WHITE PAPER

Maximizing System Performance to Manage the Cost of IT Operations

Sponsored by: Oracle

Elaina Stergiades Randy Perry

April 2012

IDC OPINION

As enterprises confront growing competitive pressures in today's global market, they increasingly rely on a complex, extensive IT infrastructure to enable the mission-critical business processes required for everyday business tasks. As more enterprise users depend on the applications delivered through these IT environments to achieve business success, any disruption in system availability can have an immediate and direct impact on the critical internal and external metrics of business performance. As a result, maximizing system availability and performance across the IT infrastructure is an enterprisewide initiative, requiring a comprehensive approach to IT system management and ongoing IT operations. IDC research regarding the cost of maintaining system availability has shown the following:

- The cost of downtime can vary significantly across industries and is generally dependent on the application in question. IDC research shows the financial services segment has the highest cost per downtime hour of all the industries in this study, while the public sector has had considerable success implementing measures to achieve fewer downtime hours over the past five years.

- Overall, the results show that organizations are deploying more applications to automate everyday business tasks and at the same time simplifying and standardizing IT infrastructure to maximize system availability. As a result of these efforts, many enterprises have successfully reduced the number of unplanned downtime hours that can adversely affect ongoing operations.

- Industrywide best practices for managing and reducing downtime include proactive support services, IT framework best practices, advanced IT management tools, clustering, and virtualization. Enterprises that implement all of these best practices can realize a significant reduction in the number of annual downtime hours.

- Enterprises should consider leveraging their support services providers for help in implementing policies, procedures, technology, and best practices that can help maximize system availability across the IT landscape. IDC recommends looking for specific details regarding how advanced tools and utilities can avoid problems and speed time to resolution, best practices for optimizing ongoing operations, and system availability strategies that include a comprehensive approach to help align IT and business unit objectives with overall enterprise goals.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

IN THIS WHITE PAPER

This IDC white paper describes the implications of managing the availability of IT infrastructure across the enterprise and looks at the typical cost of downtime for various industries and regions. It highlights the role of support and advanced support services in maximizing system performance, as well as best practices enterprises can consider to minimize downtime. The paper also examines support and mission-critical support offerings from Oracle and how they can be applied to improve system availability.

SITUATION OVERVIEW

In 2012, most enterprises have deployed an extensive, complex IT infrastructure with deeply integrated technologies to enable critical business processes. Advancements in enterprise software and hardware have allowed organizations to modernize many aspects of everyday business tasks, leading to increased employee performance and productivity. Access to advanced tools and real-time data can help enterprises respond quickly to dynamic markets and demanding customers to achieve strategic and competitive advantage.

However, as business units demand top performance from CIOs and IT managers, IT organizations often struggle with maximizing the availability and performance of IT infrastructure. Extensive legacy IT assets, integrated systems and technologies, and the pressure to introduce new solutions make ongoing IT management a difficult task. In this environment, minimizing downtime is critical both for customer-facing systems and for internal applications and systems that enable the enterprise at large.

How System Availability Affects the Modern Enterprise

In most modern enterprises, complex and sophisticated IT infrastructures with highly advanced technology are increasingly deployed to support mission-critical business processes. Employees rely on software and hardware for most day-to-day business tasks, from managing and analyzing sensitive financial information and completing payroll to tracking customer behavior and controlling complex manufacturing operations. Access to these systems on a consistent basis is critical to maintaining top performance across all business units and maximizing business opportunities in the increasingly competitive global marketplace.

Across the enterprise, the value of IT infrastructure is measured in the ability to deliver these mission-critical applications to end users. As a result, system availability can have a direct and potentially devastating effect on internal and external metrics for business success. Depending on the systems and applications affected, unplanned system downtime can lead to lost revenue, decreased employee productivity, reduced customer satisfaction, increased customer churn, slower time to market for new products — and even a declining reputation in the market.

As more and more everyday business processes require an advanced IT infrastructure to drive top performance, maintaining business continuity becomes critical for most IT organizations. However, this is not limited to extensive planning and investment for disaster recovery and business continuity solutions. Maximizing system availability requires a comprehensive approach to system optimization and ongoing IT management to minimize the impact of downtime — especially for IT infrastructures that require global availability. This can typically include an ongoing operational commitment to proactive and preventive support tools and procedures for faster problem diagnosis and resolution.

