R-13-8504
D. Prior, A. MacNeela, I. Brown, J. Krischer, D. Scott, J. Green-Armytage

Strategic Analysis Report
20 December 2001

Enterprise Guide to Gartner's High-Availability System Model for SAP

Management Summary

Enterprises face many challenges in identifying and understanding all the components of hardware, software and services that are required in a large, mission-critical SAP application. All the components need to be integrated and this leads to a complex infrastructure, even for a small project.

Gartner has assembled the required elements into a model that covers, in detail, the technology and services that such an application needs — server and storage hardware, system software, and consulting and support. Enterprises can use the model to evaluate vendor offerings for SAP projects where high availability is a priority.

As an example, Gartner specified a mission-critical high-availability system and invited Hewlett-Packard (HP), IBM and Sun Microsystems to submit proposals for a suitable system. We then scored their proposals and responses to evaluate them against the criteria in the model.

Gartner conducted this exercise at the end of 2000. The scores would have been different a few months later. By late 2001, all three vendors had introduced new high-end servers. And Sun had changed to Hitachi as its supplier for storage, which would have changed our ratings of its storage capabilities. But the methodology is more important than the result. And the dynamics of the industry mean that leading vendors will frequently change their relative positions in any ranking or selection process.

In this instance, we started by grading the vendors' responses with up to 10 points for each of 32 criteria. The methodology is designed to allow enterprises to weight the scores according to their own requirements so that, for instance, they could give more weight to proactive support services than to operating-system partitioning. Applying representative weightings for the relative importance of the criteria, the maximum possible score was 670, and we scored the vendors as follows: HP 409, IBM 355 and Sun 283.

Scoring performance against detailed criteria is only one of several stages of vendor selection. Enterprises should address their business and IT process issues first. Then they should identify all the elements required by a high-availability system (using this Gartner model for reference). Finally, they should ask prospective vendors key questions about the technologies, consulting and support service proposed.

Gartner intends to apply this model to other leading hardware vendors and, where warranted, update or modify the scores or even the selection criteria depending on shifting technology trends and feedback from enterprises running large, mission-critical SAP systems.

Key issues addressed in this report include:

- How will Unix and midrange platforms be compared and selected through 2006?
- How will enterprises make the most effective use of high-end storage?
- What technologies will be critical to achieve high availability and continuous operations?
- What strategies should enterprises use in achieving, measuring and reporting application availability?

CONTENTS

FIGURES

1.0 Introduction

Selecting the best hardware platform for a particular computer application is a recurring task. In recent years, Gartner has published a considerable amount of research to help enterprises cut through vendor hype and obtain an independent view (see Appendix A). Gartner has also developed a generic application server evaluation model (see *Research Note* R-10-0234) and a storage system evaluation model (see *Research Note* COM-11-9285).

Selecting the platform for a complex enterprise resource planning (ERP) application can be even more important, particularly if the application is mission-critical and calls for high performance and availability. Gartner has published research on ERP platform selection (see Appendix A). However, the current trend is to centralize ERP systems, and this increases the requirements for 24x7 availability and consistent, fast response times. It also means more enterprises are requiring large ERP platforms. Consequently Gartner conducted research in this area and chose SAP's R/3 ERP package as a representative application.

Most SAP R/3 users choose Unix as the operating system for the hardware platform when their projects are mission-critical and require a proven, high-availability platform technology. This technology is proven for complex R/3 systems, but enterprises still face extremely onerous and complex tasks in designing, building and operating a complete high-availability solution for an SAP R/3 production system. The system comprises the following elements:

- Server hardware
- Storage hardware
- System software
- Consulting services
- Support services

The problem is that hardware vendors vary greatly in their abilities to supply a complete high-availability system that optimizes all of these elements together. Gartner's research reveals that enterprises frequently overlook this complexity during their processes of selecting a hardware vendor and, as a result, they often encounter real difficulties during live operations. With this report we hope to help our clients by identifying criteria to use when evaluating and selecting a hardware partner for critical ERP implementations.

Gartner developed the SAP High-Availability System (SAP HAS) model to identify, in detail, all the elements which a typical, real-world, large-scale SAP project requires.

Gartner's SAP HAS model addresses the following key questions:

- In detail, which elements are needed for a complete, mission-critical, large-scale SAP system?
- What levels of SAP application availability are achievable and at what cost?
- What are the single points of failure in each hardware vendor's technology?
- What options are recommended by each hardware vendor as part of its high-availability system?
- How do the track records of the hardware vendors compare in their ability to deliver a proven SAP HAS in the real world?

