System i Optimization Best Practices

VISION SOLUTIONS

Introduction

System i is at the heart of the IT infrastructure for many organizations. Companies around the globe depend on the platform to provide reliable performance for their sales, operations, human resources, accounting, shipping and other critical applications. While it is true that System i is one of the easiest to manage business systems in the world, it is still far from entirely self-managing or self-optimizing.

Normal business growth brings with it constant changes, not the least of which are changes to applications and data structures. Under the best of circumstances, as these changes and modifications are applied, and as upgrades and migrations are performed, the probability is high that related system optimization tasks may be delayed or left undone. Over time, multiple rounds of changes and modifications can compound the reduction of system efficiency. Any poorly planned or executed changes can make matters even worse.

The need for effective system management and optimization will not come as a surprise to the IT department. Nevertheless, with IT resources under increasing strain, these tasks are often relegated to the bottom of the department's to-do list. The reasons for the strain on IT personnel are varied. Competitive pressures demand more value-enhancing application functionality, but the IT department must deliver the new features with leaner staffs. Rising downtime costs force system managers and operators to devote more time to ensuring high availability, reducing the time they can spend on other tasks. And growing IT environment complexity places yet further burdens on the IT staff. All of these demands reduce the resources the IT department has available to undertake vital System i management and optimization tasks.

A large part of the solution to this conundrum is the same as the solutions that IT has used to fulfill the rest of its productivity goals: automation. A set of highly automated System i management and optimization tools can reduce the strain on the IT department, while ensuring timely performance of vital management and optimization tasks. But before you can set out to evaluate the tools available to undertake this automation, you must first have a clear understanding of what needs to be done. This white paper examines some of the areas that you need to address in order to ensure healthy System i operations that deliver the highest possible performance.

Monitoring

Intuition alone rarely advances you far along the path toward ensuring good System i health. All aspects of System i performance and resource usage can and should be measured and monitored. The statistics you gather can serve a number of purposes. For example, you may want to compare current performance levels against company goals and/or industry norms; or your intent may be to set a baseline against which the success of future optimization efforts can be evaluated; or you may need to use the data to validate compliance with an internal or external Service Level Agreement (SLA). Whatever the specific objective, knowing the facts—facts that continually change as your System i usage evolves—is critical to maintaining System i health.

This section examines the areas of System i operations that require monitoring.

CPU Usage

What determines whether the organization will rate an application as successful? Clearly, functional completeness and correctness are the first considerations, but, beyond that, response times are also critical success factors. The difference between sub-second response and, say, a two or three second response does not seem substantial until you multiply it by many thousands of transactions a day. Any additional response time per transaction can quickly translate into a daily loss of up to several hours of productivity across the whole company.

Possibly worse, slow applications cause users to become frustrated and their idle moments weaken their concentration on the task at hand. That's bad enough when users are employees, but when they are customers interacting with, for example, Web-based sales systems, their frustration can result in significant lost revenue.

Efficient code will obviously run faster than inefficient code, but, for any given program, CPU utilization is a primary determinant of response times. To keep your finger on the pulse of your System i health, and to garner the information necessary to resolve any CPU utilization issues, you need a monitoring tool that will answer the following questions:

- ☐ What is your capacity utilization throughout the day?
- ☐ Are there regular periods during the day when utilization peaks too close to 100%?
- ☐ Are there regular times when the CPU is significantly underutilized?
  If so, can some of the jobs currently running during peak times be rescheduled to run during those slow times?
- ☐ Are one or more applications using inordinate amounts of CPU resources?

Disk Usage

The dramatic drop in the cost-per-gigabyte of disk storage over the past several years has led many IT professionals to take a somewhat cavalier attitude toward storage usage. A common outlook is, "If a disk fills up, even if much of it is wasted space, who cares? Just buy new disks. It's not worth my time to worry about it."

There are a couple of problems with this attitude. First, disk is cheap, but it's not free. True, if databases had not grown over the past decade or two, total storage costs would be trivial now. But that is not the case. Databases have grown extraordinarily as business volumes have expanded and companies have automated more functions and collected more types of data. In the not too distant past, if 30% of disk space was wasted, the total amount of squandered space would still have been a low absolute number, from the perspective of today's gargantuan databases and storage pools. But today, when multi-terabyte databases are common, 30% wasted space can create a need for significant disk expenditures.