Finally, with technology supporting more and more mission-critical business processes, the effects of system downtime are no longer limited to the IT organization; therefore, maximizing system availability is an enterprisewide initiative. Improving the relationship and alignment between IT objectives and business objectives is a key part of successfully managing downtime across the IT landscape. Cross-functional teams and processes can be helpful when scheduling planned downtime and when outlining a response plan for unplanned downtime.

Primary Causes of Downtime

Although the causes of downtime vary substantially, system downtime typically results from issues that are relatively common in ongoing IT operations. IDC research shows that the primary causes of unplanned downtime include the following:

- **Server and storage issues.** The nature of these issues is changing as organizations extend the adoption of virtualized computing. Traditionally, software problems were largely a result of version incompatibility, multiple version deployment across a single environment, and inadequate maintenance of patches and updates for current versions. Hardware issues were typically the result of unbalanced loads, storage congestion problems, and backup issues. Today, as organizations adopt highly virtualized and standardized computing environments and wrestle with 40% annual growth in data, more issues have shifted to the storage side. Server and storage issues cause 21% of downtime incidents.

- **Configuration issues.** Unplanned downtime can occur when applications fail due to improper configuration, poor application design, inadequate change management policies, and out-of-date applications. Configuration problems drive 17% of downtime incidents.

- **Network issues.** IT infrastructure problems due to networking failures are common for most enterprises, especially in today's increasingly complex IT environments. Issues are often a result of general network congestion or the performance of workloads supporting large volumes of mobile users. Network issues account for almost 30% of downtime incidents.

- **Client incidents.** As users have migrated from the office to remote and mobile locations, the causes of downtime due to client incidents have changed. For example, incidents due to faulty drives and user error account for less than 5% of downtime today. However, software issues abound, including operating system incompatibilities with hardware and applications (now exacerbated by mobile devices and mobile applications), poor change management, security issues, and the increasing availability and use of freeware. Client incidents account for 31% of downtime incidents.

- **Security.** In today's environment of escalating security threats, an external security breach often leads to extensive downtime. In addition, viruses and malware can affect an entire IT ecosystem quickly and with devastating results.

- **Aging, outdated infrastructure.** The current practice of extending PC, server, storage, and networking infrastructure beyond the three- to four-year optimal life span multiplies the probability of the above-mentioned causes. Aging infrastructure can exacerbate interoperability issues as systems and software become more incompatible and require constant maintenance to maintain reliable service levels. For example, while extending the replacement cycle for servers from 36 to 48 months can help capital cost reduction, it also increases downtime by 24%.

Relationship Between Planned and Unplanned Downtime

For many IT organizations, managing downtime typically means careful management of patches and upgrades, as well as detailed remediation for unplanned downtime. However, managing downtime successfully also requires an understanding of the relationship between planned and unplanned downtime.

In most enterprises, planned downtime is typically scheduled for a number of reasons that can span the IT life cycle, including:

- Minor updates, patches, and backups for existing systems

- Significant upgrades to existing systems

- Major implementations for new hardware and software

Unfortunately, while IT organizations execute these operations during planned downtime, that planned downtime can often lead to unplanned downtime. In addition, significant upgrades and major implementations often lack adequate planning and resources for an efficient deployment. Finally, many enterprises have no backup systems or fail-safe processes that can be put into place when system problems arise during planned downtime.

IDC research has found that the relationship between planned and unplanned downtime tends to be inverse in nature: As the number of planned downtime hours increases, the number of unplanned downtime hours decreases. That is, well-designed planned downtime that is managed efficiently can serve to reduce unplanned downtime that could negatively impact the entire enterprise.

Costs Associated with System Downtime

The cost of downtime is a product of internal and external end users of business applications being denied access to systems required for ongoing business operations. Examples include internal users unable to access email, interruption of internal business processes, or external users unable to complete a business transaction. Unplanned downtime can have a number of economically costly ramifications for the business, including the following:

- **Revenue loss.** The cost of unplanned downtime is felt most directly and immediately in revenue-producing, customer-facing systems. As businesses increasingly automate internal operations and customer interfaces, revenue loss due to downtime has become critical to industries far beyond financial services. Retail operations increasingly rely on not just point-of-sale (POS) systems but also in-store self-service systems and online etail. As government organizations significantly reduce offices and personnel to cut costs, they increasingly rely on customer-facing systems to collect fees and taxes. Lost revenue per hour can range from a few thousand dollars to millions of dollars.

