SingleStore

WHITEPAPER

# Moments that Matter: How the Modern Database Powers Data-Intensive Applications

## Uniquely enabled transformative, data-driven digital experiences

# Table of Contents

## I. Data Is the Essence of Today's Moments that Matter

It's a singular moment that happens millions of times per day—the instant in which a rideshare customer considers an offer to get to where they want to go, at a specific time, in a certain type of vehicle and at an exact price point. For the world's leading rideshare provider, each of these moments is an opportunity to win, or perhaps forever lose, a customer. Each moment is an essential decision point that determines a purchase or a pass, and make-or-break customer loyalty.

Collectively, these experiences are called moments that matter. Across industries, whether conveying ride-share options, curated content recommendations, in-game advertising or myriad other types of digital consumption, the most fruitful moments that matter invariably share a common characteristic: they are formulated blindingly fast, with vast inputs of the most relevant, freshest data possible.

Operationally, these moments similarly share another characteristic: they are enabled by data-intensive applications, critical underlying technology characterized by:

- Large data sets
- Large, fast-changing datasets
- Complex queries
- Extremely low latency requirements

### Data-intensive applications demand a Modern Database

Potent as they are, data-intensive applications present multiple back-end challenges. To put the attributes above into an operational context, data volume and complexity are on the rise, taxing current processing structures. Meanwhile, delivering the best customer experiences requires real-time analytics, adding another voluminous dimension of complexity. At the same time, data-hungry artificial intelligence (AI) and machine learning (ML) models are increasingly applied in real-time to predict outcomes and recommend products, layering on even more performance pressure.

The upshot? Powering these moments that matter, the lifeblood of data-driven businesses, requires essential database that are powered by a modern database. A modern database is:

- Distributed
- Cloud native
- Multi-model
- Relational

This paper explores the evolution of database technology and why Modern Database architectures are uniquely capable of powering moments that matter. Spanning diverse industries such as ride-sharing, entertainment, digital retail, finance and many more, moments that matter are the defining characteristic of the world's most valuable companies.

Purpose-built for data-intensive applications, SingleStore is the Modern Database helping a large and growing list of industry leaders to win and keep customers in the moments that matter—the ultimate prize in the decades-long drive toward digital transformation and data-driven businesses.

## II. The Path to Modernity: A Short History of Cloud Data Systems

Modern Database technology isn't completely new or, by extension, unproven. Its architecture is Darwinian, the product of an ongoing, long-term evolution in which the strongest of the species survives and thrives.

### The first era: Retrofitting general purpose technology

In what can be considered the primordial era of cloud data systems—1999 to 2010, the end of the dot-com era and the birth of the commercial web—businesses were scaling their commercial operations on the internet. Thousands of online visitors rapidly grew into millions. Traditional on-premises databases were pressed into service, but even enterprise-class databases became a bottleneck for throughput.

In an effort to scale out to accommodate high user loads, single-node relational databases using structured query language (SQL) such as Oracle, but mostly inexpensive open source MySQL, were retrofitted for the web. They were replicated into multiple instances and split horizontally into multiple partitions. As visitors were routed to these different replicas and partitions, the overall system could tolerate higher numbers of concurrent users and performance improved. In the first era of cloud data systems it became apparent that milliseconds matter; in 2009, Amazon found that every 100 milliseconds of latency on its website cost the company one percent of sales.

But database replication and partitioning were fundamentally flawed strategies for high-intensity web properties. With this approach, which is viable in certain applications and is still in use today, sharding middleware coordinates database reads, writes and changes across the extended environment. In high-volume web-scale applications, data quality unfortunately deteriorates. Limited by the simple wire speeds of on-premises networks and the read/write capabilities of physical disk storage, data changes can't be propagated quickly enough, resulting in dirty reads. Meanwhile, the failure of a single replica or partition can jeopardize the integrity of the entire database.

### The high price of poor data integrity

For a business such as a meal delivery service, for example, this inconsistent data can have devastating consequences. What if, during a peak period, customers ordered a special offer listed as available when in fact it sold out? Customer satisfaction would plummet, along with star ratings from angry reviewers.

Thus, general-purpose SQL could not scale cost-effectively for web-scale requirements. The other logical option—a single instance of a single-node SQL database running on a single machine—was cost-prohibitive. This configuration was available only as a database hardware appliance offered by a handful of vendors, a solution beyond the reach of all but the richest of IT organizations.

