COMMISSIONED BY

ACTIAN™
a division of HCLSoftware

MCKNIGHT CONSULTING GROUP

Edge IoT Database Testing

IoT Database Performance

Actian Zen v16 vs. MongoDB & MySQL

Related Research

Data Stream Processing Ease of Use and TCO, McKnight Consulting Group (2024)

Actian outperforms major competitors in latest TPC-H benchmark

William McKnight

Jake Dolezal

Prepared by MCKNIGHT CONSULTING GROUP

Contents

Executive Summary

Embedded databases seamlessly integrate into software, providing transparent, low-maintenance data management for end-users. The Internet of Things (IoT) proliferation empowers devices through localized database management systems (DBMS).

Embedded databases enable sophisticated applications directly on IoT and remote devices, bypassing client-server architectures that rely on database servers accessed via interfaces. This localized approach reduces latency, enables device-independent operation and optimizes resource utilization.

To fully harness data for competitive advantage, embedded databases require high-performance capabilities for real-time processing at scale. This enables instant data insights, scalable data handling and efficient operations.

High-performance embedded databases unlock real-time insights, fueling data-driven decision-making and business agility. They enhance smart device functionality, optimize edge computing performance and drive data-driven innovation for business success.

Our benchmark was informed by TPCx-IoT. Simulating a realistic real-world scenario and use case for embedded edge IoT databases is one of our benchmark objectives. We used TPCx-Iot to determine the workload for our benchmark. The TPCx-IoT benchmark evaluates the price-performance, scalability, and performance of IoT (Internet of Things) data intake systems. It assesses a system's capacity to effectively manage massive amounts of time-series data by simulating real-time data input and processing from IoT devices.

Actian Zen outperforms competitors in data ingestion rates, reaching up to 7,902 records per second. MongoDB had peak rates of 2,099 records/second. MySQL trails significantly, maxing at 162 records/second.

Actian Zen also consistently demonstrates the lowest ingestion latency across various sensor configurations. MongoDB follows with significantly higher latency, while MySQL exhibits substantially higher latency, up to 92.7 ms.

Demonstrating scalability across sensor configurations, Actian Zen proves ideal for real-time, high-volume data applications. Its superiority highlights traditional databases' limitations, such as MySQL's poor performance. Recommended for IoT, sensor data and high-performance computing, Actian Zen's architecture warrants further optimization research and comparative analysis against other NoSQL databases to explore production environment applications.

Platforms

Actian Zen

Actian Zen is a secure, modular, and highly scalable database supporting NoSQL and SQL Access. The central DBMS architecture consists of two database engines: a Micro Kernel database engine that uses a key-value store for storage and Btrieve API access and a relational database engine that can be accessed through SQL. It supports stand-alone, client-server, peer-to-peer and software-as-a-service architecture and runs across commercially available hardware and virtualized platforms. Being developer-friendly, it’s widely used in embedded applications, edge computing environments and IoT use cases where real-time data processing is critical.

MongoDB

MongoDB is a NoSQL, document-oriented database designed for scalability, flexibility and high performance. It stores data in JSON-like documents, allowing for dynamic schema changes and efficient handling of large amounts of semi-structured data. MongoDB supports horizontal scaling, replication and auto-sharding, ensuring high availability and reliability. Its query language, aggregation framework and indexing capabilities enable efficient data retrieval and analysis. Common use cases include real-time analytics, content management, IoT data storage and mobile apps, making MongoDB a popular choice for modern, data-driven applications.

MySQL

MySQL is a popular, open-source relational database management system for storing, managing and retrieving data. It supports structured query language for querying and manipulating data, offering scalability, reliability, security and flexibility. MySQL ensures data integrity, supports encryption and access control, and is compatible with various operating systems, programming languages and storage engines. It's widely used in web applications, content management systems and data warehousing, providing high performance, ease of use and customization.

Platform Summary

In this IoT database performance study, we tested the following releases of the platforms:

Vendor	**Actian Zen**	**MongoDB**	**MySQL**
Tier	Enterprise	Enterprise	Enterprise
Version	16.00.066	7.0.14*	8.0.40
Synchronization Utility	Easysync	Mongosync	MySQL Replication

*A newer version of MongoDB (8.0) was generally available; however, at the time of this study, mongosync was not yet compatible with version 8.0.

