Database Infrastructure Performance Challenges:
Approaches to Better Manage Application Database and Storage Subsystem Performance

A META Group White Paper
# Database Infrastructure Performance Challenges: Approaches to Better Manage Application Database and Storage Subsystem Performance

## Contents

- **Executive Overview** ........................................................................................................ 2
- **Introduction** ..................................................................................................................... 3
- **New Demands on Database Infrastructure** ...................................................................... 3
  - Multi-Tiered Infrastructure ............................................................................................. 3
  - Growing Storage Demands and Options ........................................................................... 6
  - Mixed Workloads and On Demand Scaling ................................................................. 7
- **The Evolving Role of the DBA** ....................................................................................... 7
  - Organized for Success ..................................................................................................... 8
- **Strategic Performance Management** ............................................................................ 9
- **Bottom Line** ..................................................................................................................... 10
Executive Overview

Corporate relational databases now manage the majority of business-critical data within the enterprise. IT organizations face continuing challenges in managing increasingly complex, data-driven application environments.

Through the end of this decade, several factors will converge to challenge the IT organization’s ability to manage its database software infrastructure, and by extension, the critical applications that are dependent on the database. The realization that data has become a corporate asset and has value beyond the recording of a business transaction is resulting in more attention being focused on the importance of improving overall data management.

META Group recommends that several steps be taken to help ensure the proper support for critical applications and specifically for the relational database infrastructure that services them:

- **Assess process maturity**: Well-defined support processes are key to improved support. The extent to which critical support tasks are acknowledged and processes are documented and repeatable creates a foundation for successful application support.

- **Take a strategic approach to performance management**: The traditional focus on tactical performance management (i.e., what is happening) does little to provide database administrators (DBAs) and other infrastructure support teams with the type of in-context data that can dramatically affect future infrastructure decisions. Demand now exists to better understand by whom and how the database is being used, and how the infrastructure might be more efficiently utilized.

- **Automate wherever possible**: With well-defined and repeatable processes and a more strategic perspective on infrastructure management, IT organizations can use management tool suites to automate operational support and improve performance problem detection. This reduces application total cost of ownership (TCO) by controlling the full-time employee (FTE) costs associated with maintenance and improves the proactive capabilities of the support process.

- **Centralize data resource management**: As IT organization demographics skew (20% currently over 50 years of age) and new demands for data usage increase, IT organizations should create a center of competency focused on data management. This will promote process maturity, knowledge sharing, and improved coverage for critical application development and operational support.

With the ever increasing challenge of data resource management, organizations must streamline their procedures. Comprehensive toolsets that can help organizations automate best practices, become more proactive, and free up
valuable DBA staff time for higher-value tasks are essential. Tools that provide visibility across the infrastructure spectrum (i.e., application, network, server, database, and storage subsystem) and reduce critical mean time to solution will be invaluable assets in achieving the organization’s support objectives.

**Introduction**

Database environments have always been critical platforms to be managed. Historically, DBAs have demanded the highest salaries among non-management-level IT workers. META Group predicts an increasing demand for skilled DBAs through 2010. Indeed, the US Department of Labor Bureau of Labor Statistics (BLS) is forecasting a 66% increase in the demand for database administrators. In META Group’s view, this number factors in only pure demographic information; it does not factor in the increasing demand for data-driven decision making and the complex mixed-workload environments (i.e., transaction processing and analytics) that will challenge DBAs to meet required service levels.

The task of managing the performance of the database, and by extension the applications that depend on that database, can be complex. The challenge of tuning a database environment is inherent in the complexity of the database subsystem and the subsystems on which the database is dependent (e.g., CPU/hardware, disk/storage). Add to that the complex, multi-tiered applications driving the workload on the RDBMS, and the tuning challenge is even more daunting. A strong tuning process and tools to implement that process are critical to managing the increased complexity of database tuning.

**New Demands on Database Infrastructure**

The organization’s ability to obtain data and use it for various purposes has never been greater. As business processes become increasingly automated, additional classes of data are being identified. META Group research indicates that 77% of organizations plan to capture and analyze even more detailed data (e.g., subtransactional) in 2004 than they did in 2003. We predict that this demand will result in data volume growth of 125% annually through 2007. In addition, new types of data and processing (e.g., unstructured, document management) are now being folded into the relational database management system. As an organization’s informational needs mature, the performance of the database becomes increasingly critical and more difficult to manage.