For the first phase of the development and testing of the SAP HAS model, the leading Unix hardware vendors, i.e., HP, IBM and Sun, were chosen and surveyed. Section 1.1 examines the European market shares of these three suppliers, while Section 2 reviews the constituent causes of system downtime. Section 3 describes the SAP HAS survey process that was utilized. Section 4 documents the SAP HAS model in detail, while Section 5 presents the SAP HAS survey results for the three chosen Unix hardware vendors. Section 6 contains the conclusions drawn from this first phase of SAP HAS research, and Section 7 presents some resulting recommendations for enterprises.

In future phases of Gartner SAP HAS research, the following additional hardware platforms will be surveyed:

- Additional Unix vendors
- IBM's OS/390 platform
- Leading NT hardware vendors

1.1 The General Unix Marketplace

Three companies dominate the market for general Unix systems: Sun, HP and IBM. In 2000 they increased that dominance, leaving only 21 percent of the European market to other vendors compared with 32 percent in 1999 (see Figure 1). Sun increased its lead from 28 percent to 37 percent of the market in 2000. HP's revenue grew only 5 percent during 2000 and its share of the market was virtually static at around 25 percent. IBM increased its share of the market by three points to 17 percent. A substantial part of Sun's success in this marketplace was due to aggressive positioning of its systems for the Internet.

Unix Server Vendors
End-User Revenue, Western Europe

Selected Market	End-User Spending, 1999 ($ in Millions)	Share of Total Spending, 1999 (Percent)	End-User Spending, 2000 ($ in Millions)	Share of Total Spending, 2000 (Percent)	Change in Spending, 1999 to 2000 (Percent)
Sun Microsystems	2,144	27.9%	3,080	37.0%	43.6%
Hewlett-Packard	1,964	25.6%	2,063	24.8%	5.0%
IBM	1,108	14.4%	1,432	17.2%	29.3%
Others	2,462	32.1%	1,743	21.0%	-29.2%
Total	**7,678**	**100.0%**	**8,318**	**100.0%**	**8.3%**

Source: Gartner Dataquest Estimates, 2001

Figure 1. Concentration in the European Unix Market

2.0 Causes of Application Downtime

SAP applications are often mission-critical and downtime can be very expensive. Enterprises should understand the causes of downtime and take steps to prevent it. Not even the most highly available system can completely eliminate it.

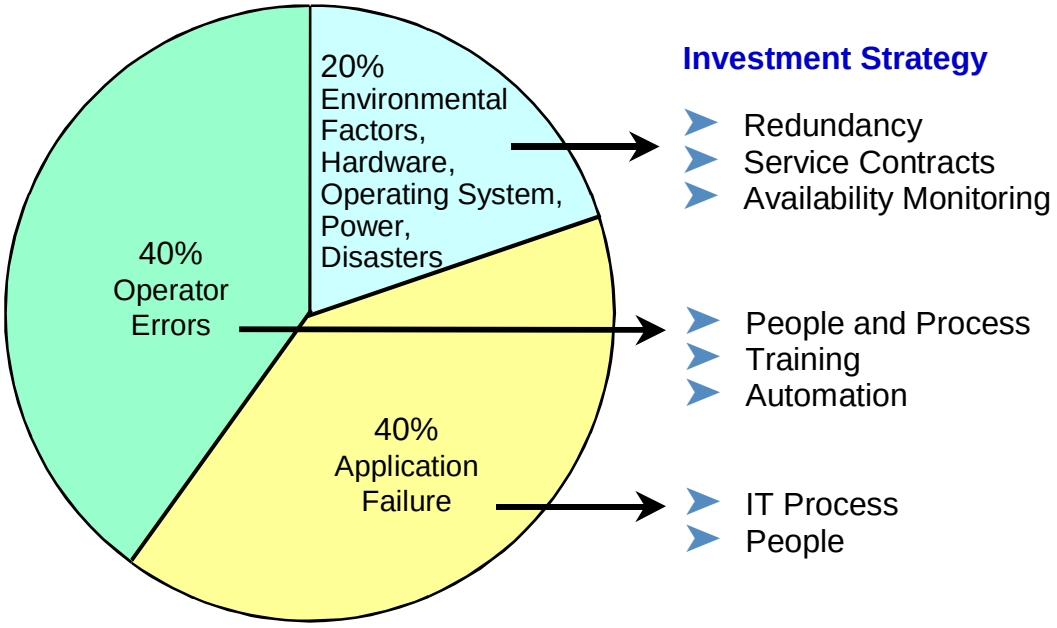
There are two kinds of downtime — planned and unplanned:

