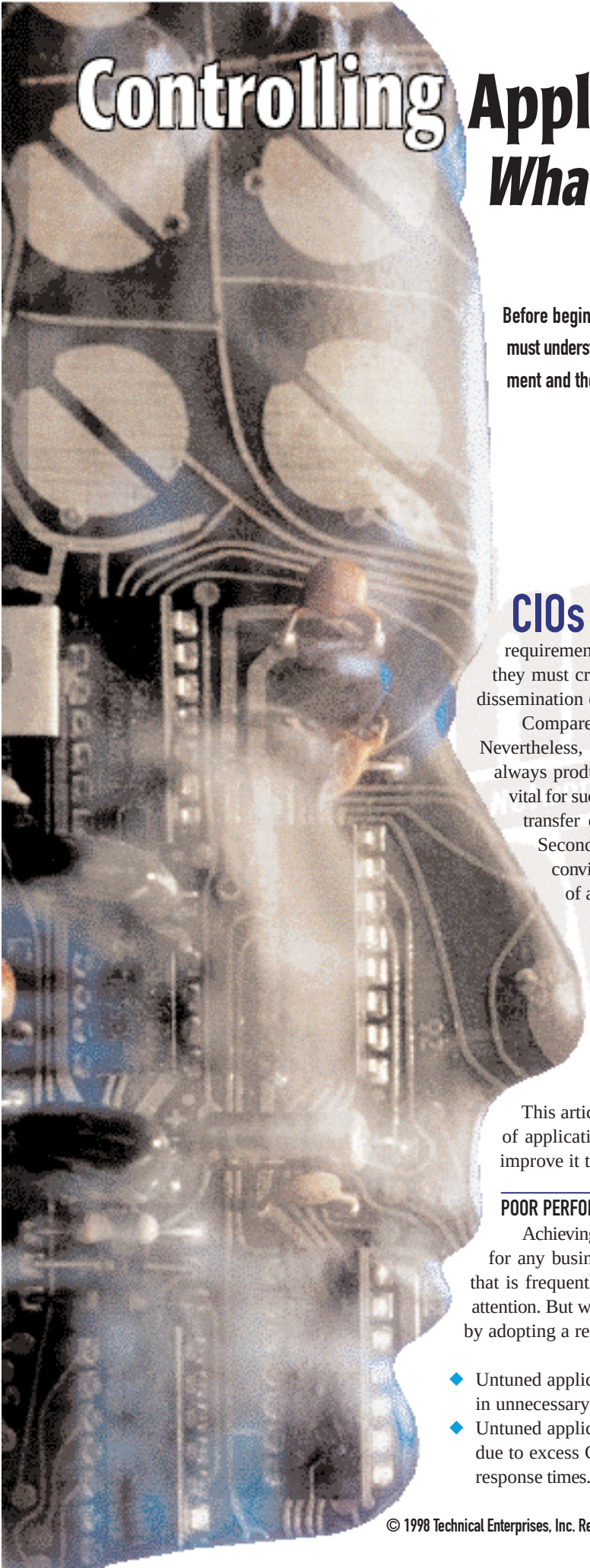

Controlling Application Performance: *What Every CIO Should Know*

BY CHRIS LOOSLEY

Before beginning a systematic approach to performance assessment and management, CIOs must understand the cost impact of inadequate approaches to application performance management and their role in addressing the problems and costs of poor performance management.

CIOs face complicated issues of data center infrastructure, a portfolio of differing application workloads, and rapidly changing business requirements. Strategically, CIOs need to react quickly to competition. Tactically, they must create common processes across business units to make collection and dissemination of information simpler and to cut the cost of new applications.

Compared with these concerns, application performance lacks glamour. Nevertheless, a systematic program of performance monitoring and tuning almost always produces surprising cost savings for large companies. CIO sponsorship is vital for such a program to succeed. In the first place, only the CIO can negotiate the transfer of funds among the capital equipment, salary, and software budgets. Second, management prioritization and employee motivation are required to convince lower-level managers and technical professionals of the importance of application performance. The CIO can provide the necessary emphasis.

