database
TRENDS AND APPLICATIONS

DATABASE PERFORMANCE FOR THE MODERN ENTERPRISE: MAXIMIZING SPEED, SCALABILITY, AND AGILITY

database
TRENDS AND APPLICATIONS

Best Practices Series

DATABASE PERFORMANCE FOR THE MODERN ENTERPRISE: MAXIMIZING SPEED, SCALABILITY, AND AGILITY

Best Practices Series

"With data collection, 'the sooner the better' is always the best answer," said Marissa Mayer, former president and CEO of Yahoo. The ability to collect data sooner is the challenge of building and configuring data systems to capture, process, and move data when and where it is needed. The spotlight is on data teams to figure this all out. Are today's organizations ready to deliver?

Ask most data managers and professionals about their greatest challenge to achieving this, and they almost unanimously mention performance. Worse yet, database performance issues continue to take a major bite out of organizational time and resources, according to an upcoming survey report of 218 data professionals conducted by Unisphere Research. Many respondents (62%) say they are impacted by database performance issues monthly or more.

In this day and age, database performance should be a problem that has been licked, right? Not necessarily.

Database vendors have been responding to the challenges faced by their customers, continually expanding their range of offerings.

Although database environments that have been developed over the past few years have all the necessary bells and whistles—buttressed by cloud capacity—they still are trying to catch up with the demands of business users in the year 2025, which encompass AI, real-time capabilities, and supporting edge environments.

Data managers and professionals need to ask these questions: Will data be available the moment it is needed by users? Is the data that is available up-to-date and available to assist decision making or drive enterprise applications?

Database vendors have been responding to the challenges faced by their customers, continually expanding their range of offerings. AI-enabled tools, stepped-up security, and support for a wide variety of data types are now part of the platforms of seasoned vendors as well as next-generation vendors.

Still, new solutions can take time to put in place. Database environments need

Storage systems often can drag down database and application performance, especially if data is making round trips between front-end systems and back-end disks.

to keep up with the speed of business—without the need to support armies of data engineers and administrators to help deliver it. Database performance—in a world where businesses are demanding real-time, sophisticated output—requires rethinking approaches to achieving speed and scalability.

Of course, it's necessary to tend to the long-established P's and Q's of extending basic database performance. This is accomplished through tried-and-true strategies such as data compression, data caching, database tuning, query optimization, managing resource allocation, data partitioning, and sharding.

The following are best practices to employ to move forward in this new environment:

Leverage cloud resources. While there has been some movement away from 100% public cloud services for data management and storage and toward on-prem or hybrid storage, it's useful to have cloud resources up and ready to support scaling as needed. This is especially true due to the relentless growth of AI and machine-learning data, not to mention basic database functions such as backup and restore. A cloud strategy needs to be carefully measured, as monthly access and egress charges can quickly add up as data environments grow. In addition, there is the potential that control of data resources and assets could be handed over to a third party. Still, it's difficult to build a well-functioning data environment with on-prem systems and hardware, along with the staffing to manage the performance of these environments.

Leverage AI and machine learning to boost data operations. It's a two-way street when it comes to AI, machine learning (ML), and databases. While robust database environments are now necessary to support AI and ML implementations, they also can benefit from these technologies as well. Database AI will now automatically handle backup and restore functions; conduct semantic searches across documents, images, and data; and scale as needed.

Look to AI for query optimization. Interestingly, today's AI-enabled data environments take out the complexity of SQL commands, enabling users to employ natural language prompts to get at the data they need. Rather than wrestle with often-complicated SQL code—or have the data department take on the task for users—queries, or more appropriately, prompts, can get users to the data they need simply and quickly.

Store data strategically. Storage systems often can drag down database and application performance, especially if data is making round trips between front-end systems and back-end disks. It's important to understand what data is needed frequently and instantaneously versus occasional access. Managing data lifecycles or retention strategies has often been a hit-or-miss proposition. Data managers are often faced with the need to assess the value of "hot" data that is immediately available on more-costly digital systems or to move it to a less-costly "cold" data repository, which usually consists of low-cost tape storage. AI tools will help track data usage and allocate appropriate storage tiers. AI can also be used to predict future storage trends and allocate space to manage the process.

Keep hardware up-to-date. For on-prem systems, hardware obviously plays a key role in performance—as it has for decades within IT environments. Nowadays, for those environments not relegated to the cloud, the need for efficient hardware is even more critical as organizations seek to grow through real-time and AI capabilities. This may call for stronger chipsets, such as those offered by providers such as NVIDIA, as well as higher-capacity storage to accommodate the various tiers of data mentioned above and to support internal deployments of both small and large language models. It may also help to offload data processing loads to edge environments that are close to data sources and users.