**Multi-Tiered Infrastructure**

The process of relational database tuning is complicated by the fact that the RDBMS is essentially a subsystem dependent on other subsystems. These subsystems include the application code itself, the network that connects the
application to the database, the server capacity on which the database subsystem executes, the database subsystem itself, and finally the storage subsystem on which the database reads and writes data.

Figure 1 — Multi-Tiered Infrastructure

When a DBA investigates performance issues, the first critical task is to identify which subsystem is the root cause of the performance problem. This requires tools that can collect information on end-to-end response time and break down the amount of time spent in each subsystem. Tools that can collect and correlate this information from multiple tiers are critical to improving the overall mean time to
solution while reducing much of the time wasted in unnecessary finger-pointing exercises among infrastructure support staff. Ongoing historical monitoring and analysis of these interdependent subsystems are crucial for implementing a strategic performance management approach.

The DBMS subsystem is complex in and of itself. With hundreds of configuration parameters and options on parameters, the DBA faces a daunting task. Although database vendors are improving their autonomic capabilities, many organizations will not implement these newer database versions in production for several years. And even then, organizations will need to support heterogeneous database infrastructures, which will require tools that span database platforms.

Since performance management is a process, which by definition should be continuous in nature, the DBA must have visibility into the following two distinctly different yet related areas.

The Database’s Use of Hardware
How the database interacts with the hardware platform it runs on must be well understood. Particular focus should be on the features offered by the DBMS to provide and manage memory, provide parallel processing, and eliminate (or speed) physical I/O (e.g., partition elimination, prefetching). Interesting features to look for are connection affinity (to reduce context switching) and I/O affinity. Users should thoroughly test the application after setting appropriate database initialization parameters to avoid unbalanced use of resources (e.g., maximum capacity of one CPU while another is idle).

From a physical I/O perspective, it is important to understand how different physical database files are used and associated with tables or tablespaces so that object placement on the disk is optimized. For example, database log files should be on their own mirrored drive with no other I/O activity. This makes sense due to the critical nature of logs in recovery and the intensive read/write activity. The sequential nature of log write activity is the antithesis of the random I/O seen in most transaction processing systems. Placing logs on separate drives reduces the disk’s head movement/repositioning and improves critical logging performance.

The Application’s Use of the Database
The workload of the application and how it uses the database should be calculated through the use of SQL tuning tools and application server monitoring tools. The goal of this exercise is to model the ratio of application “transactions” according to the volume and type of access to the database (e.g., insert, update, select, delete). Through experience, constant monitoring, and historical trending,
administrators can more accurately predict scalability requirements and become proactive in dealing with performance issues.

**Growing Storage Demands and Options**
As organizations are becoming more analytically mature, a trend is emerging around a centralized enterprise data warehouse approach. Companies investing in enterprise data warehouse construction have matured to a point where disjointed pockets of information (or marts) can no longer satisfy the organization’s increased need for higher quality and more timely business analytics. What these organizations seek is a platform that can handle the demands of multiple subject areas and large numbers of concurrent users, accessing multiterabyte databases.

As DBAs approach the task of very large database support, storage management becomes a key component of their job descriptions. Although RDBMS software inherently has the ability to process requests for data in parallel, it needs help. Indeed, the physical I/O process is the most expensive task (in terms of time) that a database performs on behalf of an application. In the process of performance management, DBAs are likely to spend a significant portion of their time tuning the database's storage subsystem. They will require tools that can identify I/O hot spots and take appropriate action. This may include "striping" a table or tablespace’s data across multiple physical disks to make use of the database's parallel capabilities. We note that striping can occur at the server level or within the storage subsystem, which can make resolving from logical to physical database object layout an extremely challenging endeavor. Appropriate action may also include identifying disk contention between heavily accessed or joined tables and separating their underlying data files into different disk arrays. Another consideration is the underlying RAID level for the disks on which the database objects reside (e.g., RAID-1 or RAID-5).