Before beginning a systematic approach to performance assessment and management, CIOs must understand two issues:

- the cost of inadequate approaches to application performance management
- their role in addressing the problems and costs of poor performance management

This article discusses both of these themes as well as some of the practicalities of application performance – how to measure it, how to manage it, and how to improve it to better meet the needs of the business.

POOR PERFORMANCE YIELDS HIGH COSTS

Achieving acceptable service levels is not an option but a fundamental requirement for any business. However, performance is a vital component of software quality that is frequently forgotten. The top 10 problem applications may receive plenty of attention. But what about the remainder? Most companies increase cost to the enterprise by adopting a reactive approach to performance management:

- Untuned application software costs companies millions of dollars annually in unnecessary processing.
- Untuned application software results in unnecessary and costly hardware upgrades due to excess CPU utilization, increased batch processing time, and/or unacceptable response times.

Can You Save Money?

The following is a simple checklist to evaluate your organization on a 10-point scale.

Evaluating Your Organization's Commitment to Performance Management

Evaluation Criteria	Yes/No
1. Does a senior manager oversee performance management activities and compliance with systems/application performance objectives?	
2. Are your organization's approaches to performance issues proactive?	
3. Is there either a separate group or identified cross-functional team whose primary mission is systematic performance management?	
4. Do you manage performance using a team approach that coordinates the different skills and perspectives of all the relevant technical disciplines?	
5. Do you have a mechanism to establish shared application performance goals?	
6. Do you have a common language (and/or tools) for discussing and measuring compliance with performance goals?	
7. On an ongoing basis, do you automatically measure and track application performance?	
8. Are development and systems staffs rewarded for reducing processing costs?	
9. Are application performance goals set and tracked during development?	
10. Must applications meet performance criteria before running in production?	

Evaluate Your Organization

Score 1 for every "Yes" answer. For most organizations, achieving any score above 5 would produce visible cost savings.

- Untuned application software disrupts business operations. Latent performance problems appear during peak business periods when mission-critical information systems become overtaxed.
- Untuned application software undermines mission-critical business practices.
- Inefficient programs waste expensive computing resources every time they are run and cost staff time to fix.
- Slow systems frustrate employees who need these systems to do their jobs.
- Productivity drops and attrition rates rise.
- Untuned applications become the real "performance disaster" story.

By failing to address performance problems or potential problems, companies most certainly face increased costs and may likely run into a full-scale performance "disaster." New systems run into performance problems so often that nearly every company has experienced such a disaster. Trade journals often print stories of millions of dollars lost due to a critical application that, at the last minute, proved unworkable. Most such disasters are quietly buried and soon forgotten, except perhaps by the CFO and those executives unlucky enough to be blamed.

Right behind disasters comes a much larger class of applications that actually are placed into production, only to become a continual source of irritation and complaints. Although they are functionally correct, their sluggish behavior has costly side effects:

Inefficient processing has indirect costs, too. Unresponsive business systems can lead to loss of business when the systems that were supposed to support a business process cannot. Companies today use computers for more than mere bookkeeping. Effective CIOs aim to have all the right information resources and services available promptly to support interactions with customers. When vital information processing applications run slowly, customers are often directly affected. We have all stood in line while the desk clerk apologizes for "the slow computers."

The bottom line: One way or another, poor performance is sure to cost money — in wasted development costs, excessive processor utilization, or lost income. Few companies can afford to pay the price.

WHO IS RESPONSIBLE FOR APPLICATION PERFORMANCE?

One of the central challenges of performance management is this: Although performance involves everyone, often no one feels fully responsible for it.

In information systems, there are many different jobs, and people must specialize to be effective. A major division exists between those whose focus is applications and those whose job is to manage the systems. Systems specialists struggle to keep up with new technology while application developers seek to understand the business.

Performance management is affected by this division of skills because application performance is determined by both internal and external influences:

- Internally, design and implementation determine the load an application places on the computer system's resources.
- Externally, the computing environment determines the time it takes to process the application's workload.

Application developers are not trained in the externals. Systems specialists know little about what an application does, even less about what it is supposed to be doing, and nothing at all about why it's doing it.

