

Data-Intensive Applications: A C-Suite Guide

Understanding the technology imperatives
at the pinnacle of market transformation



Table of Contents

| | |
|--|---|
| I. Framing the Challenge: The Physics of Data Intensity | 3 |
| Data-Intensive applications enable moments that matter | 3 |
| II. Technology Imperatives: Understanding Data-Intensive Demands | 4 |
| Identifying key requirements | 4 |
| A Modern Database is required | 5 |
| A Simpler Approach is Required | 5 |
| III. SingleStore: Purpose Built for Data-Intensive Applications | 5 |
| Unifying disparities | 6 |
| The magic of universal storage | 6 |
| Eliminating IT friction | 7 |
| IV. Cracking the Code: Doing Data Intensity Right | 8 |
| Uber: Nonstop multi-dimensional optimization | 8 |
| V. Summary: Data-Intensive Applications Can't Wait, Won't Wait | 9 |
| About SingleStore | 9 |

I. Framing the Challenge: The Physics of Data Intensity

It can be hard to grasp the enormous importance of data to today's digital businesses, even as the cleverest succeed by attaching an indispensable service atop a metaphorical mountain of it. Think Uber, Pandora and Hulu, for example. We think we know what these companies do—ridesharing, music and online TV—but do we really?

- **Uber** transformed the taxi market by using complex datasets to instantly match drivers to riders, dynamically price the ride based on demand, and control the quality of the experience. From its launch in San Francisco in 2010, Uber now operates in 10,000 cities worldwide; [93 million](#) people worldwide use the service on a monthly basis.
- **Pandora** replaced CDs and radio with curated playlists drawn from its massive music catalog and compiled by proprietary algorithms. The user simply chooses a music style and is delivered an uninterrupted flow of familiar and captivatingly new artists. Pandora has about [55 million](#) active subscribers.
- **Hulu** makes its vast catalog of TV shows easily accessible by providing users with uncanny recommendations, individualized playlists and a powerful search engine, replacing the tyranny of choice with auto-pilot simplicity. Hulu has roughly [43 million](#) paying subscribers.

All of these companies share a common characteristic: they process an extraordinary amount of data in milliseconds to present, ideally, an irresistible offer. In the moments that customers consider the offer, they will decide to make a purchase or a pass. Based on a fleeting customer experience, a lifetime of loyalty can be won or lost. These are moments that matter, and companies like Uber, Pandora and Hulu have mastered these moments by mastering data.

Data-Intensive applications enable moments that matter

Intensity is the best way to describe these data-fueled moments that matter. In physics, intensity is a measure of power over surface area over time. [Data intensity](#) can similarly be thought of as processing data over a defined period of time—typically far less than in the blink of an eye—under the simultaneous constraints of high volume, complexity, rate of change, concurrency and latency.

Uber, Pandora, Hulu and a growing number of other stratospherically valued digital businesses deliver data-intensive applications that engage, amaze and delight customers in moments that matter. Traditional companies are adopting data-intensive applications, too.

- Retailers are tracking sales in real-time, dynamically pricing and responding in the moment to supply chain delays and disruptions.
- Energy companies and telecom providers process real-time telemetry, weather and supply chain data to predict and mitigate operational risks.
- Transportation, financial services, real estate, hospitality, music, entertainment and healthcare all are being disrupted by a multitude of data-intensive applications.

Across all industries, a simple truth is clear: market transformations are data-intensive. Companies that transform markets stand to win, and win big. This whitepaper explains the technology imperatives of data-intensive applications, and why SingleStore is the modern cloud database purpose-built to meet their exacting requirements.

II. Technology Imperatives: Understanding Data-Intensive Demands

Data-intensive applications are data-hungry, leveraging real-time and historical data. They enable machine learning and drive real-time artificial intelligence (AI) decisioning. Their performance is almost always inherently transactional, with the additional challenge of analytic computation embedded within. Although every business-critical application relies on data, data-intensive applications consume multiple fast-changing data streams, creating data networks. Ultimately, they monetize data.

