splice Machine RDBMS is designed to provide high performance for simultaneous OLTP and OLAP queries. Spark provides the performance of in-memory technology for OLAP queries, while HBase can scale out to petabytes for OLTP queries.

**Advanced Optimizer.** The Splice Machine RDBMS delivers a seamless integration of both Spark and HBase. The Splice Machine optimizer automatically evaluates each query and sends it to the right computation engine: OLTP queries (i.e., small read/writes, range queries) go to HBase/Hadoop, OLAP queries (i.e., large joins or aggregations) go to Spark.

**Isolated Resource Management.** Splice Machine uses advanced resource management to ensure high-performance for simultaneous OLTP and OLAP queries. With separate processes and resource management for Hadoop and Spark, the Splice Machine RDBMS can ensure that large, complex OLAP queries do not interfere with time-sensitive OLTP queries. Users can set custom priority levels for OLAP queries to ensure that important reports are not blocked behind massive batch processes that consume all cluster resources.

## Real-Time Updates with Transactional Integrity

Database transactions ensure that real-time updates can be reliably executed without data loss or corruption. Even in analytical applications, transactions are important. For instance, transactions guarantee that the data and secondary indexes are updated atomically. Transactions also enable zero-downtime updates or ETL to data warehouses, as data can be updated while reports simultaneously see a consistent view of the data.

Splice Machine provides full ACID (Atomicity, Consistency, Isolation, Durability) transactions across rows and tables. It uses a form of Multiple Version Concurrency Control (MVCC) called snapshot isolation that does not change records — instead it creates a new version with a unique timestamp. Each transaction can use different versions of each record to create its own “snapshot” of the database. With each transaction having its own “snapshot”, transactions can execute concurrently without locking reads. This leads to very high throughput and avoids troublesome deadlock conditions.

Splice Machine developed a state-of-the-art snapshot design based on the latest research on distributed transactions. Splice Machine extended the research of Google’s Percolator system, Yahoo Labs, and the University of Waterloo on HBase transactions with support for very large transactions and increased throughput of bulk loads.

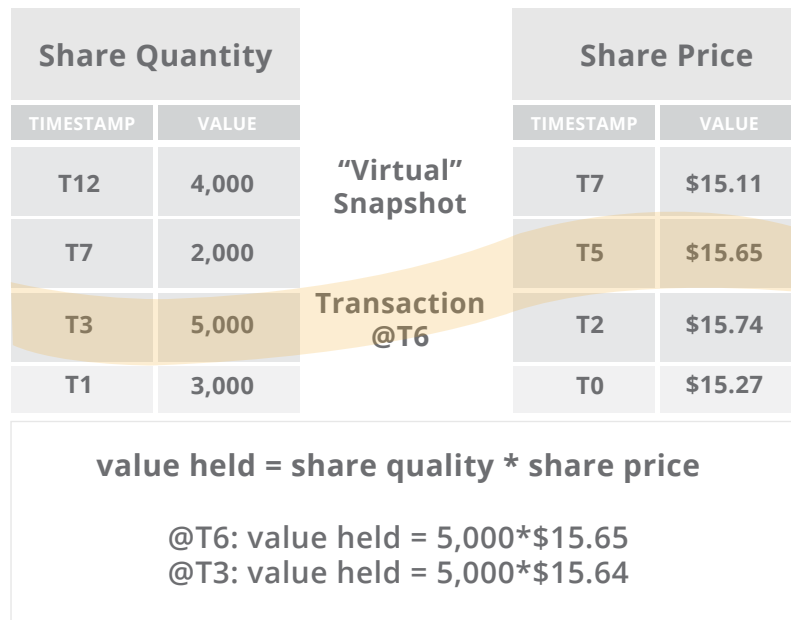


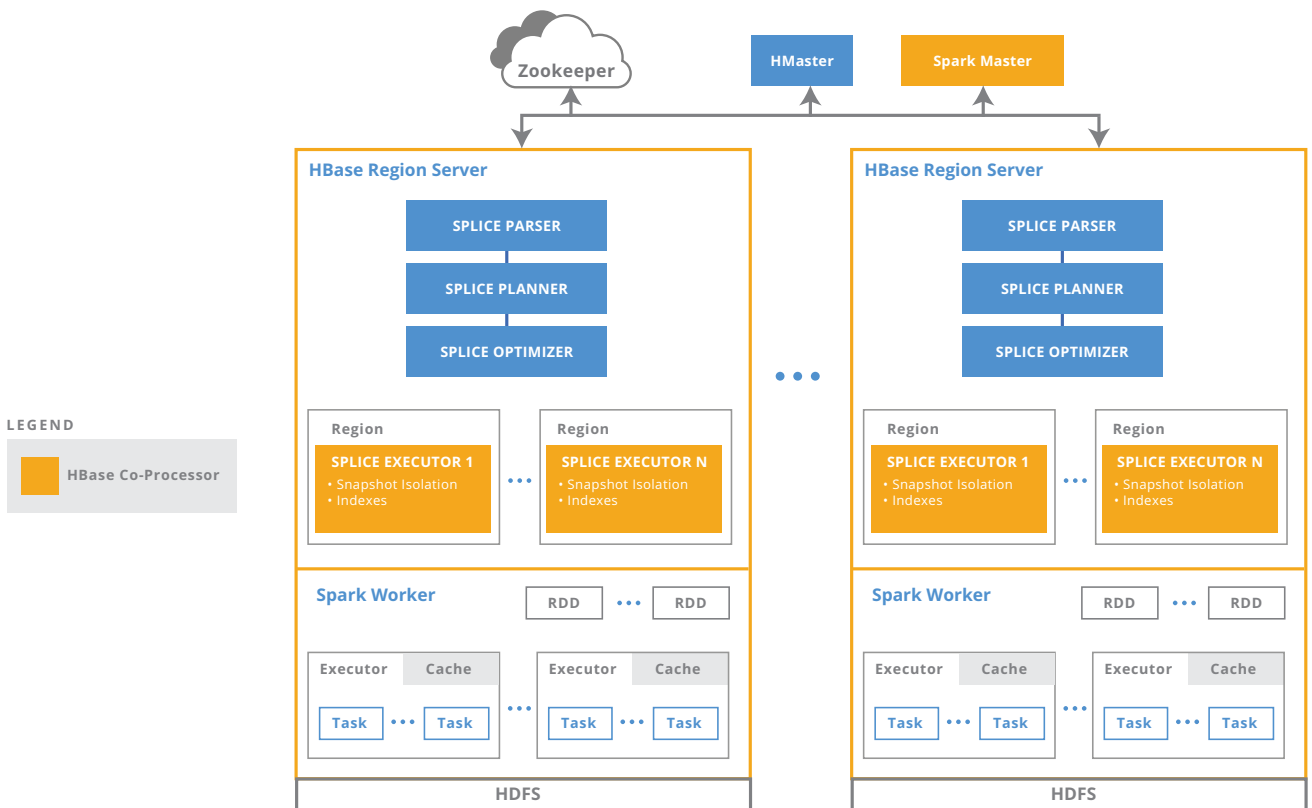
fig. 8  
High concurrency,  
ACID transactions