- **Reduced end-user productivity.** Users across and beyond the organization, from employees to contractors and partners, rely on IT-delivered services and applications to conduct business operations. Unplanned disruptions can greatly reduce productivity, and for many knowledge workers, downtime means they might as well go home (if they are not working from home already). The cost of reduced productivity is the lost time multiplied by the user's salary multiplied by a factor that accounts for the ability of the user to perform other work-related functions even though the user does not have access.

- **Cost to resolve downtime issues.** The costs of downtime also extend to the personnel and resources required to find, troubleshoot, and fix the issue as well as postmortem analysis. Resources include travel costs and third-party response costs. The drain on productive IT staff time that could be spent on more strategic projects includes the requirement to position additional staff at remote sites primarily to deal with downtime.

- **Customer churn and damage to brand.** Even when downtime occurs in systems that are not directly revenue producing (for example, in customer service or support systems), there can be significant cost as dissatisfied customers take their business elsewhere and poor word of mouth can cause future sales to suffer. The most often cited reason for customers switching telecommunication services providers is excessive downtime.

- **Overprovisioning of resources to compensate.** Many organizations guard against unplanned downtime by overprovisioning server, network, and storage resources; building redundancy into their systems; and keeping pools of hot-swappable hardware on hand. This is expensive, not only in terms of the additional equipment that must be purchased and maintained but also in terms of the additional staff hours required to set up and maintain the resources.

Figure 1 shows IDC estimates of the average cost of downtime in specific industries worldwide. The figure shows the following data:

- Revenue loss per hour of downtime (vertical axis)
- End-user productivity loss per user per hour of downtime (horizontal axis)
- Average number of annual downtime hours (size of circle)

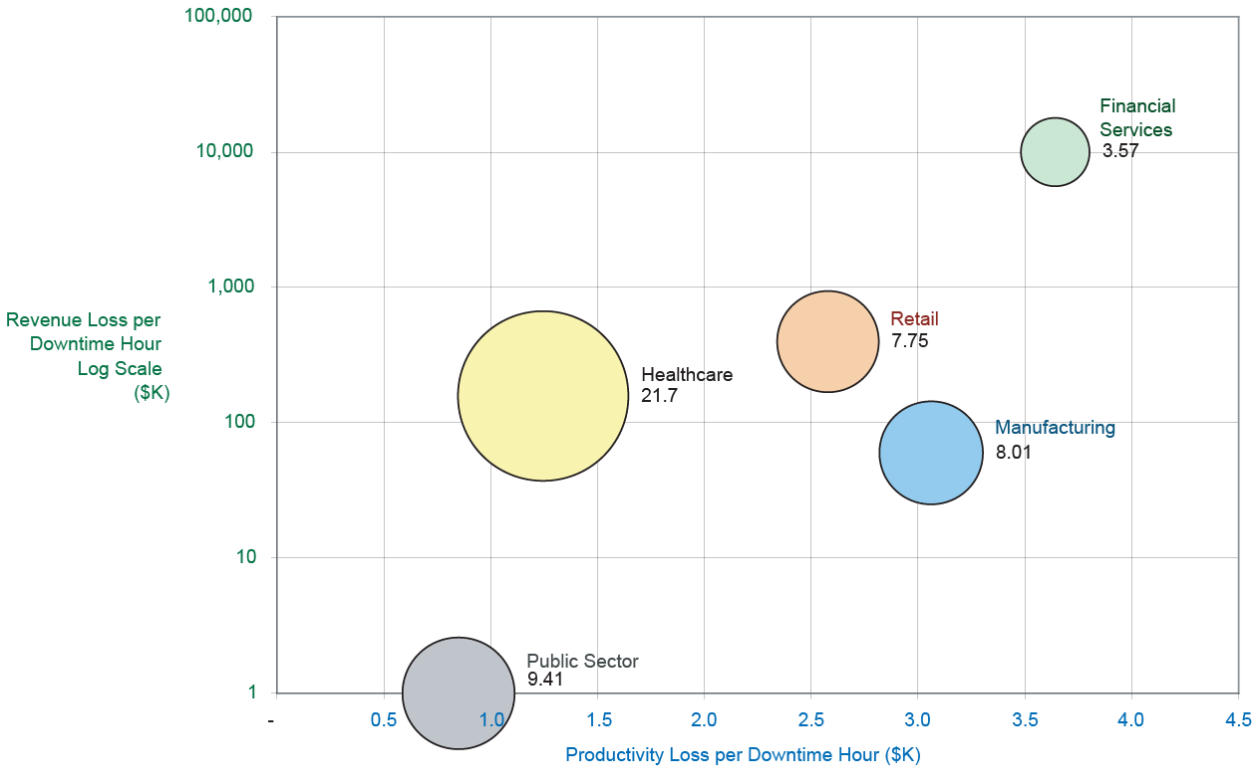
According to IDC Business Value data, the financial services industry experiences the greatest financial liability from downtime with nearly $10,000 revenue loss per hour and $3.64 in lost productivity per hour. The revenue loss reflects high-value custom applications such as trading floor applications for which 30 minutes of downtime costs well over $1 million. Productivity loss is reflective of the relatively higher salary structure. As a result, financial services firms dedicate significant resources to maintaining high reliability in their infrastructure; thus they have the lowest amount of downtime at 3.6 hours per year.