Data-Intensive applications therefore can be characterized by having any combination of these five characteristics:

1. *Large data sets* of tens to thousands of terabytes that need to be readily available to make decisions
2. Fast-changing streaming data that must be ingested at high rates such as hundreds of thousands or millions of rows per second
3. *Complex queries* which perform analytic functions, joins and calculations over those large, fast-changing datasets
4. *Low latency* with consistent response times of 10 milliseconds or less
5. *High number of concurrent users* or queries, resulting in up to 10,000 or more concurrent queries

Data-intensive applications are not constrained by compute resources, which are readily available in the cloud. Nor are they restricted by storage; the cloud has also solved that. These challenges are data-bound. Thus, the way data is processed emerges as the most determinant technology factor.

Identifying key requirements

The unique data processing mode of data-intensive applications—transactions with simultaneous analytics embedded within—translates into three data platform requirements:

1. *Ultra-fast ingest* to effortlessly collect and import vast quantities of data from multiple internal or external sources: cache, data mart, data warehouse, centralized operational data store (ODS)
2. *Super-low latency* to execute real-time analytics within high transactions, to deliver response times in single-digit milliseconds
3. *High concurrency* to support tens of thousands, or 100,000 or more users in parallel without degrading response times

Additionally, the data platform should have a simple architecture, be easy to get started with and maintain, and be cloud-based to run as efficiently as possible. It must also be cost-effective, running on industry standard hardware if on premises, with predictable cost management over time.

A Modern Database is required

The volume, variety and velocity of information required to fuel data-intensive applications expands every day. A new data stack is needed; its foundation is a new type of database, one that is optimized for response. Analytic databases cannot handle the dynamic nature of the data as it is generated, while transactional databases lack the computational context needed to rationalize complex systems.

A database for data-intensive applications must be optimized for data dynamics and speed-to-insight. It must be a modern database that delivers four necessary attributes. The modern database is:

1. *Relational*: Built on a relational foundation and all of the sophisticated query capabilities it imparts, in industry-standard ANSI structured query language (SQL).
2. *Distributed*: Conceived, designed and built as distributed systems that do not require external middleware to segment, operate and scale.
3. *Cloud-native*: Purpose-built to run on cloud infrastructure for elastic scaling of compute and storage in both hybrid and multi-cloud environments.
4. *Multi-model*: Native support for relational and non-relational data types and data models.

A Simpler Approach is Required

It is common for both large and small companies to attempt to satisfy the requirements of data-intensive applications by using multiple special-purpose databases for a single application. Often the number and variety of underlying databases for the application increases in functionality and use. It often begins with a single transactional database, either a SQL or NoSQL database. As the application becomes more widely used, performance bottlenecks start to arise, such as read requests. The conventional wisdom is that you should add a caching layer to your application to handle this, so you add the second database. Later, you find that new functionality is needed to provide analytics within the application and neither the first transactional database you added, nor the second cache datastore you added performs well for analytics. So, you add a third type of database for analytics. This is database sprawl.

The problem with this approach is that data must be synchronized and replicated between each database. This additional development work increases latency, and often it doesn't completely solve the scaling problem. There is a better way to solve for data-intensity, that doesn't create more database sprawl, and that is to use a database designed to handle all of these requirements in a single database.

III. SingleStore: Purpose Built for Data-Intensive Applications

SingleStore is the single database for all data-intensive applications, purpose-built to incorporate all four Modern Database requirements. It is the industry's premier database for data-intensive applications: a distributed, highly scalable SQL database built for the cloud that can run there, on premises or in hybrid environments. SingleStore delivers maximum performance that can simultaneously handle both transactional and analytic workloads with multiple models, including relational. Based on these capabilities, SingleStore delivers important benefits:

Exceedingly fast analytics: SingleStore can analyze more than one trillion database rows per second, accomplished with parallel streaming data ingestion at rates of one terabyte per second. This enables low-latency, single-digit millisecond responses while scaling to handle tens of thousands, or up to hundreds of thousands of concurrent users.

