Unlocking The Business Benefits of Speedb Data Engine

Unlocking The Business Benefits of Speedb Data Engine

Companies are making massive investments in advanced data-intensive applications to optimize operations, create new digital products and services, and achieve competitive advantages. However, the limitations of existing data architectures may inhibit them from realizing the business benefits of technology innovation.

Organizations today are dependent on their data to operate. Across all industries, the ability to analyze overwhelming volumes of historical and real-time data to improve decision making, achieve operational efficiencies, and much more, is an integral part of everyday business life. However, that reliance on data creates a vulnerability— any malfunction or performance degradation in the underlying data infrastructure may have a major impact on mission-critical applications and services, and consequently — on the business.

The need to handle unprecedented amounts of data creates new challenges in our new data-driven world. Existing data architectures are based on technologies and concepts that were not designed to meet the performance and scale requirements of modern, data-driven businesses. As data volumes continue to grow, these inherent limitations are increasingly manifested in reduced application performance, slower response times, crashes, and other issues, which may significantly compromise business operations and customer experience.
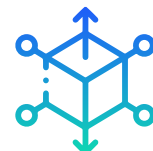
Business Benefits

Optimize performance for user operations and improve customer experience by eliminating user-side stalls and minimizing user latency

Transform your data operations without making any changes to your existing infrastructure and code

Scale infinitely to meet the data volume and throughout requirements of today and tomorrow

Maximize the utilization of existing resources and significantly reduce hardware costs

Metadata Growth Pains

This explosion of data is driving the wave of digital transformation sweeping across every industry. This new wave of data allows organizations to leverage modern, data-intensive applications and rich media content that generate huge amounts of unstructured data in the form of objects such as images, videos, and audio files.

The increase in size and quantity of objects is accompanied by a tsunami of metadata. Essentially, metadata is "data about data," which allows us to quickly find data files by identifying certain properties. The growth of metadata associated with objects is becoming a major challenge as an increasingly large number of objects that may only be a few bytes large may now be holding a metadata of about the same size, and sometimes even more.

The expansion of metadata, which used to be stored in memory behind the scenes, presents a growing problem: where to store it and how to manage it effectively. Unless these issues are adequately addressed, the systems that hold the metadata will eventually impact business operations and performance.

Business Benefits

Free developers to focus on delivering real business value

Support services and bespoke customization to address use-case specific requirements

Significantly reduce storage and networking costs

Focus on your business logic and not on the underlying infrastructure logic

Why You Should Care About Metadata Explosion

In the IT world, performance and scalability issues are traditionally addressed by adding more compute resources and/or implementing solutions to monitor and optimize the different layers of the stack, e.g., systems, applications, and databases.

This approach cannot be effectively applied when dealing with metadata. Metadata is often managed using key-value stores (KVS) like RocksDB, which offer more flexibility and speed compared with relational databases. KVSs, like other types of databases, rely on a storage engine, AKA data engine, which is the part of the software stack that sorts and indexes data. However, existing data engines have limited capacity, high CPU utilization, and significant memory consumption, which impacts their performance.