At the other end of the spectrum is healthcare, with relatively low $157 revenue loss per hour and $1.25 productivity loss per hour and the highest amount of downtime at 21.7 hours per year. This model is rapidly changing as time constraints and compliance issues drive physicians to become IT dependent and hospitals move to virtualized client infrastructure.

The public sector suffers the least productivity loss per hour and no revenue loss (by definition), and its 9.4 hours annual downtime hours are about on par with the annual downtime hours of the retail and manufacturing industries. The public sector has made the most improvements in reducing downtime over the past five years.

FIGURE 1

Average Annual Revenue Loss, Productivity Loss, and Downtime Hours by Industry



	Productivity Loss / Hour	Revenue Loss / Hour	Downtime Hours
Financial	$3.64	$9,997.50	3.57
Retail	$2.58	$397.50	7.75
Healthcare	$1.25	$157.50	21.70
Manufacturing	$3.06	$59.93	8.01
Public Sector	$0.85	$ .00	9.41

Note: The size of the circle representing the industry denotes the average annual downtime for that industry.

Source: IDC Business Value Research, 2011

A tidal wave of change is transforming the worldwide market for highly available servers as IT infrastructure evolves to become fully virtualized and to improve availability for workloads running on virtual machines. Drivers include the need to ensure end-user access to critical workloads and to preserve business continuity through the high availability of applications running across the server infrastructure.

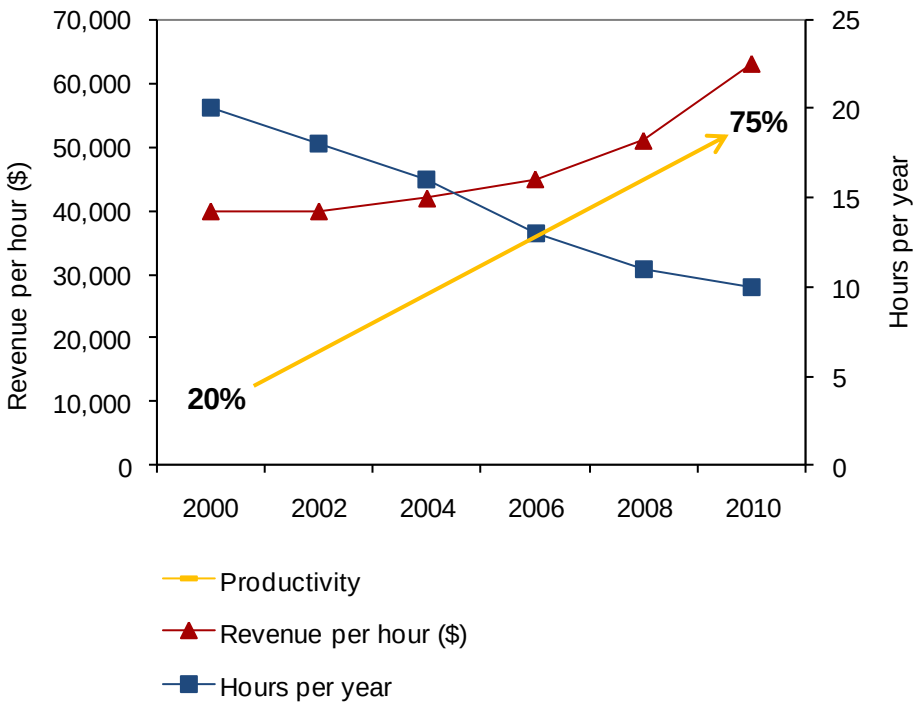
A variety of technologies and best practices are involved in this process. First, IT managers must leverage technology features such as component redundancy, memory protection, multipath I/O, and data replication to ensure that data requests have alternate resources for processing and multiple ways to access outside resources such as storage and networks. Then, they must make sure they have dual power grid feeds for the datacenter, backup power, and multiple connections to the outside world via multiple high-speed networking connections.