Both groups influence performance but neither has an explicit focus on application performance. Improving application performance demands a set of skills that cuts across departmental boundaries. For this reason, performance management is often addressed in a reactive mode.

However, some companies have adopted a more systematic approach, based on viewing performance as a team responsibility. They either establish a group of skilled performance specialists with the right backgrounds (the Performance Department) or with a matrix management approach they bring together the necessary skills from different departments (the Performance Review Board).

The performance team, however constituted, must have the support and cooperation of all the other departments whose decisions and actions affect performance. This is a requirement that demands attention at the highest levels.

REQUIREMENTS FOR A PERFORMANCE MANAGEMENT ENVIRONMENT

The astute CIO may realize the importance of application performance and even

have a sense of who should comprise the performance team. However, for truly successful application performance management to take place, four ingredients must be present:

- **Politics:** The right level of management sponsorship
- **People:** The right mix of skills in the performance team
- **Profiles:** The right set of tools to create and track application usage profiles
- **Process:** A systematic performance management process

Politics and the Management Sponsor

The importance of a high-level management sponsor cannot be overly stressed. Although performance management is primarily a technical discipline, it often involves politics: Performance issues touch many different areas of the enterprise, and disputes about service levels can lead to turf wars.

Entrenched positions can determine how and when performance issues are tackled. No one may be willing to fight a political battle when "everyone knows it costs too much to manage application performance across the board."

Service-level management approaches solve this kind of problem by establishing shared ownership of performance and service-level issues and a common way of thinking about performance and business goals.

An effective way to focus on the performance issue is to mandate that the internal cost of running the company's applications be reduced by a percentage over a defined time period.

People With the Right Skills

Many individuals must collaborate to achieve satisfactory performance:

- The user community sets performance goals and provides the ultimate measure of an application's efficiency.
- The business analyst identifies performance requirements for each business process.
- The application designer ensures that program structure supports these requirements.
- The programmer writes code implementing these structures.
- The database administrator designs and tunes the database for satisfactory performance.
- The capacity planner reviews overall utilization of system resources.

No performance management process will be completely effective unless individuals with all these skills participate actively.

Finally, there are management skills. For the CIO, these involve both the creation of talented project teams and the communication of corporate values.

Profiles: The Key Starting Point

Performance management begins with a set of application performance profiles to determine how applications use computing resources. To create application performance profiles on an ongoing basis, we need the following:

- measurement tool(s) to gather the raw data about application behavior
- analysis tools to filter and summarize measurement data and develop profiles
- a database of application performance history
- analysis tools to read the database and simplify trend analysis

To overcome communication barriers, it is essential for team members to speak "a common language." Here, a common set of tools is very helpful. If team members use the same tools to look at the same data, they are much more likely to understand each other. The tools in question need to focus on application response time, the one performance metric users care about and the one aspect of performance everyone understands.

Process: The Systematic Approach

Finally, the performance measurement and analysis activities must be carried out systematically. The ideal performance management process would involve:

- establishing application and system performance objectives
- systematically measuring application performance against those objectives
- automatically measuring a large number of applications and recording and analyzing the resulting data
- tracking a high percentage of programs in the application portfolio
- automatically highlighting exceptions to defined performance objectives or recent history
- establishing measurement and analysis controls accessible to both development and systems groups
- recording performance information in formats useful to both development and systems groups

A performance management department (or cross-functional team) should adopt this list of objectives.

PROACTIVE MANAGEMENT SCENARIOS

Measurement is the starting point for all performance management. If we cannot measure our applications, we cannot know how to control them. This is why an up-to-date set of application profiles is the first step toward gaining control of our information processing resources.

Once we have tools to measure the resources an application uses, there are different ways to utilize this information, including these possibilities:

- **For Senior Management:** Systematically locating and eliminating unnecessary processing cuts the true cost of the IS function.
- **For Systems Specialists or Performance Analysts:** A periodic audit of the longest-running batch programs and most frequently executed online transactions may reveal inefficiencies.
- **For Capacity Planners:** A capacity planning exercise combining growth projections with measured application profiles can reveal future bottlenecks before they occur. This allows time to weigh tuning and purchase options and take remedial action.
- **For Application Analysts and Programmers:** New applications are built or old ones are modified. In either case, the most cost-effective method is to consider performance as early as possible.
- **For Business Analysts and Application Designers:** Profiling tools can be used to supplement program documentation, which is often either out of date or lost altogether.