- **Planned downtime** is an deliberate interruption in application availability, scheduled for:
  - Hardware maintenance
  - Software upgrades to the operating system, database or application

  - Database maintenance operations such as backup, recovery, reorganization or archiving
- **Unplanned downtime** is an unexpected failure of an application for any reason, including:
  - Hardware failures and disasters
  - Software issues
  - People and process issues

Any of the above factors will reduce availability of the application for business users. Any downtime can mean a possible breach of a business service-level agreement (SLA) for an SAP system.

Clearly, SLAs should be defined with appropriate allowances for planned downtime. As for unplanned downtime, Gartner's research has revealed a range of typical causes, which are shown in Figure 2. Based on extensive feedback from clients, we estimate that, on average, people and process issues cause most of the unplanned application downtime for mission-critical applications.



Source: Gartner Research

Figure 2. Causes of Application Downtime and Appropriate Responses

Only one-fifth of unplanned downtime arises from technological failures or disasters. Technology failures may include hardware faults in servers or network components; operating-system problems; or environmental factors such as heating, cooling and power failures. Disasters include flood, fire and storms.

Two-fifths of unplanned downtime arises from application failures, including "bugs," performance issues or changes to applications that cause problems (including changes to the application code itself or to any layered software on which the application depends).

The remaining two-fifths of unplanned downtime arises from unexpected user behavior or operator errors, such as not performing a required operations task or performing a task incorrectly. For example, changes made to infrastructure components can cause problems and incur unexpected downtime.

Thus, approximately 80 percent of unplanned downtime is caused by people- and process-related issues, while the remainder is caused by technology failures and disasters.

Enterprises seeking to improve application availability should adopt different strategies and investments for each of the three categories of unplanned downtime.

Downtime from technology failures and environmental disasters has risks that can be limited. Enterprises should do the following:

- Monitor components for availability. Identifying a failure is the first step toward resolution. Monitoring is typically done with agents or sensors and, ideally, it should be predictive and warn the operator or vendor of potential failures before they occur.
- Buy vendor service contracts that reduce time to repair. Many vendors offer time-to-repair commitments for increased fees.
- Implement redundancy in the system to ensure alternate processing capabilities in the event of failure. Data mirroring, clustering and diesel generators are examples of redundancies that limit downtime when failures occur. In comparing potential solutions, pay particular attention to how transparent recovery is to end users and applications.

Enterprises can reduce downtime from application failures by investing in, improving and re-engineering IT processes, including the following:

- Change management — reduces unplanned downtime caused by inadequate planning and testing of application changes, and enables a more proactive approach to problem prevention
- Problem management — improves problem identification, isolation and resolution, thereby reducing time to repair
- Configuration management — tracks relationships between dependent application and infrastructure components, and enables better understanding of change impact and quicker fault diagnosis
- Application architecture and design — reduces single points of failure, aids in problem isolation and makes application failures more transparent to users
- Performance management and capacity planning — proactively identifies current and future resource shortages that impact SLAs

Downtime from operator errors can be reduced by:

- Maturing IT operations to a more process-oriented and documented approach that does not require or mandate that specific, knowledgeable people be available to perform tasks
- Hiring competent people and training operators (and vendors) on IT processes and procedures
- Automating processes (e.g., by using job-scheduling and event management tools) to reduce the likelihood of errors
- Improving change and problem management processes related to IT infrastructure and facilities

Given the complexity of an SAP system, achieving high availability involves more than buying redundant hardware. A responsible hardware vendor for such a system must advise an enterprise on the causes of downtime, and on how to minimize each cause to obtain the right level of SAP availability for the business

users of that enterprise. Furthermore, the hardware vendor should demonstrate the cost-effectiveness of each technology option it offers, and relate this to the cost of downtime for an enterprise's business.

3.0 Survey Process

Many enterprises take the likely availability of a system into account when evaluating a proposal from a system vendor. In addition, enterprises should assess a vendor's ability to integrate a highly available enterprise application. All aspects of an offering are relevant to the completeness and effectiveness of a vendor's technology.