### The second era: Specialized databases

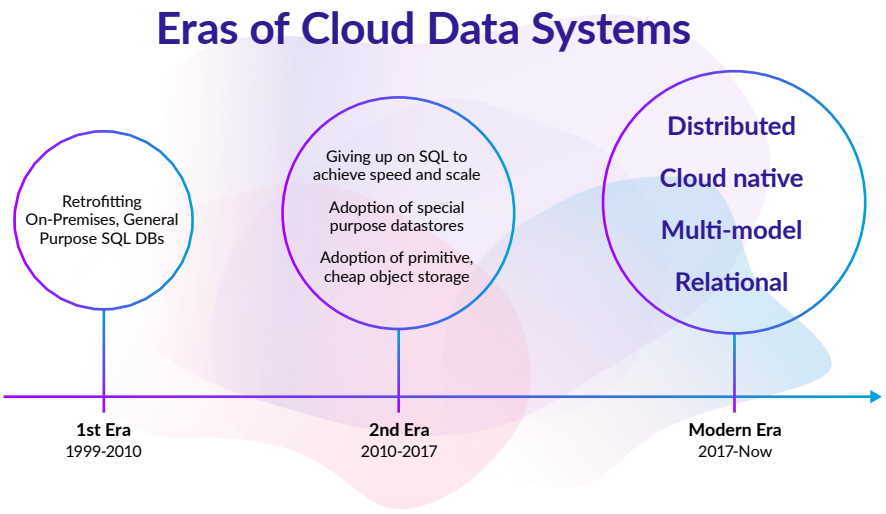
Between 2010 and 2017, web-scale businesses adopted a second era of cloud data systems, abandoning the power of SQL in favor of specialized databases. Dozens of new datastores sprang up, offering speed and scale in a single instance, without sharding middleware.

There was a steep trade-off, though. While these new architectures were self-contained and highly elastic, they were non-relational, forfeiting SQL's declarative query powers. Speed and scale were readily attainable, but second-generation datastores left behind SQL's ability to do joins, aggregations and analytics within a query. Instead, instructions on where to find the data (its physical location in the database) had to be spelled out, introducing another burdensome category of complexity to the cloud database challenge.

## Data moves to the cloud

The rise of primitive, inexpensive object storage mitigated some of the pain. In parallel with specialty NoSQL database adoption, big data projects proliferated on the open source distributed processing framework Hadoop. IT organizations found that Hadoop workloads could be readily shifted from disk storage in the data center to cloud storage, which offered limitless elasticity and scalability at lower costs. During the second era of cloud data systems, enterprise data lakes migrated at a brisk pace from on-premises to the cloud.

First- and second-era cloud data stores are still viable, in wide production and are well-suited for many applications. Data-intensive applications are not one of them; here, multiple special-purpose databases must be stitched together to meet speed, scale and customer experience requirements, albeit at significant complexity, inefficiency and cost. Still, no combination of first- and second-era databases can deliver the robust foundation necessary for data-intensive applications: a cloud-native database built upon a distributed, relational architecture that also accommodates multiple database models.




### Consequences of 1st and 2nd Era Data Systems

**Business impact**

- Poor customer experience
- Poor ratings and reviews
- Increased customer churn

**IT impact**

- Increased data movement
- Incessant data duplication
- Greater system complexity
- Increased maintenance
- Higher costs

## III. Defining the Modern Database

With data-intensive applications now the backbone of highly desirable digital businesses, multi-database patchworks are ineffective in handling their large data sets, enormous numbers of concurrent users and complex queries at extremely low latency. A single database, a true cloud data system, is imperative. This Modern Database is:

- Distributed: Conceived, designed and built as distributed systems that do not require external middleware to shard and scale.
- Cloud-native: Purpose-built to run on cloud infrastructure for innately elastic scaling of compute and storage in both hybrid and multi-cloud environments.
- Multi-model: Native support for relational and non-relational data types and data models.
- Relational: Built on a relational foundation and all of the sophisticated query capabilities it imparts.

## IV. Modern Databases Power Real-World Moments that Matter

Scores of leading companies are creating new markets with digitally transformative applications. Here are three examples of how Modern Databases power data-intensive applications—and the moments that matter they impart.

### Rideshare company: Nonstop multi-dimensional optimization

The Modern Database is a critical component of a premier rideshare company's surge pricing, a complex analytic process that is the ultimate expression of supply and demand: How much is a customer willing to pay for a ride during the busiest times? In fueling this moment that matters, the Modern Database delivers geospatial insights and processes intense analytics on each transaction, at scale. The real-time surge price is presented to the customer 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.

The Modern Database delivers all of these data points in an instantaneous, massive disbursement, enabling the rideshare company to continuously optimize surge pricing. It's a marquee example of the dynamic customer experience that, when done right, wins and keeps customer loyalty in moments that matter.

## Streaming media: Contextual recommendations

Entertainment lovers around the world look forward to the hour or two they may be able to set aside in their day to experience the joy of discovering something new. For streaming content providers such as Comcast, Netflix, Amazon and Hulu, content recommendations are pivotal moments that matter.

"Let's face it, when we're looking at TV and trying to find a show to watch, we really want to find something within ten to 30 seconds," says Mark Hashimoto, software engineering manager at WhatsApp and a veteran of Comcast's Silicon Valley Innovation Center. Multiplying that dynamic by billions of people across the globe, searching for entertainment, quickly conjures a picture of the unmitigated data analytics challenge streaming content providers face.

