

Database  
Performance  
Today: **THE  
NEED FOR  
SPEED &  
SCALE**

# Top Strategies for Improving DATABASE PERFORMANCE

## Best Practices Series

Efficient and agile data environments pave the way for business innovation and growth. A strengthening movement toward data and analytics democratization is opening up data sources and tools to a wide breadth of users. This sense of enlightenment, however, is tempered by the fact that most database teams are struggling just to “keep the lights on,” bogged down by routine maintenance activities. While data environments continue to expand in size and complexity in conjunction with new business demands, the challenge of maintaining the performance and availability of business-critical systems and applications is progressing in step.

Growing data volumes, performance concerns, and tight budgets are some of the challenges faced by data managers today. Close to half of data managers, 49%, report slow application performance as their top challenge, along with cost issues, according to a survey by Unisphere Research, a division of Information Today,

Inc. When it comes down to what counts in database management, it's performance, performance, performance. Forty-one percent of respondents acknowledged spending significant amounts of their week addressing on-premise performance levels, and 24% cited the time spent addressing performance in cloud-based environments (“DBAs Look to the Future: PASS Survey on Trends in Database Administration”).

Emerging technologies are also adding to performance issues. In another Unisphere Research survey, close to half of data managers, 47%, said they are, at some level, concerned about the impact of 5G, streaming, and IoT data straining their infrastructure. To meet these challenges, 28% expect to be running most of their data workloads in public clouds 3 years from now, up from 15% currently (“DBTA Digital Transformation and Cloud Workloads Study”).

From the rise of hybrid and multi-cloud architectures to the impact of machine learning, automation, and

containerization, database management is rife with new opportunities and challenges. Is today's data environment ready for the challenge? The following are the top nine strategies data managers and business leaders need to adopt within their organizations to overcome this growing performance deficit.

### TOP NINE STRATEGIES

**Look at data fabrics/virtualization:** A significant source of performance issues is the wide variety of data now flowing through enterprises from on-premise, cloud, and edge environments. Data tends to end up in silos or within data lakes, requiring further integration work to provide enterprise views. The challenge is to minimize the movement of data. A data fabric or virtualization approach provides layers of functionality that can bring together catalogs, repositories, and other essential components. A virtualization layer or data fabric can provide software-defined services that

*From the rise of **hybrid** and **multi-cloud architectures** to the impact of **machine learning**, **automation**, and **containerization**, **database management** today is rife with **new opportunities** and **challenges**.*

enable data replication. Once put in place, this is a more rapid alternative to pulling and transforming data from multiple sources into a centralized repository or data warehouse-type environment. In addition, a more virtualized approach retains and manages information at, or close to, original locations and platforms. Enabling data to remain with existing assets also helps reduce latency and overhead.

**Move to real-time streaming options:** Most analytics tools and platforms still rely on historical data that is maintained and pulled from either original sources or data warehouses, lakes, or lakehouses. Real-time data streaming—built on environments such as Apache Kafka or Pulsar, which are highly distributed systems that act as enterprise central nervous systems—enables collection of high-volume data consumable by enterprise users.

**Bring in the rights:** Finding people with the skills to manage increasingly complex data environments is one of the greatest hurdles facing today's enterprises, especially as they seek to evolve into data-driven enterprises. While many organizations have legacy IT skills, staff will need to be retrained to manage on-premise-to-cloud migrations, or cloud-to-cloud migrations. Deploying and managing multi-cloud arrangements also calls for new types of skill sets. The processes involved in moving to the cloud, computing, processing edge data, and building the foundation for AI and machine-learning environments demand new types of skills for increasing performance as well as meeting business requirements.

**Rely on automation:** A capability that is increasingly making its way into database operations is robotic data automation (RDA), which employs bots

to manage many data preparation and integration activities. Such automation handles the manually intensive aspects of provisioning and monitoring databases. Automation can also alleviate the bottlenecks that may occur because of version control issues and testing and can also help to free up database staff for more high-level activities.

**Consider database lifecycle management:** Just as application lifecycle management is beneficial for reducing confusion and clutter, adopting a database lifecycle management approach will help move the platform through all phases of adoption and upgrades until its eventual retirement.

**Look at new roles for data warehousing, data lakes, and data lakehouses:** Data warehouses have been robust repositories for analytics-ready data, but often do not have the up-to-the-minute data that a data-driven business may require. Data lakes are places for unstructured data but often have data governance issues. A third approach using data lakehouses combines requirements for analyzing both structured and unstructured data. The data lakehouse architecture offers a compelling option that merges the best qualities of data warehouses and lakes to provide a single solution for all significant data workloads, ranging from SQL analytics to business intelligence, data science, and AI.

**Review cloud options:** Leveraging cloud services can provide rapid on-ramps to increased processing power, storage, and analytical tools, thus preventing any lags resulting from spikes in usage or queries. Cloud services can also provide for automatic balancing and scaling of workloads.