Setup

This section analyzes the methods we used in our IoT performance testing. The benchmark was executed using the following setup, environment, standards, and configurations.

TPCx-IoT like Workload

An aim of the benchmark is to simulate a typical real-world scenario and use case for embedded edge IoT databases. In our benchmark, we derived a workload from TPC Express Benchmark™ IoT (TPCx-Iot). The TPCx-IoT benchmark, developed by the Transaction Processing Performance Council (TPC), measures the performance, scalability, and price-performance of IoT (Internet of Things) data ingestion systems. It simulates real-time data ingestion and processing from IoT devices, evaluating a system's ability to handle large volumes of time-series data efficiently.

However, this is NOT an official TPCx-IoT and the results are not comparable with other published TPCx-Iot results. The TPCx-IoT is an express benchmark that provides a working kit, and the TPC specification is limited to providing high-level information. The official TPCx-IOT kit could not be used to test our platforms of interest.

Thus, we developed our own TPCx-IoT-like benchmark driver using Python for the purposes of this study. To build our driver, we referred to the 2018 IEEE International Conference on Data Engineering (ICDE) paper by Poess *et al*. that describes the TPCx-IoT benchmark in detail. In developing our own driver, we desired to maintain the spirit of the TPCx-IoT, but more importantly, we wanted a fair, objective, “apples-to-apples” comparison of the platforms under test.

AWS EC2 Instances

To simulate our IoT testing, we used pairs of Amazon Web Services EC2 instances as client and server in the same availability for close network proximity. We used the following EC2 instance types:

16-core Virtual Machine	**c5.4xlarge**
Processor	16 vCPU 3.6GHz Intel Xeon Scalable Processors (Cascade Lake)
RAM	32 GB
OS	Ubuntu 22.04 (Jammy Jellyfish)
Disk	EBS gp2

For each client-server pair of EC2 instances, we installed the database software as specified in the previous section.

Test Methods

The test driver was a custom-built Python application. It consists of three operational functions:

- **Insert**: Insert IoT data into the Client database and measure insert throughput and averages
- **Query**: Query IoT data on the Server database and measure query throughput and averages
- **Synchronization Latency**: Query the latest data that has been synced to the Server and measure synchronization latency

We used Python multiprocessing pools to simulate multiple sensors.

The TPC-IoT 2018 IEEE ICDE paper specifies that each IoT sensor reading should consist of a timestamp, sensor identifier, measurement, and padding to bring the record size to 1024 bytes. Thus, the ANSI SQL DDL statement that describes the data is as follows:

```
CREATE TABLE iot (
  ts DOUBLE NOT NULL,
  sensorid VARCHAR(63) NOT NULL,
  measurement DOUBLE NOT NULL,
  padding VARCHAR(943) NOT NULL
);
```

where `ts` is the current UNIX timestamp, `sensorid` is a random UUID assigned to each driver Insert process, `measurement` is a random normal variate with a mean of 150 and standard deviation of 20, and `padding` is a string of 943 random ASCII characters to bring the total bytes per row to 1,024.

For the Insert process, we use the following methods for each platform:

- **Actian Zen Btrieve2 API**: .RecordCreate() method
- **MongoDB**: .insert_one() method
- **MySQL**: SQL INSERT + COMMIT

We then simulated various numbers of sensors (16, 32, 64, 128, 256, 512, and 1,024) each simultaneously producing sensor readings and fed them to the client database for ingestion. We measured the number of records ingested per second.

To send the ingested records to the server database for each platform, we configured each system's native synchronization utility (listed in the previous section).

On the server side, we also implemented the analytical queries used in the TPCx-IoT. These queries were run concurrent with the ingest operations in order to simulate a live and active system. The queries are filtered on a random five second interval (e.g., SQL BETWEEN *x – 5* AND *x*, where *x* is a random UNIX timestamp within the testing time interval). There are four queries, which represent typical dash-board-like queries:

- *Maximum Reading*: compares the maximum sensor measurement (e.g., SQL MAX()) within the interval.
- *Minimum Reading*: compares the minimum sensor measurement (e.g., SQL MIN()) within the interval.
- *Average Reading*: compares the average sensor measurement (e.g., SQL AVG()) within the interval.
- *Reading Count*: compares the number of sensor measurement (e.g., SQL COUNT()) within the interval.

