Cutting the Complexity out of Hadoop

Augmenting or Replacing Hadoop with SingleStore

While Hadoop came to the limelight in the 2010s with the promise of offering fast, efficient, and cheap storage and processing of loosely structured (big) data, it's sheen has largely worn off given the new realities of today's data requirements. Hadoop is not built for today's speed or agility of data and has failed to deliver on its promise of better performance or faster analytics. Moreover, customers are dealt with an environment that is incredibly complex, cumbersome, and costly to manage. And with the rapid evolution of the Cloud, leading organizations across the globe are looking to migrate workloads off Hadoop – to drive analytics and insights faster with a simplified architecture.

Today, we have a number of our customers who are migrating workloads off of Hadoop into **SingleStore** to drive **faster analytics** at a **fraction of the costs** with **zero complexity**. SingleStore offers a fast, distributed, highly-scalable, hybrid data platform designed to power today's data-intensive applications. It is designed to deliver maximum performance for both **transactional** (OLTP) and **analytical** (OLAP) workloads in a single unified engine to drive maximum performance for your modern applications.

With SingleStore you can ingest millions of events per second with ACID transactions using SingleStore Pipelines, while simultaneously delivering blazing-fast SQL queries on billions of rows of data, to deliver analytical performance that is up to **1000x** faster than Hadoop. And best of all, you can deploy SingleStore anywhere - **On-premises**, **in the Cloud** on AWS, Azure, or GCP or in a **hybrid mode**. And with Managed Service, SingleStore removes the need for users to manage infrastructure, provision clusters, handle upgrades, or troubleshoot failures.

Why SingleStore

1000x Faster	**50% Lower Costs**	**Zero Complexity**
Accelerate the time-to-insights by 100-1000x	Half the costs or more compared to Hadoop deployments	Fully-managed data platform that runs on-premises, or on the Cloud. Easy to deploy, manage, and integrate compared to Hadoop

Ultra-Fast Ingest
Millions of events/sec with immediate availability

Super-Low Latency
Millisecond latencies with immediate consistency
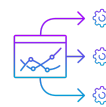
High Concurrency
Millions of real-time queries across tens of thousands of users

Fast Analytics
Accelerate analytical performance by **1000x**

Key Challenges with Hadoop

Performance	**Complexity**	**Costs**
Hadoop is not built for fast analytics or modern applications. Lagging query performance and limited real-time ingest.	Incredibly complex to manage, maintain and upgrade. Multiple projects strung together with a complex, failure-prone architecture	Hadoop costs can run as high as $10K/node, with 3X the hardware resources and a dedicated team to manage and operate

SingleStore as the Hadoop Replacement Engine

SingleStore is built for parallel streaming data ingestion, super low-latency queries, and high concurrency to help you ingest, process, analyze, and act on data instantly.
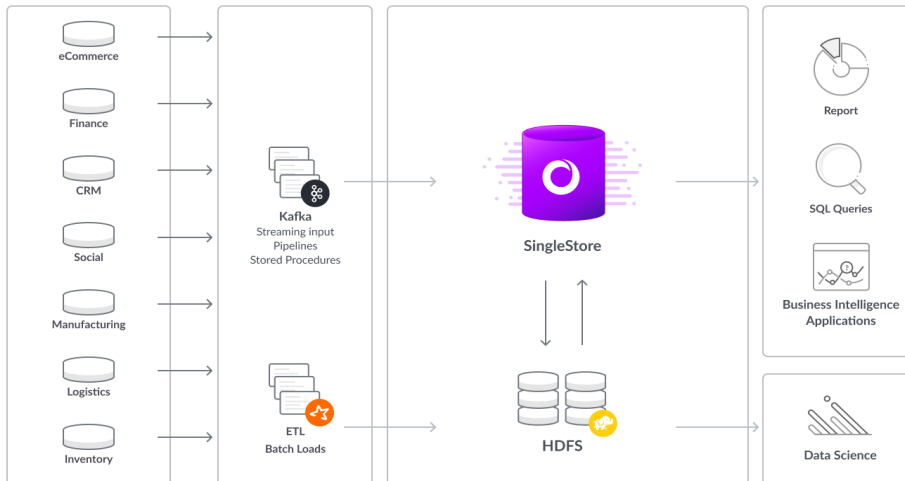
	hadoop		◯ SingleStore

Data Ingestion (NiFi, Flume, Kafka)	>	**SingleStore Pipelines** with connectors to Kafka and Spark Optimized for fast real-time ingest - up to millions of events/sec or batch uploads
Data Storage (HDFS/ HBase)	>	**Tiered Storage** - In-memory, on-disk, and the Cloud object store with separation of storage and compute (Universal Storage)
SQL Analytics (Hive, Impala, Kudu)	>	SingleStore delivers the world's **fastest SQL engine** for both transactions and analytics - Up to **100x faster** and supports multiple data types (JSON, time-series, geo, full text-search, relational)
Real-time analytics (Storm, Flink, & Spark Streaming)	>	**Kafka + SingleStore** - With our built-in Kafka connector, you can stream data directly into SingleStore from Kafka or other sources using **Pipelines** and run analytics on data as it lands, in real-time
Operational Database (HBase, Phoenix, Solr)	>	SingleStore **Universal Storage** supports both operational and analytical workloads, is ACID compliant, and SingleStore natively handles full text search as well
Batch Processing (MapReduce, Spark)	>	**Spark + SingleStore -** Batch processing in **Spark** and using **Spark Connector** to get the data into SingleStore for storage, analytics, and querying
ETL (MapReduce, Spark)	>	**Spark + SingleStore** - Transformations and data cleansing in Spark and connect to SingleStore using the SingleStore **Spark connector**

Using Singlestore to Augment or Replace Your Hadoop Infrastructure

SinglesStore can effectively augment or replace your Hadoop infrastructure to enable ultra-fast ingestion, unlimited storage, and processing, with blazing fast queries to drive fast analytics on any data anywhere.