"The consumer is faced with the challenge of finding great content that's hopefully personalized to them, and will resonate with them," Mark elaborated. "The first part of the discovery portion—presenting curated content options—is really data intensive. It's based on what that individual has watched or what their interests are. From the streaming media company's perspective, the challenge is how to provide those pure magical moments in the moment, in the user journey, to say, 'Here's a great show that we think will really captivate you, and that you should watch.'"

As with the surge pricing, content recommendation moments that matter are in constant flux, typically occurring in waves as evening hours wash over the world. An extraordinary amount of analytics is poured into each milliseconds-long calculation, which factors in:

- What the customer has viewed in the past, and when.
- The viewer's stated content genre preferences.
- Which content in the provider's catalog bears the right level of similarity or adjacency to previously watched content.
- The viewer's persona (adult, teen, kid or young child) and any viewing restrictions.

Modern Database technology powers a leading streaming provider's content recommendation capability, helping to win each opportunity to retain a customer by offering the right recommendation at the right time, anywhere in the world.

## Cloud gaming: In-game advertising

Popular digital games such as Fortnite are the newest frontier for native advertising, presenting players with opportunities to interact with brands in-game. For the companies creating this entertainment, in-game ads are moments that matter, a chance to engage or potentially alienate the customer. Approximately 70 percent of US consumers are gaming, affirming this pastime's status as mainstream entertainment.

The revenue potential is similarly enormous. Gaming is predicted to reach more than 2.8 billion people globally in 2021 (one-third of the world's population) and gamers are predicted to generate over $189.3 billion in revenue. That's more than the American sports industry ($75 billion) and the global film industry ($100 billion) combined.

Unlike streaming media content recommendations, in-game advertising is often less apparent, presented in more subtle ways. Here, the emphasis is more often on monetizing the game, not the gamer, with enhancements such as billboards in open world games, or hoardings at virtual sports stadiums, adding to the realistic features of the environment without breaking the immersive experience.

Games are increasingly delivered via cloud technologies, providing an expansive metaverse in which brand affinity can be built with surgical precision. Modern Databases are at the heart of numerous premier brands to deliver personalized gaming experiences at scale, to engage and reward consumers in a non-intrusive way.

## V. SingleStore: The industry's premier Modern Database

SingleStore is the single database for all data-intensive applications, purpose-built to incorporate all four Modern Database requirements. As such, it is the industry's premier database for data-intensive applications: a distributed, highly scalable SQL database that can run in the cloud, on premises or in hybrid environments, delivering maximum performance for 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 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.

### SingleStore's secret sauce

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.
- 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.

Rick van der Lans, database consultant at R20/Consultancy and influential business intelligence analyst, provides insightful commentary on SingleStore's "secret sauce":

- A "single store" is a row store. It's there to support transactions. With row stores, that means the data is kept in memory, so the transactions are also processed in memory. They are kept in memory in a record-oriented format, which is still the best. SingleStore uses skip lists, a construct almost designed for transactional environments.
- SingleStore offers strong consistency. If two users access the same data at the same time, they'll see the same result. There's fast data injection through pipelines. If we want to get into streaming applications, SingleStore is built for the cloud, to support a massive transactional workload. It's also built to support a massive analytical query workload.
- Of course, SingleStore has all the advanced query optimizations or techniques that you would expect from an analytical database. But in addition to that, it compiles queries. The advantage of compiling queries is that when users enter similar queries, the optimization doesn't have to take place again; the optimization, the column, the plan, or the query strategy is already there. That speeds up certain queries: IT support reference tables, for example, speed up the parallelization of join processing. SingleStore also supports geospatial data types and geospatial functions; more users are not only interested in knowing who bought what, but where they bought it.

## The magic of Universal Storage

"But the magic of the SingleStore product," Rick explained, "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, it 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!"

He continued, "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… As we're getting more and more into the world of real-time analytics, what's important is that the data architectures become much simpler, and therefore much more agile. That's important for organizations, and that's what SingleStore delivers."

## VI. Conclusion: From Here to Ubiquity

Data-intensive applications like ridesharing, streaming content and gaming are the zenith of digital transformation, a now-quaint term that has evolved far beyond digitizing existing ways of doing business. Transactional and analytic elements are the essence of the most successful digital products, and can be delivered only by Modern Databases that are distributed, relational, multi-model and cloud native.

SingleStore is today's premier Modern Database, poised to become companies' ubiquitous choice as they seek to create and sustain strong customer loyalty in moments that matter. For both challenger companies looking to disrupt the status quo, as well as established companies defending against competitive disintermediation, SingleStore uniquely delivers the ideal cloud data system to rule a data-intensive world.

## 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 ⅓ 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 or visit at www.singlestore.com.