Furthermore, the cost of DASD is often the least of the problems. The threat to application performance may turn out to be much more serious. As an example of how response times may be degraded by superfluous data, records that have been logically deleted, but not yet physically removed (which is normally the case until the file has been reorganized), are still brought into the buffers during read operations. The logically deleted records are immediately filtered out, but not before burdening your system with unnecessary I/O operations.

Obsolete data that is neither physically nor logically deleted from a database may also result in performance degradation. If the column that determines whether the data is obsolete is not indexed or if that determination is made in the program logic outside of the SQL query, I/O operations will be required to read in and then filter out the obsolete data. And when the column is indexed, the obsolete data creates indexes that are larger than necessary.

Obsolete objects that are still hanging around on the system usually don't have the same impact on application performance as obsolete data in a database, but they can rob you of considerable storage space nonetheless.

It is, thus, critical to monitor disk usage not only to keep storage costs down, but also to ensure that your applications continue to perform optimally. As the above discussion illustrates, monitoring should not be limited to simply looking at raw disk capacity utilization. On an ongoing basis, you must also examine how that space is being used and how much of it is being wasted.

In a typical company, 15-40% of disk space is normally consumed by obsolete and unnecessary data.

Your storage monitoring tool should provide a complete inventory of all objects on your system and answer questions such as:

- ☐ How much disk space is being consumed by each application?
  Is one or a small number of applications consuming the majority of the DASD space?
- ☐ How is your DASD being used?
- ☐ How much space is being consumed by each type of object and each data file?
- ☐ Are there any objects that have not been accessed for a long time and, therefore, are likely obsolete?

- Are the organization's data retention and archiving policies being followed? (Which begs the questions: Has your organization established such policies? Are the policies known by the appropriate people? Is the execution of those policies automated?)
- How much disk capacity is being wasted by logically deleted records that have not yet been physically deleted?
- Which files are "bound" by deleted records? That is, which files have become unwieldy and burdensome to process because embedded deleted records are continually being passed in and out of buffers?
- How much disk space could be made available using data compression?

Growth Trends

It has been said so often throughout history as to have become trite, but it bears repeating all the same: the one constant is change. If things go well, businesses will grow and add new operations. And, whether things go well or not, organizations modify their processes from time to time. All of these changes will have an impact on system and storage usage.

Change happens, but new servers and disk space can't be added instantly to accommodate new requirements. Gaining approval for the purchase of new DASD, a new server or a server upgrade requires hard evidence of need and lack of other options. Ordering and installing that equipment typically takes at least weeks (or possibly months) no matter how vital the purchase may be. So it becomes critical not only to be able to utilize what you already have to its fullest, but also to be able to recognize well in advance when expansion or upgrade is going to be required. Accurate, detailed reports on conditions before and after system optimization efforts become a vital tool, both for ongoing management and forward system planning.

Even when the solution does not involve a new purchase, but rather the execution of system optimization tasks, those tasks must be scheduled to be undertaken when the necessary IT personnel are available and when the optimization functions won't unduly impede business operations.

It is, therefore, important to monitor CPU and storage usage trends so that you can be proactive. To do so, your CPU and DASD monitoring tools must be able to answer the following questions:

- What is the CPU and DASD growth trend for each application?
- Which batch applications have significantly longer run times than they did, say, six months ago?
  What is the rate of growth in those run times?
- After considering existing CPU and DASD utilization and the projected utilization for planned projects, should you be budgeting for more CPU and/or DASD capacity? If so, how much?

Statistics Storage & Reporting

SLAs introduce a new reason to gather and keep system metrics: Those metrics are required to verify SLA compliance.

The monitoring functionality of your system management tools should store all of the key metrics and make them easily accessible. This is especially important for IT departments that have entered into SLAs with their user departments. Access to this data may be the only way to accurately validate SLA compliance.

Your system monitoring and reporting tools should also allow you to easily analyze the data to identify and resolve any issues concerning the health of your System i. The tool should provide you with a high level, aggregated view of the information, but it should also allow you to drill down into the details in order to look more closely at specific areas and further investigate problems when necessary. The tool should also offer graphical representations of the statistics because issues that are nearly impossible to spot when buried in massive rows and columns of raw data often become instantly obvious when plotted on a graph.

The ability to download the data to a spreadsheet for further analysis with other BI tools is also very valuable. This is especially helpful when reporting on compliance with SLAs. These exported stats can also feed reporting and support analysis for IT projects and planning in general.

