



WHITEPAPER

MemSQL Overview

DATE: August 2018

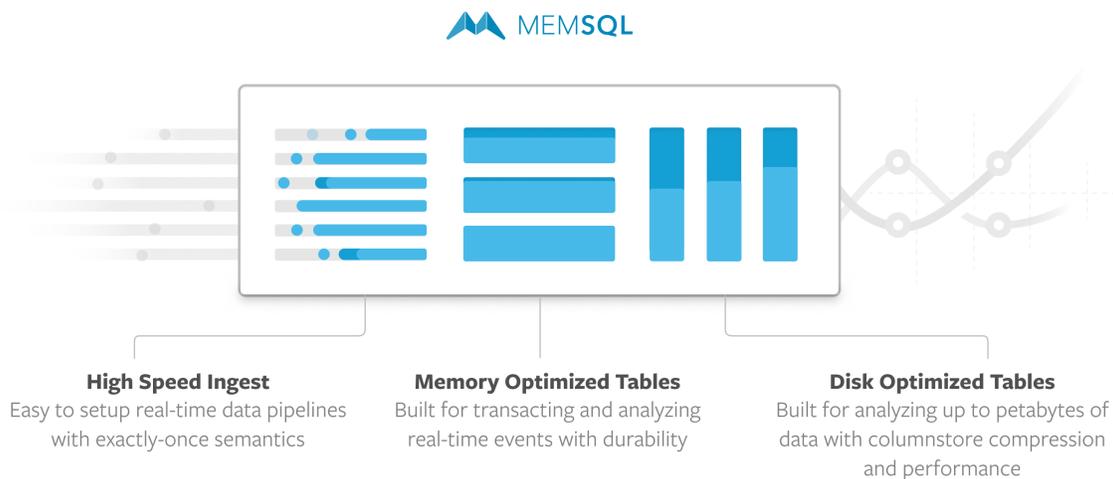
The No-Limits Database™ for Modern Applications and Analytical Systems

MemSQL delivers a fast, scalable, distributed relational data store for handling the demands of modern applications and analytical systems. The platform is ideal for ingesting large volumes of data, have a low tolerance for latency, require high concurrency while maintaining the familiarity of SQL running on commodity hardware.

MemSQL provides a single unified database platform that delivers:

- Data warehouse workloads with disk optimized columnstore tables for added compression and performance
- High performing transactional workloads leveraging memory optimized tables
- Scale-out distributed architecture that runs efficiently on commodity hardware for on-premises or cloud deployments
- Built-in pipeline capabilities for Kafka, S3, HDFS, and filesystems for stream ingestion
- Easy integration with existing data warehouses and applications based on ANSI SQL and ODBC/JDBC

The No-Limits Database™ Explained



MemSQL enables businesses to identify and act to critical events in real-time. This is accomplished by processing transactions and performing analysis simultaneously in a single scalable database with a lock-free data structures. By loading everything into scalable DRAM and leveraging a disk-based columnstore, data can be executed across extremely large datasets in real-time.

With immediate standard SQL access to fast changing and historical data, MemSQL opens new opportunities to accelerate analysis, create new experiences with analytic applications, and serve as a general purpose data store for managing mixed workloads on a single system:

- Capture Real-time data at point of ingest and fuse with other operational data
- Power Operational Analytics
- Accelerate Applications

Sub-Second Transactional Speeds

By leveraging in-memory rowstore tables, MemSQL can serve query results from billions of records at the same moment data is being created. Insert latency is less than a millisecond and has been lab-tested to be faster than any other database for processing transactions and analysis in real-time. Customers routinely use MemSQL to ingest 100k to millions of events per second.

Simple and Scalable Distributed Architecture

MemSQL leverages a distributed computing model that makes it easy to set up, maintain, and scale. Reduce both upfront and long-term maintenance costs with scale out on commodity hardware from your data center to the cloud.

Single Integrated Database Management System



- Multi-cluster manager for easy administration
- Parallel loading for high-performance streaming
- Workload manager for maintaining performance under load
- Distributed ANSI SQL leveraging relational rowstore and columnstore tables
- Scale-out durable storage for easy addressing growing workloads

MemSQL features an in-memory rowstore and a fully-integrated disk and memory based columnstore that eliminates the need for separate databases and enables MemSQL to achieve extremely low latency and execution times while allowing for data growth.

The row and column stores can easily transfer data and perform complex SQL queries with tables in either store, expanding far beyond an in-memory workload and making it extremely powerful in high velocity Big Data environments.

Shared Nothing Architecture

By distributing data across multiple system resources with no single point of failure, MemSQL can leverage thousands of memory cores for extremely fast data loading and analytic workloads.

High Availability

MemSQL is ACID-compliant and distributes multiple copies of data on shared system resources, both within and across data centers. Transactions are always committed to disk and later compressed into full-database snapshots. If a server goes down, data can be recovered in-memory, so queries continue to run with no noticeable degradation.

Deploy Anywhere

Choose from on-premise or cloud deployments that align with your existing infrastructure. With the freedom to choose, you can horizontally scale on commodity hardware without a large up-front investment in your data center or in the cloud.

Easy Integration With Existing Applications

MemSQL is a relational database management system that uses an ANSI SQL interface to access all data types on a flexible schema. While the database is online, MemSQL can alter table schemas, join tables, or manipulate JSON data in a relational environment all at the sub-second speeds of real-time performance. This enables easy integration with existing applications without costly query rewrites or custom connectors. MemSQL supports the MySQL wire protocol providing a near seamless porting of applications with minimal code changes.

MemSQL integrates easily with your existing enterprise data warehouse. Organizations that have already invested in AWS S3, Hadoop or Apache Spark can circumvent the performance issues in a real-time environment by ingesting streaming data and running time-sensitive analytics that traditional data warehouses cannot handle.