MINIMIZING PERFORMANCE SURPRISES

The biggest risk companies face by deferring performance management issues is that severe bottlenecks usually surface at inconvenient times. When this happens, it is difficult to fix the problem

10 Key Challenges of Performance Management

Since no company really wants to waste money, why are performance problems still so common? Here are 10 reasons:

1. **No senior management sponsor exists.**

   It's hard to demonstrate the benefits of application performance management to someone until he or she experiences the cost of not doing it. Risk denial ("There won't be any problems!") and cavalier attitudes ("Someone else will fix them!") are common.

   Clear direction from senior management determines attitudes lower in the organization. When senior management deem good service levels essential, lower-level managers and technical staff will treat performance problems seriously.

2. **Performance management requires commitment and coordination.**

   A key difficulty in effective performance management is that no one person can do the whole job. Effectiveness requires an array of skills and perspectives that span several departments.

   In companies that have already made commitments to coordinate performance management at the highest levels, the newest technology has been extended to include performance concerns.

3. **Communication barriers abound.**

   Whenever people with different skills and backgrounds need to cooperate, communication is a challenge. Ironically, the world of "information systems" is a proverbial tower of Babel:

   - Business analysts speak and think in terms of business information flows and processes.
   - Application developers speak the language of COBOL or C, of loops and procedure calls.
   - Developers know how applications work but view the systems environment as a black box.
   - Database administrators work with indexes, block sizes, buffer pools, and access paths, but they don't know what the applications are doing.

   Unless an organization can establish shared performance goals, a common language for discussing those goals, and a way to bridge the gaps in skills between departments, there is little hope of effective cooperation.

4. **Those responsible adopt reactive management styles.**

   Traditionally, IS departments have created new application solutions in response to business demands. The CIO's reactive approach trickles down to lower-level managers and affects the response to performance management issues, which include these:

   - we are too busy fighting fires.
   - we don't have the time.
   - we don't have the people.
   - we don't have the budget.
   - computers are so cheap that we don't have to plan.

   Introducing systematic performance management tools and processes and requiring key managers adopt a more proactive approach is the only way to break this pattern.

5. **Systems managers see no immediately visible return on systems performance management.**

   Systems managers are governed by "the tyranny of the urgent." It is difficult for systems managers to assign skilled staff to track the performance of application software that is not visibly causing a problem. More urgent and visible problems demand immediate attention. Performance management can wait until tomorrow.

   The only effective counter to tactically oriented resource allocation is a more strategic view by the CIO.

6. **There is little focus on performance during development.**

   Many performance problems arise because of developer inexperience or lack of tools. Necessary tools do exist, and organizations possess the skills to use them, but they choose to focus instead on other priorities such as function, schedules, or costs. Although acceptable application performance is certainly not the only goal of software development, it should not be ignored either.

7. **Development managers see no visible return on application performance management.**

   A major obstacle to application performance management is that development managers are asked to invest time and resources early on, with no visible return on the investment.

   Many managers say in effect, "There won't be any problems, and even if there are, someone else will have to fix them." These attitudes persist as long as the focus is on delivering timely application functions within budget. But there is an important caveat to software development: Software can be good, quick, and cheap — pick two!

   Only the CIO can alter this equation by factoring in the later costs of poorly performing applications.

8. **For developers, performance is not exciting.**

   Many developers want to use the latest tools and techniques. In the rush to embrace new technologies, application performance concerns can be overlooked.

   Performance needs can be given their proper importance if senior management enforces the right attitudes. A CIO has the opportunity to address the separate issues of staff attrition and costs at once. By giving more visibility to the real costs of poor application performance, CIOs can show developers that the quality of their technical efforts has a direct impact on the bottom line.