We ran each platform at each sensor count for two hours to ensure the system could keep up its performance over time.

Test Measurements

Our benchmark collected two statistics.

- **Sensor reading throughput –** Measured as IoT records per second. This is similar to the TPCx-IoT metric IoTps.
- **Ingest latency –** Measured as the average time elapsed for each ingested row. This is not captured as part of the official TPCx-IoT but was measured for it comparison.

We did not track the sync "latency" or the measurement of how far away from "real-time" the data being synced to the server is. "Sync latency" is not a TPCx-IoT measurement, and it is not an apples-to-apples because the ingest rates on each platform. One platform's synchronization may be handling many more records that another.

The timing or number of these queries was also not tracked or measured because one platform may be handling many more records that another putting it as disadvantage. These queries are mostly there to add query "pressure" to the systems under test.

Results

This section analyzes the results of our IoT workload scenario for the platforms we tested.

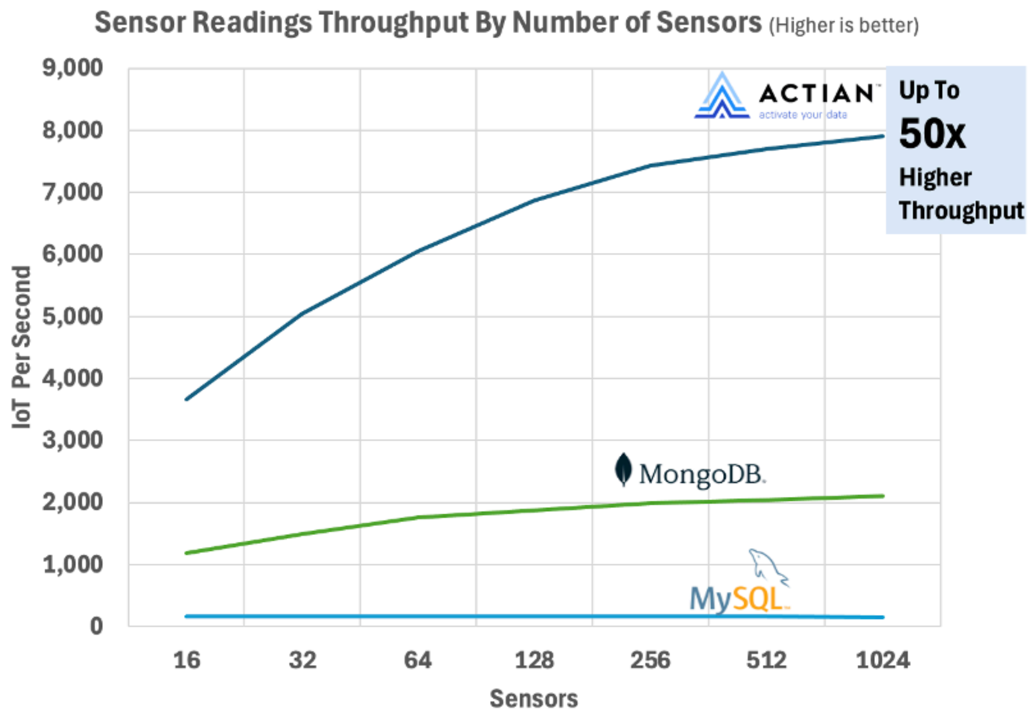

Figure 1: Sensor Readings Throughput by Number of Sensors

Actian Zen, MongoDB, and MySQL demonstrate varying levels of effectiveness in handling IoT sensor data ingestion. Actian Zen consistently outperforms competitors, showcasing scalability and high ingestion rates. Conversely, MySQL exhibits limited scalability and low ingestion rates.

MongoDB's performance plateaus beyond 64 sensors, indicating constraints in handling high-volume IoT data. Actian Zen's superiority stems from optimized data processing and efficient architecture.