With this is mind, Gartner began this research project in the last quarter of 2000. The aim was to analyze the key attributes of a best-of-breed, high-availability system for a large-scale SAP implementation, and to evaluate what vendors could supply according to those attributes. We drafted a specification of requirements to send to selected vendors and defined a schedule of what the vendors should include in their responses.

3.1 Specification Presented

The specification for a representative project was:

- It should be an SAP R/3 production system using release 4.6B.
- It should allow for 3,000 concurrent users of the R/3 system.
- The business workload rating should be 36,000 SAPs, assuming 100 percent CPU utilization. Note that it is customary to express an SAP business workload in units of "SAPs" (see *Research Note* TU-08-0030) at a standard CPU utilization of 100 percent, although real SAP production systems are never sized to operate at more than 70 percent.
- The database should be 850GB.
- Average response time should be guaranteed at less than two seconds in any online window.
- The disaster recovery scenario should be excluded.

3.2 Submission Required

The vendors were required to submit:

- A clear schematic of the recommended solution
- A parts list with prices in pounds sterling
- A list of all software products and tools included in the configuration
- A statement of what consulting services would be supplied
- A list of five reference sites running such a configuration
- A statement of the different levels of overall application availability, with costs for each option
- A description of everything needed to achieve each level of application availability

We then invited HP, IBM and Sun to submit proposals to meet the specification. We chose these vendors because they are leading suppliers of the large Unix systems typically used for SAP applications (see Section 1.1).

3.3 Evaluation Criteria and Review Process

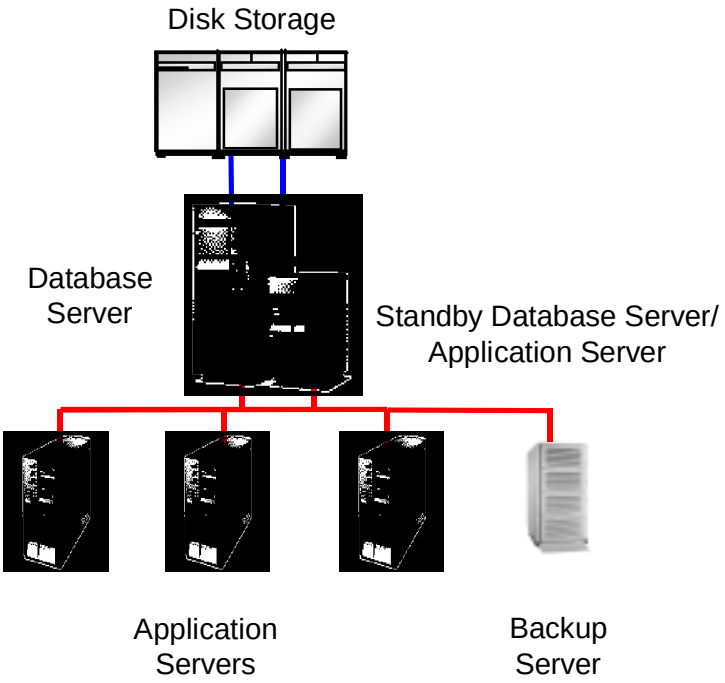
Using the model described in Section 4.0, we identified 32 criteria by which to rate the proposals. These criteria are listed and explained in Section 4.1. We examined and rated the vendors' responses, giving them scores on a scale of zero to 10 for each criterion. This gave each vendor a basic score out of a possible 320 points. Since each enterprise has individual requirements, we then adjusted the scores by multiplying each score by a weighting factor to produce weighted scores. In this case, the weighted scores did not change the relative position of any of the vendors. In other cases, however, weighting can change the overall preferences because different enterprises place different values on the various aspects of a proposal.

We sent our assessments to each vendor for comments. Where appropriate, vendor comments were considered before the final draft of this report was further reviewed by each vendor before publication.

We limited the study to technology that could be delivered immediately and made no allowances for vendor promises of future products. Appendix B reports subsequent developments in hardware.

4.0 Gartner's SAP HAS Model

Conceptually, the hardware configuration for an SAP system is simple. Users access application servers that are linked to a database server, which, in turn, accesses the disk storage system. Typically there is a failover database or application server, and another application server dedicated to backups (see Figure 3).