Customers can deploy SingleStore **on-premises** or in any of the leading **cloud environments**.

Example: Hadoop Augmentation with SingleStore

Key Outcomes with SingleStore

- ✓ 100-1000x faster analytics & insights on all data
- ✓ 50%-60% cheaper than Hadoop implementations
- ✓ Easy and simplified performance architecture
- ✓ Relational data for everyone

Example of a Hadoop Augmentation with SingleStore:

Fork streaming data into SingleStore and other batch data sources into Hadoop and combine them, as needed, with a Serving layer for analytics

When is this ideal?

Transitioning analytics from **batch to real-time**, with high concurrency and query response

Uber COMCAST Kellogg's D&LLEMC Akamai SAMSUNG CISCO pandora n. monday

SingleStore - Key Features

Patented Universal Storage: Both large-scale OLTP and OLAP are supported on this single, default table type. Universal Storage gives you the best qualities of row stores and column stores while reducing data duplication, data movement, and data latency.

SingleStore Pipelines: Built-in parallel data ingestion technology natively ingests high-throughput real-time data from external sources such as Apache Kafka, Amazon S3, Azure Blob, Filesystem, Google Cloud Storage, and HDFS data source.

MySQL Compatibility: SingleStore is wire-protocol compatible with MySQL/MariaDB which offers access from hundreds of languages, 100%compatibility on data types, and 95% coverage of built-in functions, easing migrations.

Security & Compliance: Delivers Enterprise-grade security with integrated user authentication, full encryption of data in transit and at rest, and SOC2, ISO27001, HIPAA, GDPR, and CCPA compliance.

Separation of Storage and Compute: Allows users to effortlessly scale compute resources to meet the needs of any workload, while managing the storage needs completely independently.

Distributed Ingest, Bulk or Streaming, with Lock-free/ Non-Blocking Reads and Concurrency: Offers a lock-free architecture that efficiently processes transactions and updates without locking or blocking concurrent reads, resulting in delivering the capability to perform bulk and/or streaming ingestion online, simultaneously with query workload.

Suspend & Resume Workloads Effortlessly: Clusters can be Suspended and Resumed nearly instantaneously, making all of your data available when you need it, and minimizing cost when workloads are inactive.

Flexible Credit Pricing Model: Provides flexibility of on-demand or with monthly credit bundles to handle dynamic and growing compute workloads at reduced TCO.

Latency-Free Analytics: SingleStore lets you achieve ultra fast query response with high concurrency across both live and historical data using familiar ANSI SQL.

Ultra-fast Event-to-insight Performance: Deliver against the toughest service-level agreements using parallel, distributed lock-free ingestion and real-time query processing.

Scale Limitlessly: Elastic scale-out architecture with distributed massively parallel data processing delivers consistent, predictable response under high ingest and user concurrency.

Ease of Use and Flexibility: SingleStore™ Brings simplicity and ease to your data processing by allowing OLTP and OLAP workloads to be processed using a single table type.

Drop-in Compatibility: MySQL wire-protocol-compatible, enabling you to plug in directly with existing technologies like Amazon S3, Spark, Kafka and Hadoop.

Customer Case Studies

COMCAST
With SingleStore, Comcast is able to drive real-time operational analytics by ingesting and processing up to **300k events/second**. By using SingleStore to augment Hadoop, Comcast can drive real-time anomaly detection to proactively diagnose potential issues, and reduce technician visits. Their Lambda architecture writes one copy of data to a SingleStore instance and another one to Hadoop.
Read Case Study >

Uber
SingleStore enables Uber to connect **thousands of drivers and passengers in real-time.** Uber relies on SingleStore to power its real-time marketing analytics and customer segmentation. Uber's entire business model is based on real-time data. Because Uber augmented their Hadoop setup SingleStore, the company is able to make real-world decisions based on analytical results in real-time.
Read Case Study >

FORTUNE 50 Leader
A Fortune 50 company is augmenting Hadoop with SingleStore for processing the financial close cycle to provide real-time visibility into their enterprise operations and revenue flow. With SingleStore, the access to **real-time data** gave the C-Suite, asset managers, analysts, and advisers a single source of truth for investor and quarter-end reporting. For the Fortune 50 company, analytics now run against near real-time data, and information that used to take days to arrive now takes just minutes.
Read Case Study >

Tier-1 Wealth Mgmt. Firm
SingleStore helps a Top five North American financial services organization provide interactive wealth management experiences to high net-worth investors. The legacy Hadoop-based architecture was struggling to provide fast and interactive query responses with high concurrency in the face of dramatic or unexpected market events. With SingleStore, they were able to provide nearly **40,000 users** with **10ms response time** and increase access to **historical data by 5X** for better analyses.
Read Case Study >