Data Tuning and Optimization

Effective monitoring is necessary, but it isn't sufficient. Once you know what needs to be done you then have to do it before you can reap the rewards. Data tuning and optimization have two primary objectives: minimize storage usage and maximize application performance. Because data optimization is often the most straightforward and best way to boost application performance, it should be performed on a regular basis. Achieving the highest of performance and storage utilization objectives requires that you regularly complete a variety of tasks, including data archiving and purging, file reorganization and data compression.

Archiving and Purging

Many companies retain considerable historical data that is no longer used for ongoing operations and rarely, if ever, used for analytical purposes. When this data is kept in production databases it consumes storage space and can slow down applications, without adding significant value for the organization. Archiving old data to near-line or off-line storage and then purging it from the production databases, helps to keep production systems healthy, while still keeping the data available when needed for reporting, analysis or auditing purposes.

File Reorganization

Limited or non-existent maintenance windows make it difficult or impossible to schedule RGZPFM file reorganizations.

When records are deleted from a database they are deleted logically, but they usually continue to occupy space and are included in I/O operations until they are physically deleted by file reorganization. This is usually accomplished by an RGZPFM (Reorganize Physical File Member) command.

In larger systems, the more serious issue may be ensuring that several thousand libraries or thousands of individual files are being utilized and run in the most efficient manner. The effort to manually monitor and manage systems with large numbers of files and libraries can be daunting.

Database administrators recognize the importance of reorganizations, but they may be reluctant to perform them because the RGZPFM command normally requires dedicated access to files. Because the reorganization process may require several hours to complete, it must be scheduled during "off-hours" to avoid seriously impacting business operations. Even though the file reorganization utility that comes with System i is now capable of being executed while production applications remain active, the strain on the system and the locks placed on database files and records typically results in an unacceptable degradation of business application performance if a reorganization is run during peak hours.

Organizations with particularly large databases and long operating hours, possibly 24/7, face a serious dilemma as their "off hours" are invariably considerably shorter (zero for hectic 24/7 operations) than the file reorganization job run time. Companies that operate under this scenario struggle to find an acceptable time to perform file reorganizations. As a result, DASD continues to fill up and application performance continues to degrade.

It is possible to bypass some of these issues by setting the "re-use deleted records" parameter on the physical file or table. When this parameter is turned on, new records can be written over logically deleted records, thereby gradually reclaiming wasted space. This is a workable solution, but many application packages don't make use of the re-use deleted records parameter and customers are rarely successful in convincing vendors to modify the database to take advantage of it.

Another problem with the re-use deleted records option is that it only reclaims space as fast as new records are created. That should not be an issue during normal operations because, on an ongoing basis, more new records tend to be created than old records are deleted. However, when archiving deletes millions of records in one pass, it might take months to reuse all of that deleted space. Furthermore, if the archive is run only infrequently, it might delete as much as a few years' worth of data at a time. In that case, months or years may be required to reclaim that much space solely by adding new records.

Resolving this dilemma requires third-party software that both reduces the impact of file reorganizations on production applications and allows the reorganization to be divided into smaller sub-tasks that can be scheduled at times when demand on the system is normally low.

Data Compression

Compression can shrink objects by as much as 60%.

Compressing objects can dramatically reduce the amount of storage they consume, but there are also other benefits to be derived during backup and restore operations. Most backup routines instruct the operating system to compress programs before writing them to tape. When recovering from a disaster, programs that were uncompressed before being backed up are de-compressed automatically as they are loaded onto disk, but programs that were compressed to start with aren't de-compressed. Therefore, by compressing unused or little used programs, you can reduce both backup and recovery times as there is no need to compress those programs before a backup or decompress them during the restore process.

Operations Management

Every change to a live production environment represents a risk. Poorly executed or omitted maintenance procedures can result in unplanned downtime or, in extreme cases, cause data loss. These risks are significantly increased when, as is often the case, manual maintenance procedures are "improved" by using ad-hoc programs written in control language (CL). These CL procedures typically receive only minimal testing before they are run in the production environment and may not be fully effective.

There are many common operations management tasks that invite such "partial" solutions. Well designed, proven, and reliable, system management tools are necessary when automating these tasks. Following are some specific examples that illustrate this point more fully.