**Ramp up resources:** Of course, not everything is ready for the cloud, particularly data that may be sensitive

or must be confined within particular regions. From a hardware point of view, there are areas that can be expanded to speed up database performance, such as processing and memory. Upgrading processors will help accelerate database performance. In-memory processing now offers a powerful, nano-second option for processing and delivering data right from internal memory, versus swapping data from disks. Storage is also seeing dramatic decreases in latency, thanks to the availability of flash storage.

**Tighten up queries and indexes:** End users are accustomed to the speed by which online services such as Google deliver search results and expect the same within their enterprise systems. Inefficient code or commands may burden SQL-based queries. There are SQL optimization tools on the market that can help streamline a corporate inventory of queries. At the back end, data managers must consider the indexing structure that helps—or, perhaps, hinders—end users' ability to get at the information they are seeking.

## MAINTAINING PEAK PERFORMANCE TO MOVE BUSINESSES FORWARD

Data continues to move across enterprises and through applications at blazing speeds, and this is not going to let up. An enterprise-scale approach to database performance helps address the many pressing challenges with managing complex data environments. It's also important to seek to resolve the critical skills requirements that will ensure that data environments operate at peak performance to move their businesses forward in the digital era. Maintaining security, accessibility, and timeliness requires new tools, platforms, and approaches. ■

—Joe McKendrick



# The Impact of a Data Hub on Performance



## INTRODUCTION

With data being one of the most critical assets for many enterprises in today's world, it is fundamental for businesses to act quickly on delivering accurate and up-to-date data. New usages and opportunities around data bring diverse and complex data architectures involving applications, cloud infrastructure, and database technologies. The increasing data volumes, combined with the growing need for real-time data, makes it difficult for database teams to maintain performance, availability, and to scale when necessary.

## THE DATA HUB APPROACH

With the increasing complexity of data environments and the growing data volumes coming from heterogeneous sources, it is essential to keep enterprise key data under control. A Data Hub is the essential piece for an IT landscape to achieve this objective: it centralizes information across applications and provides services to consistently manage the enterprises' core data (parties, places, and things).

*"With data being one of the most critical assets for many enterprises in today's world, it is fundamental for businesses to act quickly on delivering accurate and up-to-date data."*

Semarchy xDM is the Intelligent Data Hub. This data management solution empowers business users to become data champions. It enables organizations of any size and helps them quickly bring together the critical information scattered across applications into a single Data Hub, with fast time to value. Data can be discovered, mastered, governed, and centrally managed in a non-intrusive way.

By centralizing data management and governance, Semarchy xDM plays a key role in rationalizing IT infrastructures and ultimately optimizing their performances.

## DATA RATIONALIZATION AND QUALITY IMPROVEMENT: A CORNERSTONE OF APPLICATIONS PERFORMANCE

Enterprise key data must be structured and cleaned for an IT landscape to provide the best performance: duplicate, invalid, or inconsistent data artificially increase the workload on IT infrastructures and impacts performance, while making it difficult to achieve business objectives.

Semarchy xDM allows modeling any enterprise core data and designing data model-driven applications to manage and govern enterprises' core data. The Data Hub then becomes the central point where data quality issues are resolved: it validates, cleanses, enriches, and de-duplicates data coming from all sources and provides a central point of access for other applications.

By optimizing and rationalizing core data, Semarchy xDM Intelligent Data Hub contributes to reduce the workload on database and application servers and increase the performance of data environments.

## THE NEED FOR OPTIMIZED DATA APPLICATIONS

Any application that manipulates data must be built on optimized source code to provide optimal performances.

Semarchy xDM is a low-code platform that generates optimized applications, without requiring any manual coding.

Data applications created with Semarchy xDM provide the best performance, even with high volumes of data. Semarchy xDM's engine optimizes performances with built-in caching, multi-threading, batching, and uses in-database processing in order to benefit from database optimizations (such as indexing capabilities and query optimization), while reducing the workload on the application server tier.

## SUPPORT THE DIVERSITY OF DATABASE TECHNOLOGIES, ON-PREMISES AND IN THE CLOUD

A Data Hub must adapt to the diversity of infrastructures. Semarchy xDM is compatible with many database technologies that include Oracle, Microsoft SQL Server, and PostgreSQL, and is available on most popular cloud platforms, such as Microsoft Azure, Amazon Web Services (AWS), and Google Cloud. Combined with its native support of high availability and scalability, Semarchy xDM allows administrators to get the best performance from their infrastructure by choosing the backend database technology and size, as well as configuring and sizing virtual machines with load-balancing and high availability. ■

## 30-day Trial

Ready to start with the Intelligent Data Hub™?

<https://www.semarchy.com/download/>

## SEMARCHY

For more information: <https://www.semarchy.com>