Several options beyond direct-attached storage now exist that add further complexity to this aspect of database tuning. Certainly, storage-area networks (SANs) have simplified storage management in many ways and are the recommended storage platform for most database processing. However, SANs can be expensive, and some organizations are evaluating the less expensive network-attached storage (NAS) option for some application database infrastructures. Hosting databases on SANs and some file systems on NAS remains a best practice. However, the line between SAN and NAS will become increasingly blurred as block-level and file-level access converge. Although this distinction may be immaterial by 2005/06, it currently remains important. We urge organizations to carefully monitor and test the application’s workload on NAS before going into production, to ensure that the network bandwidth is not an I/O bottleneck for database performance.
Mixed Workloads and On Demand Scaling

Through 2010, requirements for the use, processing, and dissemination of data will drive database performance innovation by combining the realm of online transaction processing (OLTP) with traditional data warehousing without being defined or constrained by the elements of the two original paradigms. This signals a shift toward operational (or active) warehouses. This trend presents a tremendous challenge for DBAs, since the level of effort needed to manage database performance in a mixed workload environment is 3x that of a typical OLTP-only environment. In addition to the previously mentioned challenges, mixed workload environments combine the needs of perhaps thousands of concurrent users executing tactical queries and hundreds of analytical processes executing complex queries while data is being loaded, inserted, or upserted. Another significant difference between OLTP and data warehousing is that the workload for OLTP is generally repetitive and fairly predictable, which is not the case with data warehousing.

In these environments, it is critical that DBAs have the tools to monitor key performance metrics and capture and analyze SQL statements and trends. The DBA has always played a role in capacity planning, but for many organizations, this task has represented an event in the development cycle, not a process that continues throughout the life of an application. This practice has resulted in extremely poor utilization rates for distributed servers, which in turn have cost organizations millions of dollars in wasted capacity. This is true for both servers and software licenses, especially those of database software, which often is licensed by number of processors.

The promise of on demand infrastructures will likely spur a great deal of infrastructure consolidation. In fact, much consolidation of servers and storage has already occurred. This continuing trend will make the challenges of tuning the RDBMS environment more complex. Certainly, capacity planning will be squarely on the shoulders of the DBA, and it now must be continuous in nature. In addition, consolidation means that contention will exist between applications sharing the consolidated infrastructure. Although performance scaling can be handled by the on demand adaptive nature of next-generation infrastructures, management will be at a premium as it has always been with shared infrastructure (as with the mainframe). Therefore, tuning tools will have to able to resolve both intra-server and inter-server performance issues.

The Evolving Role of the DBA

As technology improves, less-sophisticated jobs are eliminated while higher-level positions are invariably created elsewhere. Progress continues to push the labor force to acquire new and much more challenging skills. Yet there are problems
with this pattern. First, it assumes that the prime labor force (ages 25-54) will continue to grow. Second, those entering the labor force are better educated than the preceding generation. According to some studies, the size of the prime labor force in the US will not grow during the next two decades. In addition, the percentage of the labor force attending four years of college is likely to peak at 60%. Statistics like these lend credence to the current figures from the Bureau of Labor Statistics, which predict a 66% increase in demand for DBAs by 2010.

Of course, best-practice organizations will be in a position to use database management tools to automate many of the highly repeatable management steps, once they are documented and implemented. However, as previously mentioned, new challenges and demands are already emerging, forcing the DBA role to evolve into higher value-add responsibilities. This evolution will require higher-level skill sets and more experience than typically have been required of the DBA. This also means that the number of DBAs required relative to their percentage of overall IT organization staff will also grow, perhaps to as much as 6% of the overall IT organization versus the current average of 2%. In fact, we may be witnessing the birth of a new role altogether — not simply a database administrator, but a data manager. The data manager must possess a deep knowledge of database engines, how data flows through the business application infrastructure, and how data should be handled efficiently, securely, and reliably.

**Organized for Success**

There are essentially two methods that most Global 2000 organizations deploy: 1) use of an application DBA position with a decentralized model; and 2) creation of a data resource management (DRM) group.

The application DBA position implies a decentralized model in which individual DBAs are placed within the reporting structures of the various application development groups throughout the company and report up to an application development manager. There are several issues with this reporting structure. The first is the lack of skills and process sharing. This model perpetuates parochial attitudes toward data ownership and slows analytic maturation of the organization. In general, this model leads to uneven levels of support for the application group and will require additional headcount as we move into a period of greater demand for skilled DBAs. Certainly, strong process definition and automation can reduce DBA headcount within these parochial organizations, but we recommend that performance tool rationalization be pursued across all support groups to reduce training and improve knowledge sharing.