Collaborate. Database performance is just as much a human and organizational effort as it is technical. Dropping expensive hardware and AI-based solutions on top of a static or dysfunctional organization only means dysfunctional processes move faster. Here, adopting collaborative methodologies such as DevOps and DataOps can keep teams and products aligned. DevOps provides for the timely deployment of applications as they move from development to production stages. DataOps, meanwhile, focuses on ensuring that data moves smoothly through the pipeline.

These practices help assure a high-performing data environment that can grow or change as the business grows or changes. These days, database performance is a moving target, especially as organizations move to being data- or AI-driven enterprises. Data managers and professionals must take the lead and deliver the capabilities needed for this new generation. ■

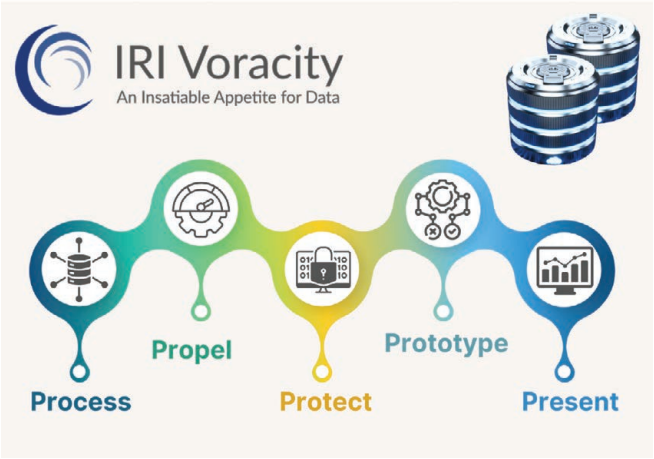
—Joe McKendrick

Speed & Security for Every Database

IRI Voracity
An Insatiable Appetite for Data

Because of the volume of operational data now available in on-premise and cloud databases, and the risks associated with PII, the speed and safety of database operations are more critical than ever. In use cases involving database migration, ETL, data cleansing, testing, and privacy law compliance, IRI Voracity is a proven solution platform for DBAs, GRC teams, and developers.

Powered by the CoSort data transformation engine and featuring mask/test tools like FieldShield and RowGen, Voracity streamlines database DataOps and DevOps. Voracity consolidates data discovery, integration, migration, governance, and analytics. Its combinable tasking reduces job steps, runtimes, and the costs of multiple tools.



MULTIPLE SOLUTIONS

Process

Voracity rapidly performs high-volume structured data processing operations, including: searching, profiling, and classifying data, change data capture/replication, query integration and acceleration, bulk table extraction and pre-sorted loads, offline reorgs, database migrations, external transforms, and cross-platform SQL operations.

Propel

Voracity optimizes the performance of several of the above tasks through task consolidation, state-of-the-art algorithms, and the multi-threading of extraction, classification (PII search), and sorting. Voracity sort/join/aggregation outperforms SQL transforms, ELT appliances, other ETL tools, and even mainframe utilities.

Protect

Voracity governs data through built-in data quality and data masking functions. Validating, cleansing, enriching, and standardizing data improves its utility and reliability. Encryption, pseudonymization, redaction, scrambling, hashing, and anonymization functions de-identify PII in production and test environments for breach nullification and legal compliance.

Prototype

Voracity can also subset and synthesize smart, safe test data for prototyping, DevOps, demos, and benchmarking. Its "RowGen" component parses DDL to rapidly generate and populate—or otherwise provision—pre-sorted, structurally and referentially correct test data for an entire schema.

Present

Voracity can report on data in databases and files as it searches, transforms, cleanses, and/or masks that data. Its embedded BI includes custom detail and summary report formatting with cross-calculation and statistical functions. Alternatively, Voracity users routinely transform and wrangle data to rapidly prepare and feed data to their own analytic, visualization, or AI environments.

USE CASES

Optimizing Database Performance

Organizations in multiple industries have leveraged Voracity to optimize their database performance.
For example, a financial services company used Voracity to accelerate their ETL processes, reducing data processing times by 50%. This improvement enabled them to deliver real-time insights to their clients, enhancing customer satisfaction and competitive advantage.

Ensuring Data Privacy and Compliance

Healthcare providers and financial institutions leverage Voracity data masking and auditing features to comply with HIPAA and PCI rules. By masking sensitive information and providing robust audit logs, IRI customers are protecting patient and customer privacy, and avoiding costly data breaches.

Streamlining Data Migration

A global retail company faced the challenge of migrating their data to a new cloud-based database. With Voracity, they were able to discover, map, and replicate their data efficiently, minimizing downtime and ensuring a seamless transition.

CONCLUSION

The ability to accelerate and govern database-related operations is essential. IRI Voracity is a popular and affordable data management platform that combines high-speed data integration and cleansing with multi-platform test data masking and generation. By leveraging the power of Voracity, you can accelerate your database performance, ensure data quality and compliance, and drive business success.

Learn more at https://iri.com/solutions/database-acceleration and https://iri.com/voracity. ■