Actian Zen achieves up to 50 times higher throughput compared to MySQL, underscoring its suitability for demanding IoT applications. This significant performance gap empowers businesses to harness real-time insights efficiently. By selecting optimal databases like Actian Zen, organizations enhance operational efficiency, foster timely decision-making and establish a competitive edge in IoT-driven markets.

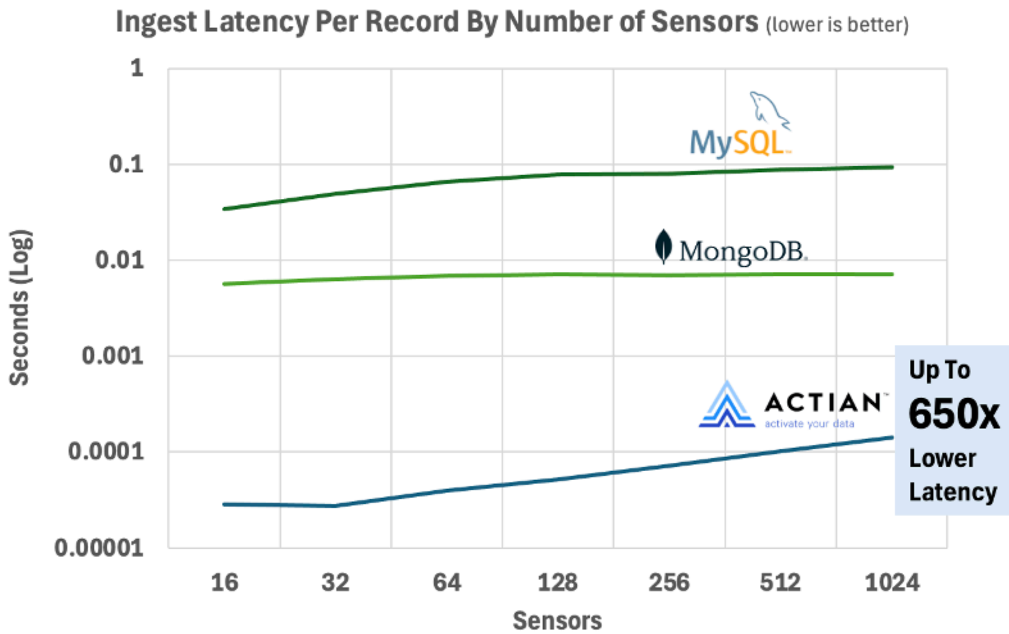

Figure 2: Ingest Latency Per Record by Number of Sensors

The provided data compares ingest latency across Actian Zen, MongoDB, and MySQL databases for varying sensor counts. A key observation is that ingest latency increases with the number of sensors for all platforms. Notably, Actian Zen consistently demonstrates the lowest latency, while MySQL exhibits the highest. MongoDB shows similar performance, with MongoDB's latency stabilizing at 512 sensors.

A closer examination reveals significant performance disparities. Actian Zen maintains remarkably low latency, averaging 3.98132×10^{-5} seconds for 64 sensors. In contrast, MySQL's latency averages 0.066329649 seconds for the same sensor count. This disparity underscores Actian Zen's scalability and efficiency. Furthermore, MongoDB falls between these extremes, emphasizing the importance of database selection for applications requiring swift data ingestion.

The data conclusively demonstrates Actian Zen's superiority in minimizing ingest latency. Remarkably, Actian Zen achieves up to 650 times lower latency compared to MySQL, particularly evident at higher sensor counts. This substantial performance advantage underscores Actian Zen's suitability for real-time data-intensive applications. Optimizing database configuration and exploring data partitioning strategies can further enhance MongoDB’s performance. However, for applications demanding ultra-low latency, Actian Zen emerges as the optimal choice.

Conclusion

Actian Zen can process up to 7,902 records per second, which is the fastest we tested. Peak speeds for MongoDB and an unidentified competitor are 2,099 and 3,554 records/second, respectively. MySQL lags much behind, with a maximum speed of 162 records per second.