The recommended model is for the IT organization to create a data resource management (DRM) group. In addition to DBAs, this group should contain data architects, data analysts, and administrators to bring all the data management-
related disciplines under a single umbrella organization. The DRM group creates a center of competency for the organization that is a marriage of the logical and the physical world, of business vision and infrastructure capacity. It instantiates the relationship between application development and data management, enabling organizations to build better performing, more secure, and highly available applications than are possible in the decentralized model, where development and production support are often separate organizations.

Staff organization is only the first step. The next step for the DRM group must be to define and document its best practices, standards, and procedures. This is an essential step toward process maturity, because once a process is documented it can be automated. This is where organizations will truly realize the value of performance tool suites that can capture and coordinate performance information from all infrastructure tiers. Through codification of best practices, application monitoring tools can be configured with proper alarm thresholds and automatic responses to many of the most common issues. This will significantly reduce the work hours required for production support and free up valuable DBA resources to focus on higher-value development projects.

Although many organizations focus on the ratio of DBAs to the database, long-term planning must also focus on the ratio of senior DBAs to junior staff. When centralizing the DBA staff under the umbrella of the DRM organization, the objective is the development of repeatable processes and procedures that, in turn, enable use of less staff to better support more databases. Although this theory is sound, it does not account for the need to maintain the necessary ratio of junior staff to develop the next generation of senior DBA staff.

We recommend maintaining a ratio of 1:2 between senior DBA staff and junior DBA staff whenever possible. We also recommend that junior DBA staff be paired with a senior DBA and, in turn, be assigned a subset of the senior staff person’s database responsibilities. In this way, junior DBAs can learn on the job and become steeped in the process and procedure of the DRM organization under the guidance of more experienced staff. By developing skilled DBAs internally and leveraging management tools well, organizations hopefully can stay ahead of the curve and avoid the need to pay exorbitant recruiting fees or salaries to outside candidates that, regardless of experience level, will still require some training.

**Strategic Performance Management**

To manage strategically, organizations must consider moving beyond the interrupt-driven management cycle to address performance and scalability issues from a long-term business-value perspective. Tools that provide historical context for
everyday performance issues (a.k.a. “performance warehouses”) are invaluable. They enable DBAs to evaluate whether recent performance issues (e.g., within the past hour or day) were an aberration or a trend signaling a need for change. Indeed, even seemingly simple infrastructure changes (e.g., creating an index) can become costly to a business application database with an expectation of 24x365 availability. It is important to know in advance that the change will not have negative ramifications for performance. Also, the historical context allows current problems to be solved in full context; for example, comparing current unacceptable performance data with prior acceptable performance data to determine whether an occurrence is a trend or a one-time event requiring action. Reactionary changes lead to bloated infrastructures (e.g., disk, servers) that raise not only the cost of the infrastructure itself, but also management costs.

IT organizations and business owners should expect more visibility into their infrastructure and how it is being used. Strategic management actually means the evolution of SQL tuning, performance monitoring, and capacity planning tools pulled together into an easy-to-use integrated suite that stresses usable reporting that is accessible to both the DBA and management.

**Bottom Line**
Application database environments are becoming increasingly complex and the demands for new data and new processing will continue to grow. A convergence of organizational, technological, and economic forces is pressuring organizations to “get out in front” of the issue of database tuning and performance management. Although there is never a single silver bullet that solves complex problems, strong, repeatable, and managed process can help improve data handling and bring solutions to light in a more efficient manner.

Organizations should evaluate their current database process maturity, ensuring they are taking the correct steps and applying them consistently across the organization. Management tools must be leveraged that can span the infrastructure from application through database and storage, with as much of the process automated as possible. Organizations should use the information (both current and historical) to make better use of the their IT infrastructure. Finally, IT organizations must think strategically about the management process and the company’s overall goals and ensure that they manage to those objectives.

*Charlie Garry is a senior program director with Infrastructure Strategies, a META Group advisory service. For additional information on this topic or other META Group offerings, contact info@metagroup.com.*
About META Group

_Return On Intelligence℠_

META Group is a leading provider of information technology research, advisory services, and strategic consulting. Delivering objective and actionable guidance, META Group’s experienced analysts and consultants are trusted advisors to IT and business executives around the world. Our unique collaborative models and dedicated customer service help clients be more efficient, effective, and timely in their use of IT to achieve their business goals. Visit metagroup.com for more details on our high-value approach.