Source: Gartner Research

Figure 3. General SAP Configuration

Such a configuration is used for SAP R/3 ERP applications and most mySAP application products. The Web infrastructure layer is omitted for clarity. There can only be one database server, so this is a single point of failure that can be mitigated with failover clustering.

The update process usually runs on the database server. Many application servers can be connected to it over a high-speed network. Application servers contain large amounts of memory to buffer user context data. All application servers contain a dispatcher plus multiple work processes for dialog, batch, printing and intersystem communication. When an application server fails, all in-process work for that application server is rolled back. Users can log onto another server, although this increases the burden placed on the remaining servers. Enterprises are wise to plan for excess application server capacity to accommodate occasional outages (planned and unplanned). This also provides additional flexibility for performance tuning and load management.

In an SAP system there are two critical internal processes — the Message process and the Enqueue process — which are combined into a single "central instance." The Message process manages application server coordination; it is "stateless." The Enqueue process manages all application data locking via memory-based tables; it is "stateful." If there is a failure in the server running the Enqueue process, all in-process transactions from every application server will have to be re-entered, and in-process batch jobs will need to be resubmitted. Recovery from such a failure can be problematic.

It is standard practice when designing an SAP Unix production system to cluster the R/3 database server to eliminate the database server as a critical single point of failure. Typical implementations are for failover of the main SAP database server to a dedicated standby database server, to an application server in the same production system or to another SAP system, such as one usually used for quality assurance or testing. However, this clustering can be complex and difficult to implement and manage. Typical failover times range from 10 to 20 minutes depending on the amount of "in-flight" data that needs to be rolled back. (Oracle Parallel Server is not a standard option for SAP systems.)

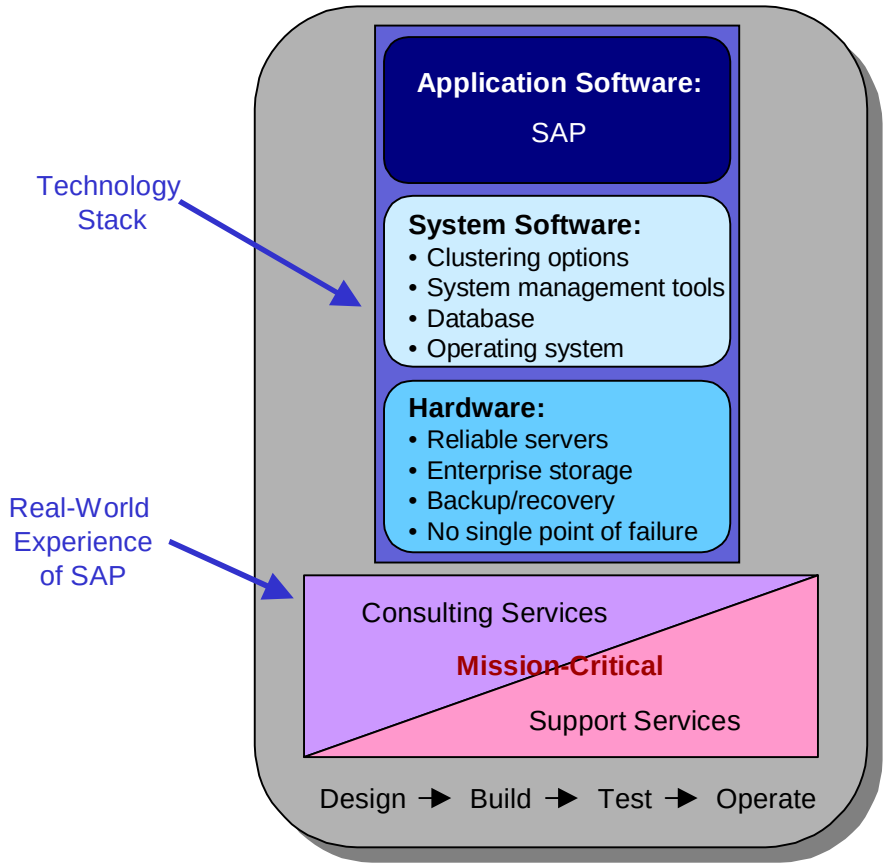
The configuration in Gartner's SAP HAS model makes no provision for disaster recovery either for a corrupted database (that can't be recovered by a failover cluster) or for site outage. Disaster recovery was excluded for simplicity, but it must be addressed in business-critical SAP projects. Some discussion of disaster recovery service offerings is included in the analysis of criteria related to system support services (see Section 4.1.5).