9. **Developers regard performance as a "systems" issue.**

   Many developers act as if performance is unrelated to functional design. While an enterprising programmer may take the initiative to do "extra" tuning work, others assume that performance issues are the responsibility of the database administrator or systems specialist.

   These misconceptions interfere with making performance an integral element of application quality. An application development process that incorporates efficient resource use and response objectives will counter such misconceptions.

10. **There are too many programs to track manually.**

    Systems specialists often work on performance-related issues, but applications receive scant attention. There are so many programs in the typical production environment that only the largest, most frequently run, most poorly performing programs are tracked on a regularly basis.

    A proactive approach demands a comprehensive, automated way to monitor applications coupled with a systematic review process.

and provide the service levels the business needs. The remedy is a more systematic approach to performance management throughout an application's life:

- an early investment in performance planning and design will maximize return on investment in development resources and pay off in applications that deliver the performance users need to meet their business goals.
- an investment in ongoing application performance management will maintain desired service levels and minimize the cost of the resources.

REENGINEERING TO SAVE MONEY

Knowing what the applications do and where they spend time is the starting point for all other performance management activities. Having a systematic way to obtain and review application performance profiles helps an organization do a better job of the following:

- minimizing resource costs needed to support the production systems
- tracking and controlling the performance of production applications
- removing inefficiencies in pre-production applications

Organizations can automate the application performance measurement process to pinpoint inefficiencies and introduce a directed form of reengineering. The outcome will be faster applications, improved perceptions of corporate responsiveness, availability of precious computing resources, and a direct impact on the company's profitability.

A FINAL WORD FOR THE CIO

Everyone knows we live in the information age, in which a company's success hinges on the quality of its systems. But in today's marketplace, picking the right information technology is no longer enough. At least as important are innovation and customer-focused operations.

With the right balance of business and technical insight, a CIO can take a leadership role in strategic planning. The CIO must aim not simply to support a strong partnership among the business, its customers, and suppliers, but rather, to make the corporate information systems the cornerstone of that partnership.

Developing efficient, responsive information systems that run smoothly is a key step in reaching that goal. Studies have shown that users' perceptions of a system's responsiveness are based only on the slowest 10 percent of its responses.

As the business becomes ever more synonymous with the business information systems, it is essential for the CIO to take steps to minimize the occurrence of severe delays and other performance-related problems. ts

This article, which is an excerpt from a White Paper, is available in its entirety by faxing your request to Emerald Software at (617) 254-8979.

Chris Loosley is the managing director of the Institute for Computer Capacity Management (ICCM). He has more than 25 years of experience in the design of enterprise information systems. His technical background is in DBMS design and performance, and his central interest is in software architecture and design and software performance engineering. For more information about ICCM, visit their web site at www.iccmforum.com.

Pay Me Now or Pay Me (More) Later

One myth of performance management is that proactive application performance management is not cost-effective. Reality is quite the opposite. In the long run, a proactive and systematic approach to performance management:

- does not raise normal application development and systems management costs significantly
- produces significant savings in the costs of production processing
- leads to a reduction of staff time spent on performance fire-fighting, application redesign, tuning, and rework

The key to these cost savings is a systematic, ongoing program of monitoring at three levels:

- **Level 1:** Continuous, low-cost exception monitoring. Most well-engineered software subsystems like operating systems, databases, and transaction monitors maintain their own counts of both normal activities and exception conditions. Sample these periodically. Abnormalities often indicate performance changes without the need to start any special traces that might affect normal processing.

- **Level 2:** Regular, targeted performance tracking. Monitor key performance indicators for a cross section of critical applications. Do this regularly during both peak and average periods. Record application response times, processor utilizations, and other critical resource utilizations. Track these key indicators over time for signals of future performance problems.

- **Level 3:** Occasional, focused performance audit. Occasionally, focus attention on a single performance-critical application. Create a profile and compare it with previous profiles to detect changes and hidden or potential problems as workload volumes increase.

Managers tend to overestimate the cost of systematically monitoring performance across the board. A performance crisis demands an immediate, massive, and widespread performance audit (level 3). The routine activities of levels 1 and 2, by comparison, are more focused and less disruptive; they make lower demands on resources.