Archiving and Purging Spool Files

Spool files are a regular facet of all system environments, but, if not managed carefully, obsolete spool files can accumulate to consume an ever expanding quantity of disk space. Consequently, regularly reviewing, archiving and purging spool files can free up significant space. If spool file archiving and purging policies are established, system management tools can perform these tasks automatically, thereby relieving the IT staff of the need to do that work.

Optimizing Programs

The change program (CHGPGM) command includes an application parameter that allows values of *NONE, *BASIC or *FULL. Using the CHGPGM command, all programs written in a high-level language, such as RPG, COBOL or C, are translated during the compile process into W-code, a backend intermediate code. From there the code is translated into machine instructions. The optimization process, when called for in the CHGPGM command, works on this machine code to remove redundant instructions and replace some inefficient code with more efficient code structures. At higher optimization levels, the program should run faster, but whether that improvement is negligible or significant depends on the nature of the program.

One drawback of this feature is that it lengthens compile times. And, the higher the optimization level, the longer the CHGPGM command will take to run.

Beyond any aversion to longer compile times, use of this feature is not widespread due to a lack of knowledge about it and due to the time required to set up the task needed to run programs through it. Furthermore, unless the feature is used when new releases of software are produced, the optimization benefits will be lost when the upgrades are implemented. Thus, optimizing programs is not a one-time event. It is necessary to keep on top of all software updates to maintain the benefits.

System management tools can overcome these problems by automating the use of the CHGPGM optimization feature, thereby eliminating the need for expert knowledge about it and reducing the person-hours needed to run it.

Monitoring Logs

In the course of normal operations, logs generated by i5/OS and application programs can consume a sizeable volume of storage. These logs may come in handy when diagnosing any system or application problems that arise. Nonetheless, because most of these types of issues will become apparent immediately or within a few hours or days, the online storage of log files typically provides little or no value beyond a few days after their creation. Consequently, it is good practice to delete most types of old log files (possibly archiving them first so they remain available for audit purposes) every week or month.

Finding and then purging old log files can consume considerable IT staff time, particularly when this task needs to be repeated every week or month. System management tools can automate the process, eliminating the IT workload.

Restricting Task Scheduling

Demands on a system are rarely constant throughout the day and week. CPU loads can be balanced somewhat by running system administration tasks and batch jobs in off-peak periods. The alternative, running them during busy periods, would amplify the peaks and accelerate the need to upgrade servers.

Permanently scheduling system managers and operators to work in the middle of the night and/or on weekends can be an expensive proposition—and it can make it more difficult to retain good system managers and operators. A better option is to employ automated system management tools that are capable of running "while active" without the need for operator attention or the interruption of applications during normal business hours. In addition, to take advantage of the additional cycles available during non-peak hours, these tools must offer the ability to schedule tasks during slow periods.

Reaping the Rewards of a Healthy System i

Like any purchase, the acquisition of system management tools should provide an adequate return on investment. When performing the cost-benefit analysis, consider the following benefits that can be derived from the use of the tools:

- Make better use of CPU resources. Program optimization, database reorganizations and other system management tasks can reduce the amount of CPU resources that are consumed to perform a set volume of the business' work. This provides a financial reward by deferring the need to upgrade servers.

- Improve application performance. Improved application performance translates into faster application response times, which, in turn, translates into higher employee productivity and lower labor costs.

- Smooth CPU spikes. Scheduling administrative tasks and low-priority batch jobs for normally slow periods can spread CPU utilization more evenly throughout the day. Because, in order to protect application response times, it is necessary to size servers sufficiently to handle peak loads, shaving the peaks and offloading some of the work to the valleys allows you to execute the same total volume of work with smaller servers.

- Reduce storage costs. Storage is cheap these days, but it's not free. System management tools that offer practical options for reorganizing files and archiving obsolete data, spool files and logs can reduce storage utilization. In addition, for the reasons stated in this white paper, some of these storage optimization tasks can also serve to improve application performance.

- Address issues. When problems occur you need ready information to help you solve them. You need answers to questions such as: What jobs ran or failed? How long did they run? What subsystems were active and when? Were system values changed? If so, by who, when, and what was the old value? System management tools can provide those answers.

- Plan for growth. Because considerable lead times are required before new equipment can be acquired and installed, you can't wait until your systems and/or storage are overwhelmed before doing something about it. System reporting tools that provide an accurate, easy to interpret picture of system utilization trends allow you to overcome approaching challenges before they become crises.

