database
TRENDS AND APPLICATIONS

DATABASE
PERFORMANCE FOR THE **MODERN ENTERPRISE:**
SPEED, SCALE, AND FLEXIBILITY

database
TRENDS AND APPLICATIONS

Best Practices Series

DATABASE
PERFORMANCE:
SOLUTIONS
FOR SPEED, SCALE, AND FLEXIBILITY

Best Practices Series

Businesses—especially more full-blown digital enterprises—are constantly evolving and changing, sometimes on a daily basis. That means unending stresses and strains on databases, their data environments, and the people who run them. If their businesses are growing at a rapid clip, last month's capacity has already become outdated. If the business took on a new acquisition or merged, it may bring about thousands of new users and gigabytes' worth of new datapoints.

Performance is everything—a data-driven business depends on a high-performing data environment. It's all built around expectations—how quickly data will be available to users, how rapidly it can move between sources and storage, and how efficient data environments run. A high-performing data environment is needed within today's emerging real-time businesses.

The emphasis on data performance was borne out in a recent survey of 217 data managers conducted by

Businesses—especially more full-blown digital enterprises—are constantly evolving and changing, sometimes on a daily basis. That means unending stresses and strains on databases, their data environments, and the people who run them.

Unisphere Research, the research arm of Information Today, Inc. Database performance is the third-ranked top-of-mind concern that takes up the time and resources for database administrators and managers, right behind security and keeping databases up-to-date, the survey found. (Security was cited by 46%, updating was cited by 39%, and performance was cited by 38%.) In addition, availability and uptime were also cited by 37% ("2023 Database Priorities Survey," June 2023).

Furthermore, 67% of data managers in the survey to a large extent agreed

*There are many ways to describe the challenges of
ever-expanding data environments. 'Sisyphean challenge,' 'Gordian knot,'
and 'Rabbit hole' are a few terms shared by Felipe Cardeneti Mendes,
co-author of* Database Performance at Scale: A Practical Guide.

that the number of resources spent on ongoing database management is severely limiting their companies' competitiveness. Solutions most often sought are embedded within cloud architectures: virtualization or cloud solutions (50%), Database-as-a-Service (37%), and Infrastructure-as-a-Service (31%).

There are many ways to describe the challenges of ever-expanding data environments. "Sisyphean challenge," "Gordian knot," and "Rabbit hole" are a few terms shared by Felipe Cardeneti Mendes, co-author of *Database Performance at Scale: A Practical Guide.* "It's complex. And that's just the tip of the iceberg. Then, once you feel like you're finally in a good spot, something changes. Your business experiences 'catastrophic success,' exposing the limitations of your initial approach right when you're entering the spotlight. Maybe market shifts mean that your team is suddenly expected to reduce latency—and reduce costs at the same time, too."

As data managers build a data architecture and data services, Mendes and his co-authors urge that it's important to ask the following questions:

- "How well you know your workload access patterns and whether they are a good fit for your current or target database."
- "How your database interacts with its underlying hardware, and whether your infrastructure is correctly sized for the present as well as the future."
- "How well your database driver understands your database—and how well you understand the internal workings of both."

To deliver to their businesses, data managers and professionals are under relentless pressure to enable greater efficiency, agility, and innovation at organizations in a variety of ways. Data must be available when and where it's needed to support business goals— improving decision making, optimizing processes, increasing productivity, accelerating product development, and boosting customer satisfaction.

Here are some of the factors shaping database performance, identified by Mendes and his co-authors:

- **Workload mix (read/write ratio):** "Some databases shine with read-heavy workloads, others are optimized for write-heavy situations, and some are built to accommodate both. Selecting, or sticking with, one that's a poor fit for your current and future situation will be a significant burden that will be difficult to overcome, no matter how strategically you optimize everything else."
- **Mixed workloads:** "Databases are essentially made for just two things: reading and writing. The way that different databases handle a variety of competing workloads is what truly differentiates one solution from another. Be sure to consider nuances like whether your reads are from cold data (data not often accessed) or hot data (data that's accessed often and likely cached)."
- **Competing workloads (real-time versus batch):** Ideally, databases should "dedicate more resources to the more latency-sensitive workloads to keep them from faltering due to insufficient resources. This is commonly the case when you are attempting to balance OLTP (real-time) workloads, which are user-facing and require low latency responses."
- **Latency expectations:** "There's no simple solution for reducing latency. The lower the latency you need to achieve, the more important it becomes to understand and explore database tradeoffs and internal database optimizations that can help you shave milliseconds or microseconds off latencies. Database internals, driver optimizations, efficient CPU utilization, sufficient RAM, efficient data modeling … everything matters."
- **High-availability expectations:** "To prepare for the worst, start by understanding what your use case and business can tolerate if a node goes down," the co-authors point out. "It's important to note that replication and consistency both come at a cost to performance. Get a good feel for your business's risk tolerance and don't opt for more than your business really needs."

Along with technical adjustments and awareness, the organization needs to be part of database performance approaches. This includes comprehensive training from database professionals to keep them up-to-date with the latest tools and techniques. There are tools and platforms continuously emerging that provide ever-greater intelligence to database practices. Key resources are available from cloud services. Automation—and AI—means issues or changes can be handled almost instantaneously without manual intervention. In addition, executive buy-in will pave the path to high-performing data environments.

Again, in today's enterprises, performance is everything. From AI and machine learning to IoT and real-time analytics, the need for high-performing databases has never been more acute. Enterprises need greater speed, scalability, and flexibility in how data is processed, stored, and accessed. ■

—Joe McKendrick

Speed & Security for Every Database

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.
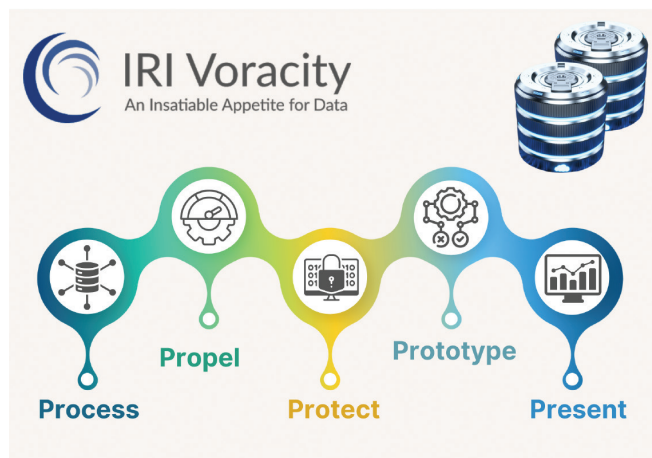
MULTIPLE SOLUTIONS

Process

Voracity processes data in relational and NoDB 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. ∎