Trends in Downtime

Two of the top drivers for IT strategy today are automating business processes to reduce costs and create efficiencies and mitigating risk and variability in business operations. While the first trend has pushed rapid growth in new applications and mobile systems, the second has spurred equally rapid growth in virtualization, datacenter consolidation, and private cloud solutions. As a result of these efforts, many enterprises are having success in reducing the number of unplanned downtime hours that could adversely affect ongoing operations. Figure 2 includes additional details.

FIGURE 2

Trends in Downtime



Source: IDC, 2011

Despite these recent successes addressing the current causes of unplanned downtime, the accelerated pace of change in technology means constant change in IT operations. To minimize future disruptions, CIOs and IT managers should be vigilant in the following areas:

- **Big data.** With the explosive growth of data storage needs, processes and policies to manage storage will be more critical than ever. With data growing at 40–45% annually, the requirements to create, manage, replicate, back up, and store data can often create significant degradations in IT operations.

- **Application virtualization.** While server virtualization continues apace, IDC anticipates a significant lag in application virtualization. As a result, organizations are delaying virtualization, continuing to run their most custom and complex workloads in unvirtualized environments, and/or running these applications in virtualized environments. In all of these cases, new opportunities for application failures are created.

- **Automated decision making.** Enterprises are increasingly adopting business analytics tools to enable automated decision making. As these tools drive improvements in product development, manufacturing, and supply chain operations, the negative impacts of potential downtime are escalated as well.

FUTURE OUTLOOK

As IT environments become increasingly heterogeneous and complex, the process of monitoring and supporting day-to-day IT operations is a careful exercise of reducing costs, maximizing availability, improving staff productivity, and meeting internal service delivery requirements. The ability to manage and reduce downtime is a critical element of these initiatives, especially with increased focus on total cost of ownership and return on investment for IT assets over the life of the IT solution.

IDC anticipates that enterprises will value services that can integrate across the IT environment to lower operating costs, most substantially by improving overall system availability through careful management of planned and unplanned downtime. Additional benefits of these services include more optimized IT infrastructures, more efficient IT environments, and the ability for IT to anticipate and respond to future requirements of the business as markets and customers change.

Best Practices to Optimize Business Continuity

Business leaders today face a high level of uncertainty due to the global nature of their business, government regulation, and changing competitive models. In this atmosphere, they become naturally risk averse and seek to mitigate uncertainty in any way they can. Business continuity becomes more important because it can be measured and, to a degree, managed. Organizations are initiating best practices to reduce unplanned downtime in their core business; ensure smooth, reliable operations; and protect the investments they have made in ebusiness.

Key best practices include:

- Deploying virtualization technologies; standardizing on fewer server and storage platforms; and converging storage, computing, and networking systems to simplify and consolidate the infrastructure to reduce causes of downtime.
- Using systems and network monitoring tools to proactively monitor and manage performance. Monitoring and resolution solutions from point products to integrated platforms enable IT staff to be proactive in identifying potential downtime issues and to quickly troubleshoot and resolve problems.
- Adopting best practices such as ITIL, CobiT, etc. Enterprises can also adopt standardized policies and practices throughout the organization to support operations continuity.
- Maximizing the opportunity for proactive and preventive support services. The tools, utilities, and procedures required for proactive and preventive support can be adopted as part of standard IT processes for problem avoidance, as well as faster issue diagnosis and resolution. Advanced proactive support services from an established support provider can deliver high performance and availability and free up IT staff to focus on more direct business-related activities.
- Implementing failover clustering and other real-time backup and data recovery strategies.
- Deploying deduplication technologies to optimize data footprint. The rapid growth of data is one of the biggest threats to business continuity. Deduplication technologies enable IT to better manage the costs and streamline the demand caused by storage growth.
- Implementing an extensive application testing program. As businesses automate, the rate of new application introduction accelerates and the management of applications requires more IT assets. Application testing reduces the downtime caused by new application incompatibilities with server and client operating systems and other applications.
- Refreshing technology every three to four years. Regular refreshing of server, client, networking, and storage systems reduces the incompatibility issues caused by running multiple software versions on hardware not optimized for the software.