- Improve system health. A comprehensive set of system management tools can help you to ensure the overall, continuing health of your System i server. In addition to optimizing system performance, this can also help to ensure your systems' reliability and robustness.

- Easily satisfy SLAs. Today, many companies offer customers SLAs as part of sales agreements. Often, the ability to comply with those SLAs is dependent on the continuous availability of high-performing computer applications. Because failure to comply with an SLA can trigger large penalties, keeping systems healthy is critical to the profitability of many companies.

Work Smarter

Although we are all aware of the benefits of a healthy system, fulfilling the daily needs of business application development, maintenance, operation and support usually takes precedence over the performance of system management and optimization tasks. As mentioned in the introduction to this white paper, the greatest impediment in the way of maintaining a healthy System i environment is normally not a lack of skills or knowledge, but a lack of time, which forces monitoring and tuning tasks to be deferred or ignored. To overcome the time impediment it is important to find and implement tools that will automate these tasks as much as possible.

On the other hand, sometimes the reason for a lack of optimal System i health is a lack of knowledge about a particular optimization technique. System management tools can overcome this challenge as well by encapsulating the necessary knowledge and then automating the application of that knowledge.

Because of the time-freeing and knowledge-augmenting benefits automation can offer, whenever you evaluate a system management tool it is important to consider how much of your currently manual workload the tool will automate. All other things being equal, the tool that goes the furthest to reduce the IT workload offers the greatest value.

Individual system administration and optimization functions described herein can be delivered by a number of separate tools from a variety of vendors, but to optimize IT staff productivity, they should all be provided by an integrated toolset from a single vendor. That way, system managers will need to learn only a single user interface, switching between tasks will be seamless and the functionality across the toolset is much more likely to interact in valuable ways.

The Benefits of Optimization

The bottom line in this discussion is that optimizing your System i server is necessary in order to, at a minimum, improve system performance and reduce storage utilization. But for optimized operation to become the normal state, and not just an occasional event, system optimization must be done regularly, consistently and correctly every time. Done manually, it takes considerable time, research and knowledge to perform and, even in the most skilled hands, is still subject to human error.

Automated software tools that ensure continuous, programmatic maintenance and optimization of your system i can deliver substantial benefits. An initial reduction in data storage requirements can sometimes justify the investment in the tools by itself, by allowing you to defer upgrading your storage capacity or by breaking through a performance bottleneck that is hindering your business. But it is important to understand the long-term benefits of continuous optimization. In other words, it is valuable to improve performance, but even more valuable to remain at peak performance, reaping the perpetual operational and financial benefits of continuously heightened productivity and cost efficiency.

Automation is the only path to achieving this continuous state of optimum System i efficiency. It can only be achieved when critical system administration and management tasks run repeatedly, programmatically and in an unattended manner while users remain active on the system. When this is the case, considerable tangible ROI can be achieved through faster application response times, reduced storage costs and higher IT staff and management productivity.

Easy. Affordable. Innovative. Vision Solutions.

Vision Solutions, Inc. is the world's leading provider of high availability, disaster recovery, and data management solutions for the IBM System i and System p markets. With a portfolio that spans the industry's most innovative and trusted HA brands, Vision's iTERA™, MIMIX®, and ORION™ solutions keep business-critical information continuously protected and available. Complementing Vision's availability offerings, Vision Director™ delivers a highly integrated set of applications that proactively monitors, manages and optimizes System i servers, databases and application environments to help ensure the continued health of System i servers.

Affordable and easy to use, Vision products help to ensure business continuity, increase productivity, reduce operating costs, and satisfy compliance requirements. Vision also offers advanced cluster management, data management, and systems management solutions, and provides support for i5/OS®, Windows® and AIX® operating environments. As IBM's largest high availability Premier Business Partner, Vision Solutions oversees a global network of business partners and services and certified support professionals to help our customers achieve their business goals. Privately held by Thoma Cressey Bravo, Inc., Vision Solutions is headquartered in Irvine, California with offices worldwide.

MGBS inc.
marketing group
business solutions
BUSINESS CONTINUITY SPECIALIST

VISION SOLUTIONS®

IBM Premier Business Partner

iTERA HA **MIMIX** HA **ORION** HA

MGBS Inc • 800-323-5404 • www.mgbsinc.com
sales@mgbsinc.com