Additionally, across a range of sensor configurations, Actian Zen consistently exhibits the lowest ingestion latency. MySQL has much higher latency, up to 92.7 ms, while MongoDB and the unknown opponent follow with far higher latency.

Actian Zen is ideal for real-time, high-volume data applications since it exhibits scalability across sensor configurations. Its excellence draws attention to the shortcomings of traditional databases, like MySQL's subpar performance.

In our evaluation of various embedded databases, Actian Zen emerged as a very compelling solution for enterprise-grade IoT workloads.

About Actian

Actian, the hybrid data management, analytics and integration company, delivers data as a competitive advantage to thousands of customers worldwide. Through the deployment of innovative hybrid data technologies and solutions Actian ensures that business critical systems can transact and integrate at their very best – on premise, in the cloud or both. For more information about Actian Vector and the entire Actian portfolio of hybrid data management, analytics and integration solutions on-premise or in the cloud, visit www.actian.com.

About McKnight Consulting Group

Information Management is all about enabling an organization to have data in the best place to succeed to meet company goals. Mature data practices can integrate an entire organization across all core functions. Proper integration of that data facilitates the flow of information throughout the organization which allows for better decisions – made faster and with fewer errors. In short, well-done data can yield a better run company flush with real-time information... and with less costs.

However, before those benefits can be realized, a company must go through the business transformation of an implementation and systems integration. For many that have been involved in those types of projects in the past – data warehousing, master data, big data, analytics - the path toward a successful implementation and integration can seem never-ending at times and almost unachievable. Not so with McKnight Consulting Group (MCG) as your integration partner, because MCG has successfully implemented data solutions for our clients for over a decade. We understand the critical importance of setting clear, realistic expectations up front and ensuring that time-to- value is achieved quickly.

MCG has helped over 100 clients with analytics, big data, master data management and "all data" strategies and implementations across a variety of industries and worldwide locations. MCG offers flexible implementation methodologies that will fit the deployment model of your choice. The best methodologies, the best talent in the industry and a leadership team committed to client success makes MCG the right choice to help lead your project.

MCG, led by industry leader William McKnight, has deep data experience in a variety of industries that will enable your business to incorporate best practices while implementing leading technology. See www.mcknightcg.com.

Disclaimer

McKnight Consulting Group (MCG) runs all its tests to strict ethical standards. The results of the report are the objective and unbiased results of the application of queries to the simulations described in the report. The report clearly defines the selected criteria and process used to establish the field test. The report also clearly states the data set sizes, the platforms, the methods, etc. that were used. The reader is left to determine for themselves how to qualify the information for their individual needs. The report does not make any claims regarding third-party certification and presents the objective results received from the application of the process to the criteria as described in the report. The report strictly measures performance and does not purport to evaluate other factors that potential customers may find relevant when making a purchase decision. This is a sponsored report. The client chose its configuration, while MCG chose the test, configured the database and testing application, and ran the tests. MCG also chose the most compatible configurations for the other tested platforms. Choosing compatible configurations is subject to judgment. The information necessary to replicate this test is included. Readers are encouraged to compile their own representative configuration and test it for themselves.

MCKNIGHT CONSULTING GROUP

Our services span strategy through implementation for turning information into the asset it needs to be for your organization. We offer a wide range of services related to the important asset of enterprise information.

For Enterprises

- ✓ 2-Week Data Initiative Jumpstart
- ✓ Proof-of-Concepts and Minimum Viable Products
- ✓ Information Management Action Plan
- ✓ Data Platform Selection
- ✓ Data Maturity Assessment and Improvement
- ✓ Data Warehousing and Analytics
- ✓ Big Data
- ✓ Master Data Management, Data Governance and Data Quality
- ✓ Enterprise Business Intelligence Architecture
- ✓ Predictive Analytics Accelerator
- ✓ Compliance Accelerator
- ✓ Artificial Intelligence Accelerator
- ✓ Education and Workshops

For Vendors

- ✓ Benchmark Services
- ✓ Competitive Market Research
- ✓ Engineering Projects
- ✓ Research Development

For Your Audience

- ✓ Keynotes, Webinars, Seminars and Podcasts