Oracle Support Programs

The Oracle Support Services portfolio includes support and advanced support services that feature deliverables specifically designed to help enterprises improve system availability across the organization. The primary offerings are detailed in the following sections.

Oracle Premier Support

Oracle Premier Support is Oracle's baseline support package, available for software, systems, and operating systems for consistent coverage across products in the Oracle stack. Although the specific features vary according to product coverage, Oracle Premier Support includes access to the following deliverables targeting system availability:

- Health checks and patch recommendations, including detailed patch configuration validation and testing utilities
- Upgrade advisors and upgrade planners, featuring a wizard-driven approach with specific steps and measures that walk customers through the upgrade process
- Advanced remote diagnostic tools that allow collaborative problem solving between customers and Oracle support engineers
- Reporting tools focused on identifying and managing known issues before they affect system performance, including diagnostic data capturing for full service request reporting
- Streamlined service request processes through configuration sharing and automated service request capabilities (currently available for Sun Systems only)

Oracle Premier Support also includes access to the Oracle communities, which feature product news, support features, and product-specific forums. The portal also features industry-specific communities, including legal, utilities, insurance, financial services, and retail.

Customers can benefit from Oracle Premier Support in several ways. "We use the knowledge base extensively and take advantage of self-service technical information, which saves us hours a day," said Chandan Ghosh, senior database administrator, Exel. "In addition, My Oracle Support's personalized configuration manager capabilities cut down the time required to submit a service request by 70% because the system already has critical information on Exel's environment and logs. It also has helped to improve — by approximately 30% — the productivity of IT team members when managing support requests, since Oracle's analysts can look at a ticket and get started rather than coming back to Exel multiple times to clarify request details."

Oracle Advanced Customer Support Services

Oracle Advanced Customer Support (ACS) Services is a global business unit within Oracle Support focusing on mission-critical support for complex IT environments. Oracle ACS Services provides expertise in supporting the Oracle stack through a strategic partnership between Oracle and the customer. Specific ACS offerings with capabilities to maximize system availability include the following:

- **Solution Support Center.** This service features a dedicated support team, proactive guidance, and preventive services across the entire Oracle IT stack. The Solution Support Center features an advanced support delivery manager, as well as a team of advanced support engineers who can work both onsite and remotely to provide 24 x 7 tailored support for the customer's IT environment.

- **Advanced Monitoring and Resolution.** Monitoring and resolution services are provided across the entire IT landscape, from servers to applications. This service can deliver real-time proactive notification of potential issues using a diagnostic tool set to resolve critical issues that could impact system reliability or cause unplanned downtime.

- **Production Optimization Services.** Oracle engineers target improved efficiencies and optimizations for existing IT infrastructures, including strategies for ongoing measurement and performance tied to business value.

- **Production Readiness Services.** This portfolio of services for the entire Oracle stack focuses on installing, configuring, and testing systems based on Oracle best practices.

- **Advanced Support Engineers.** Oracle support engineers have expertise in supporting every phase of the Oracle technology life cycle. ACS advanced support engineers can assist with problem resolution, implementation issues, or optimization processes in the IT environment.

Customers find that Oracle ACS Services can help them maximize system availability. "Oracle ACS has contributed to our success by making the application available. In addition, even when we had issues, they ensured that the system returned to service as quickly as possible," said Bryon Rickey, director of production operations, AT&T. "ACS has helped AT&T reduce unplanned downtime. We work with them well in advance, a few months in advance, before any major releases. We work with them very closely and partner with them to develop deployment and support plans. It has been very successful."