Superior price-performance: SingleStore is designed for distributed environments, scaling out to meet demand with no external middleware. It delivers a minimum of 10 times the performance of legacy databases at one-third of the cost, for a return on investment (ROI) of 289 percent according to a recent [Forrester Total Economic Impact](#) (TEI) report.

Familiar and flexible SQL: SingleStore is SQL compatible, allowing it to directly replace any legacy SQL database without friction. Its MySQL wire-protocol compatibility facilitates quick adoption and go-live production. Powerful, industry-standard relational SQL queries are supported natively.

Unifying disparities

In addition, SingleStore extends the core attributes of the modern database to uniquely unify database disparities in three ways:

Data types: SingleStore unifies all data types in a single multi-model database, supporting relational, document, key-value, time-series, geospatial, full text and streaming data—essential flexibility for data-hungry applications.

Data models: SingleStore brings together multiple data models to create a multi-model environment. It can house a relational data model, document data model, key value model and more, using SQL to join data across the various types.

Data storage: SingleStore unifies traditional database row stores and column stores into one type of storage, an industry first called universal storage in which a row store contains a column store within. This design breakthrough allows transactions and analytics, processing tasks that were previously mutually exclusive, to be executed simultaneously on the same data.

The magic of universal storage

[Rick van der Lans](#), database consultant at R20/Consultancy and influential business intelligence analyst, provides [further insight](#) on universal storage, SingleStore's "secret sauce." He explains, "The magic of the SingleStore product is that it has a new table type called universal storage, a table that combines the advantages of a column store with a row store.

"On the outside, [universal storage] looks like any other table. But on the inside, if we insert data in this table through transactions, it will be stored in the row store part of the table, which is in memory and which supports [skip lists](#). It's really, really fast!"

van der Lans continues, "Eventually, data moves automatically from the row store to the column store. If you want to do analytics, you'll organize your data in a column-oriented format. And that's exactly what the column store is. So, the table type called universal storage combines the strengths

of both; it allows us to run analytics on tables, on which we can also do transactions. That is a very, very unique feature that SingleStore offers...”

He concludes, “As we’re getting more and more into the world of real-time analytics, it’s important that the data architectures become much simpler, and therefore much more agile. That’s important for organizations, and that’s what SingleStore delivers.”

Eliminating IT friction

As a Modern Database built for the cloud, SingleStore offers distinct advantages to IT organizations struggling to manage separate transaction, analytic and specialty databases. Many of these databases were built with previous generations of technology; they were architected to run on premises with dedicated hardware and storage, and simply ported over to the cloud. In contrast, SingleStore offers:

- *Cloud ecosystem integration:* SingleStore managed service can be deployed in Amazon AWS, in Google Cloud, and on Microsoft Azure with very tight integrations with those clouds’ multiple components. For example, SingleStore offers daily backup to S3, real-time data ingestion from S3, and the ability to write data back to S3. Google Cloud storage and Azure blob storage are handled identically. Similarly, SingleStore can easily and securely connect to popular data streaming platforms such as Confluent Kafka.
- *Easy to get started with:* SingleStore is wire-compatible with MySQL, affording IT organizations an easy migration to the new database. SingleStore is also extremely easy to scale up or down; the managed service offering can be edited in the SingleStore portal in minutes.

SingleStore offers a wide range of [documentation](#) to help IT organizations easily connect the database with cloud data sources and cloud application development frameworks. This allows data-intensive applications to be quickly connected with SingleStore, up and running in a short period of time. SingleStore’s support for a large and growing number of third-party integrations additionally speeds deployment.

- *Pricing:* SingleStore offers attractive managed service pricing for any sized application, from minimally viable product (MVP) stage to enterprise-scale. SingleStore’s S00 tier starts at 69¢ per hour and goes up to hundreds of virtual CPUs, terabytes of memory and very large storage volumes.