Given these shortcomings and the rapid pace of metadata growth, the practice of adding more resources will result in escalating costs that cannot be sustained for long. As metadata accumulates, the shortcomings of existing data engines become apparent and lead to an understanding that metadata performance and scalability challenges can only be adequately addressed by redesigning of the underlying data architecture.
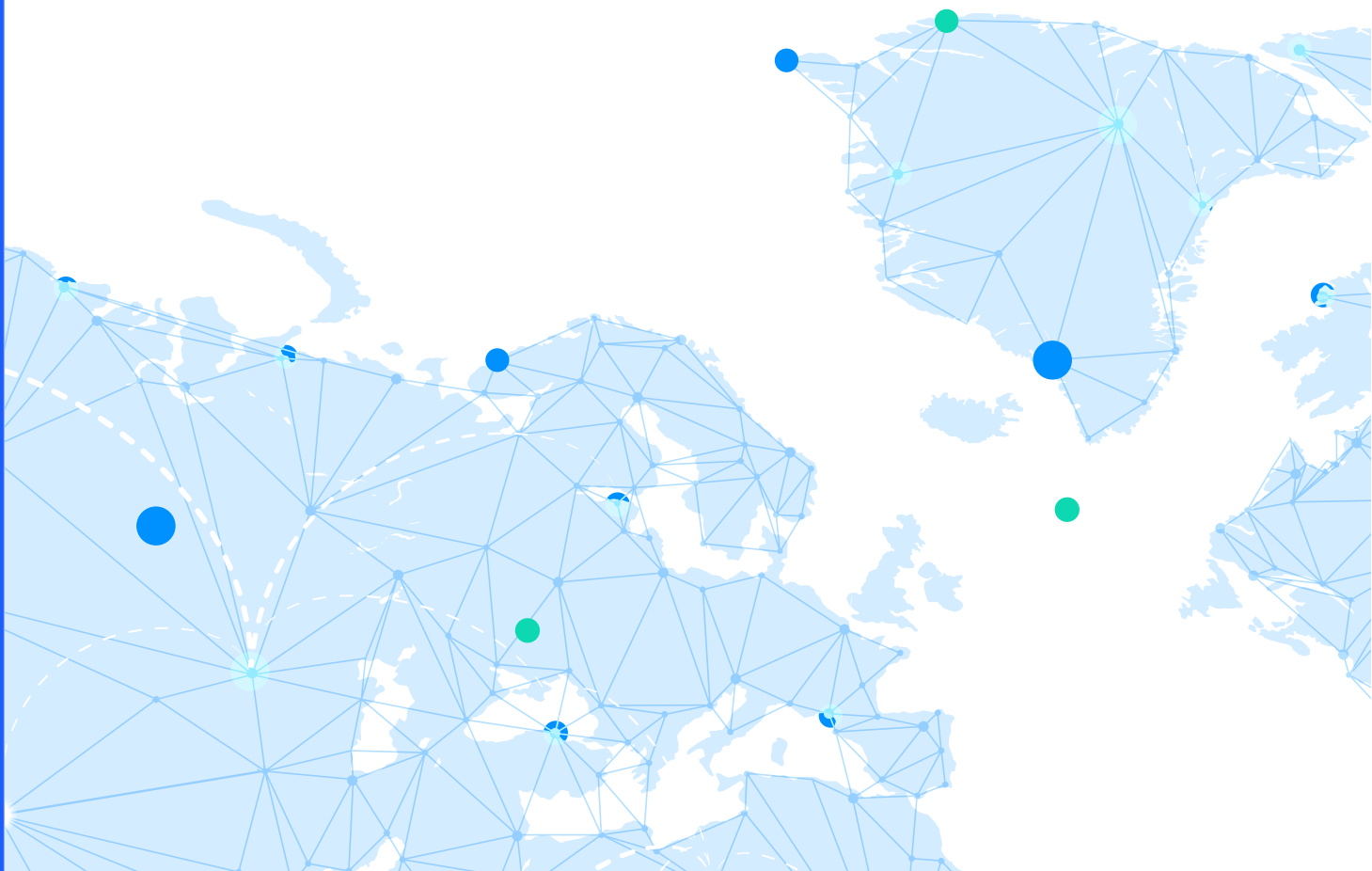
The Business Impact of Data Engine Limitations

Without an adequate alternative, organizations are struggling to balance the conflicting demands of delivering high-performance services at scale while keeping cost and resource utilization to a minimum. Due to the limitations of existing data engines, these factors often trade-off against each other, causing organizations to sacrifice performance for scale or vice versa. This could be extremely risky in a world where quality of service and customer experience are key to cultivating brand loyalty and remaining competitive.

While the degraded performance and limited scalability of existing data engines have direct and immediate impact on customer satisfaction, there are other, more indirect effects that may further compromise business performance. Most notably, tuning and configuring a data engine like RocksDB to match specific performance and scalability requirements is a complex undertaking that even skilled developers may struggle with. As a result, developers spend a significant amount of time carrying out related operational tasks such as sharding, which is the process of splitting up larger databases into smaller components that are easier to manage and thus improve the overall application performance.

Sharding is just one example of the operational tasks that are required to tackle data engine scalability and performance issues, keeping developers from focusing on delivering business value. And in cases where organizations lack the required technical expertise to optimize their data engines, they will likely use default settings that are usually not ideal for their unique needs and therefore cannot effectively support their operational and business objectives related to data.

Technology Highlights

Multi-dimensional compaction that reduces the write amplification factor (WAF) from ~30 to ~5

Fast writes even on large datasets while keeping a B-Tree like read performance

Probabilistic index that consumes less than 3 bytes per object, enabling indexes for large datasets to reside in DRAM

Mapping of hundreds of billions of objects while ensuring nearly one media access per read

Speedb - Changing the Equation

Due to the limitations of existing data architectures, it doesn't take too long before organizations reach a point where they start trading-off between performance, capacity, and cost. These compromises prevent them from reaping the benefits of their investments in advanced technologies.

To address the needs of modern, data-driven organizations, data infrastructures must evolve accordingly and change the equation where improved performance comes at the expense of scalability or vice versa. This can only be done by redesigning our data architectures at their core, starting from the data engine.

Speedb aims to do just that. Founded by a team of seasoned data and storage veterans, Speedb offers a drop-in replacement embeddable solution for RocksDB tailored to hyperscale data processing and management needs. Without changing a single line of code, Speedb enables businesses to dramatically improve the performance, scalability, and the resource utilization of the data infrastructure. Using Speedb, organizations can reduce management overhead by allowing applications to run on a single database, freeing developers from having to constantly deal with sharding, database tuning and other time-consuming operational tasks.

To accomplish that, Speedb has developed a revolutionary data engine that supports petabyte scaling of datasets with billions of objects while maintaining high performance and low hardware requirements. The Speedb data engine is based on technological breakthroughs that makes it the ultimate choice for hyperscale data operations that require superb performance. Using Speedb, organizations can finally pursue their digital growth strategies without worrying about the underlying data infrastructure.

Technology Highlights

Hierarchical data map that significantly accelerates the search time of the index

Auto-tuning of the system parameters to ensure high performance for any workload

Redesigned I/O and job schedulers to guarantee stable performance with no stalls