OPPORTUNITIES/CHALLENGES

As part of its extensive services portfolio, Oracle has the opportunity to build on its history of investment and expansion in proactive and preventive support services to maximize system availability across the IT stack. With the integration of My Oracle Support and Enterprise Manager, Oracle has demonstrated a commitment to building a comprehensive suite of IT management tools that span IT operations and support services. In addition, the "single stack provider" model for the Exa family of products could pose interesting opportunities for minimizing downtime. IDC anticipates that with deep engineering and technology expertise at all layers of the solution, Oracle will expand proactive and preventive support capabilities to minimize system downtime.

As mentioned earlier, the Oracle Advanced Customer Support Services offerings are focused on ongoing optimization and advanced support for ACS customer mission-critical IT environments. As such, ACS engagements provide a key opportunity for Oracle to try new methods of customer engagement for support services. IDC recommends that Oracle continue to roll out new utilities and tools focused on reducing downtime through the ACS offerings, which can help refine and expand these features before they are introduced to a larger customer base. This step is critical for new technology focused on reducing downtime to help achieve scale and scope quickly.

In addition, IDC believes that Oracle can leverage its growing online My Oracle Support Communities presence to include considerable focus on how planned downtime can be used to reduce unplanned downtime. Enterprises struggling to improve IT processes often lack the resources required for extensive projects targeting problems such as system availability. Using the My Oracle Support Communities to spread best practices and recommendations around reducing downtime in smaller, incremental steps can help these customers achieve better results without additional resource commitment.

However, Oracle will also face some challenges in the market as its services evolve over time. With the increasing focus on the Exa family of products and the "single stack" strategy, Oracle could miss partnership opportunities around services targeting downtime reduction. As IT organizations increasingly look outside for help managing downtime, they are likely to consider a variety of sources for assistance — from managed services providers to network providers to local support providers. Without the partnerships required to deliver integrated support services, Oracle could be left out of initiatives in which customers are targeting downtime in their Oracle environment.

In addition, IDC believes that new technologies that can help improve system availability must be easily implemented and utilized by IT staff. Most enterprise IT organizations are facing fewer resources and increasing responsibilities in their daily tasks, making it very difficult to implement and learn new technologies — even for vendors that have been long established in their IT landscape. Any new solutions that require significant investment in procurement, implementation, or training will face an uphill climb in terms of adoption and utilization, which will leave the potential benefits unrecognized throughout the market.

Finally, IDC recommends that Oracle work closely with enterprise IT customers to help align business managers and IT managers as part of initiatives focused on system availability. Discussions focused on downtime should always include both the IT organization and the business units because system availability is an issue throughout the enterprise. Aligning the downtime initiatives can help mitigate any potential issues up front and also demonstrates tangible benefits across the IT and business landscape. The downtime management discussion now extends far beyond IT managers and CIOs, and the processes and policies around system availability need to be updated to reflect that important change.

CONCLUSION

The enterprise IT environment will continue to grow in size and complexity over the next five years, and enterprises will increasingly look to advanced technologies such as virtualization, business analytics, and the mobile enterprise to gain strategic competitive advantage. The introduction of additional enterprise software for mission-critical business processes will result in sharper focus on improving the performance and reliability of that software to enable business process continuity. To better manage and reduce overall system downtime, enterprises should take advantage of support services features and capabilities that focus on industrywide best practices for IT management. To better meet the needs of demanding customers looking to maximize system availability, Oracle offers a portfolio of support and advanced

support services with specific features and deliverables to manage and reduce downtime across the enterprise. IDC believes these offerings are well suited for customers looking for outside guidance and expertise on how to improve system availability for both the IT organization and the enterprise as a whole.

APPENDIX

Cost of Downtime Methodology

This model was based on data from IDC's Business Value Database with information collected over the past five years from over 3,000 companies in 43 countries and over 25 industries. From this total data set, we selected the data associated with ERP from 100 companies in North America and 50 companies in EMEA.

The cost of downtime comes from both internal and external downtime and includes three elements:

1. Internal business process, which is the cost to the business to have ERP processes delayed and disrupted and includes costs to redo work or fix errors

2. Business revenue loss, which is the delay or loss of revenue resulting from ERP downtime (To be able to combine this value [top line] with business process loss [bottom line], we account for only the lost operating profit using a 10% margin. For example, a $1 million revenue loss reduction [revenue gain] translates to a $100,000 operating profit.)

3. Cost for IT to fix downtime in IT labor hours

Copyright Notice