For enterprise customers with data-intensive applications, a managed service fee of \$1,500 per month can serve 50,000-plus customers, and ingest and analyze more than 200 million unique events, while delivering a hands-off database experience for IT organizations.

IV. Cracking the Code: Doing Data Intensity Right

SingleStore powers many of the digital companies that are delighting customers and transforming markets, meeting their data platform requirements for:

1. *Ultra-fast ingest*: For example, [Akamai](#) executes 12 million upserts per second in its billing application. [Comcast](#)'s data-intensive content recommendation application processes 300,000 events per second, presenting viewers curated options in real-time.
2. *Super-low latency*: A [major US bank](#) scans credit card transactions “on the swipe,” in less than 50 milliseconds, to detect and protect customers from potentially fraudulent transactions. [Uber](#) meets a two-millisecond service level agreement (SLA) for market segmentation, a key component of its surge pricing models.
3. *High concurrency* Royal Bank of Canada keeps customer wait times low by processing 100,000 queries per second in its wealth management application. Uber's market segmentation application handles surges of 40,000 concurrent users per second within the two-millisecond SLA.

Uber: Nonstop multi-dimensional optimization

Uber's surge pricing is one of the world's preeminent examples of digital market transformation, bearing further discussion. This complex analytic process is the ultimate expression of supply and demand: How much is a customer willing to pay for a ride during the busiest times? What are the limits of surge pricing—at what point does the customer experience degrade to negatively impact loyalty?

SingleStore's data platform is the foundation of this data-intensive application, which delivers geospatial insights and processes intense analytics on each transaction, at scale. When a customer requests a ride during surge periods, such as New Year's Eve, Uber presents a real-time surge price within milliseconds, factoring in:

- The quantity and proximity of vehicles in a given location
- The number of customers in a location requesting rides to other locations
- The number of customers requesting rides in the destination locations
- The origin and destination of concurrent requests from the rideshare company's related delivery businesses
- Historical location-based pricing
- Predictive analytics as to the price elasticity of customers in that location, at that moment.

SingleStore delivers all of these data points in an instantaneous, massive disbursement, enabling Uber to continuously optimize surge pricing. It is an example of how data-intensive applications are at the heart of market transformation—and illustrates the critical importance of choosing the right data platform.

V. Summary: Data-Intensive Applications Can't Wait, Won't Wait

Cloud-based, data-intensive applications are pushing the boundaries of technology, and business itself, with their need to access and generate enormous amounts of fast-moving data. They must be able to ingest and process large quantities of data, from diverse sources, streaming in real-time, to answer complex questions at single-digit millisecond speeds. Users are intimately involved with data-intensive applications; crucial positive customer experiences hinge on data being highly available, consistent and secure—all within the same single-digit millisecond window.

SingleStore is the industry's only modern database built to meet the requirements of data-intensive applications. Distributed, relational, multi-model and built for the cloud, SingleStore provides the necessary data infrastructure to scale seamlessly as user numbers grow.

Enterprises increasingly rely on data-intensive applications for “can't wait” analytics and dashboards, and every day, market-transforming disruptors are building consumer-facing businesses that deliver irresistible “won't wait” experiences. Across the full spectrum of data-intensive applications, SingleStore uniquely conquers the physics of data intensity: the industry's premier modern database processes vast quantities of data in microscopic periods of time, under the simultaneous constraints of high volume, complexity, rate of change, concurrency and latency.

For more information visit www.singlestore.com.

About SingleStore

SingleStore is the first modern cloud database designed for data-intensive applications. From digital service leaders like Uber and IEX Cloud to Tier 1 banks, leading enterprises are adopting SingleStore to deliver the moments that matter. SingleStore unifies all data by combining transactions and analytics into one powerfully simple, modern cloud database which delivers 10x-100x performance at 1/3 the cost. SingleStore is available on all the leading clouds as SaaS, hybrid and on-premises deployments through SingleStore Managed Service and SingleStore DB. Follow us [@SingleStoreDB](https://twitter.com/SingleStoreDB) or visit at www.singlestore.com