## High-Performance, Distributed Computing Architecture

Splice Machine is a high-performance, distributed computing architecture that delivers massive parallelization for predicates, joins, aggregations, and functions by pushing computation down to each distributed data shard. On each HBase physical node, the Splice Machine RDBMS has a separate process and memory space for HBase and Spark. It places its parser, planner, optimizer, and executor in the HBase Region Server process and uses HBase co-processors to distribute OLTP computation across regions (i.e, shards).

Built on top of HBase and Hadoop, Splice Machine also provides proven auto-sharding, failover and replication. Splice Machine supports any standard Hadoop distribution (i.e., Cloudera, MapR, Hortonworks).

fig. 8  
The Distributed Computing Architecture of HBase with Splice Machine



## Flexible, General-Purpose Database Platform

Splice Machine is a flexible database platform that can power both OLAP and OLTP workloads and efficiently handle sparse data. The storage layer for Splice Machine is HBase, which is a schema-less, key-value store. This provides Splice Machine with greater flexibility in storing data and updating schemas.

**Sparse Data Support.** Splice Machine can efficiently support super-schemas with sparse data. For instance, if a table has 1000 columns and a row only used two columns, that row would only store the data for those two columns. No extra storage would be used to store nulls for the other 998 columns. Splice Machine also enables fast, flexible schema updates. Columns can be added to tables without full table scans.

**MapReduce Input/Output Formats.** Parallelized Splice Machine queries utilize HBase co-processors and not MapReduce for distributed computation on data stored in the Hadoop Distributed File System (HDFS). However, Splice Machine does provide an API to run MapReduce, Hive, and Spark jobs against data stored in Splice Machine. This enables users to perform custom, batch-oriented analyses across the entire data set in Splice Machine using other technologies such as Storm, Kafka, Pig, MLLib, or Mahout.

**ODBC/JDBC Support.** Splice Machine also provides many options to access its data and to deploy across servers. Applications and analysts can use SQL through a command-line interface (CLI) as well as JDBC/ODBC connections. Developers can therefore embed SQL in their code in a variety of programming languages including Java, Scala, Python, C++, and Javascript.

**Access External Data and Libraries.** The Splice Machine RDBMS includes the ability to easily access external data and libraries. It can execute federated queries on data in external databases and files using Virtual Table Interfaces (VTIs). It can also execute all pre-built Spark libraries (over 130 and growing) for machine learning, stream analysis, data integration and graph modeling.

**Compatible With Any Hadoop Distribution.** Operations teams can continue to use their existing Hadoop distributions (e.g., Cloudera, MapR, Hortonworks), as Splice Machine installs quickly on each HBase Region Server. Through HBase's integration with YARN, Splice Machine can share a Hadoop cluster with other workloads.

# Conclusion

As the only hybrid, in-memory RDBMS powered by Hadoop and Spark, Splice Machine presents unlimited possibilities to application developers and database architects. Most of all, it eliminates the compromises that have been part of any database platform selection to date. You can have proven scale out with commodity hardware, the performance of in-memory technology AND standard SQL to maximize your ability to leverage current staff, operations and applications.

No longer will you be locked into specialized hardware costing millions per server. You will have a massively scalable general-purpose database that can handle operational AND analytical workloads. Splice Machine can also help you unlock the value of the data trapped in your current Hadoop deployment.

Compared to traditional RDBMSs such as Oracle, IBM DB2, or MySQL, the Splice Machine RDBMS provide the following benefits:

- **10-20x Faster** – leverages HBase, the distributed NoSQL DB, as well as in-memory cluster computing from Spark
- **¼ the Cost** – scales out on commodity hardware using Hadoop and HBase
- **ANSI SQL** - leverages existing SQL-based analysts, reports, and applications without rewrites
- **Distributed Transactions** – ensures reliable updates across multiple rows and tables, based on advanced research by Google
- **Flexible** – provides excellent performance for simultaneous OLTP and OLAP workloads
- **Elastic** – increases or decreases scale in just a few minutes

With an innovative, hybrid architecture, Splice Machine is uniquely qualified to power data-driven businesses that can harness real-time insights to take better actions. Companies can use Splice Machine to make decisions in the moment to leapfrog their competition.

# Contact Us

Please contact us at [info@splicemachine.com](mailto:info@splicemachine.com) to learn more or see a demo.