A hardware configuration is only a small part of a complete system for an SAP HAS. Such a project requires complete integration of hardware, system software and the SAP application software in a technology stack. In addition, enterprises will require consulting and support services derived from practical experience. All are required as enterprises step through the processes of designing, building, testing and operating the system. Gartner combines all these elements into its SAP HAS model (see Figure 4).

Each project typically has to be treated as a one-off for the complex tasks of integration. Nearly always it is the hardware vendor that has the necessary expertise to manage this integration.

For consulting at the design and build stages, enterprises should use only SAP Basis consultants that have been formally certified by SAP. Once the production system has been built and tested, it must be managed by the enterprise (or outsourcer), and the complete system infrastructure must be supported by the hardware vendor. Special support services from the hardware vendor for SAP projects may be needed to achieve high levels of overall SAP application availability.

Enterprises whose business can justify the highest levels of SAP availability will require that the right technology, consulting and support options be included in their projects.



Source: Gartner Research

Figure 4. Gartner's SAP HAS Model

4.1 Criteria for Evaluating an SAP HAS

Gartner identified 32 criteria for evaluating offerings suitable for a high-availability SAP system. These criteria covered hardware and software technologies, support services, consulting services and the business practices of the vendors. The criteria are explained in Sections 4.1.1 through 4.1.7. The scores we assigned to vendors for each criterion are presented in Section 4.2.

4.1.1 Technology: Database Server Availability

- *Database server clustering options:* This refers to the ability to protect a live R/3 system against a failure in the R/3 database server, which is a single point of failure.
- *Clustering adoption for SAP, Europe:* Does the vendor know how many of its SAP installations in Europe use a clustered database server?
- *Speed of system recoverability:* This is the time required for a live R/3 application to recover from a failure of the R/3 database server.

4.1.2 Technology: Application Server Availability

- *SAP Enqueue process protection:* The Enqueue Service has two main components: the Enqueue Server and the Enqueue Table. R/3 Enqueue process protection is the mechanism provided, if any, to remove a single point of failure.

- *Hardware availability:* This criterion refers to the application servers' standard high-availability features. These include features such as memory protection — e.g., error-correcting code (ECC), memory scrubbing and memory de-allocation — as well as hot-swap components and subsystems (e.g., disks, power supplies and cooling).
- *Online upgradability:* Vendors have started to add functionality to their operating systems, which enables the operating system to recognize that new hardware has been added and then to add it to a running configuration. This capability is often called "dynamic reconfiguration."
- *Operating-system availability:* As well as designing the components so that they can be easily removed and replaced, vendors also have to ensure that operating systems and system software are aware of such removals and replacements and have routines to manage them.

4.1.3 Technology: Disk Storage Subsystem

- *Availability:* Some systems require scheduled downtime for reconfiguration.
- *Performance:* This criterion examines average performance and throughput for different workloads.
- *Functionality:* This refers to the number of platforms with which the storage subsystem can be used; the effectiveness of storage management and load-balancing facilities; and the availability of tuning features.
- *Scalability:* This refers to the maximum capacity of the system and its cache facilities; the number of Fibre Channels the system can support; and its ability to sustain high data transfer rates.

4.1.4 Technology: System Manageability

- *Server manageability:* This refers to the low-level server management tools and utilities included as part of the base operating system or specified for this SAP HAS configuration. Typically, this means a system manager based on a graphical user interface (GUI) is used to complement the Unix command line interface (CLI).
- *Operating-system partitioning:* Partitions enable a single server to be divided into a number of "logical" servers, each running its own operating-system instance and software stack. In some cases these are "hard" partitions, i.e., the partition configurations are determined by hardware boundaries.
- *Cluster management:* How well does the cluster management software integrate with other system management tools? Does it have a global file system for all the nodes in the cluster? Can all nodes be managed as though they constituted one integrated system?
- *Capacity on demand (COD) functionality:* Does the server support COD? COD refers to the possibility of buying a server configured with more processors or memory than are initially activated.
- *Storage management:* This is defined as the ability to back up and recover a complete copy of the online storage.

4.1.5 Services: System Support

- *Response time — standard:* This is the normal response time for hardware or software maintenance or remediation.
- *Response time — optional:* This refers to enhanced response time for hardware or software maintenance or remediation.

- *Proactive software patch management:* How proactively does the vendor manage the software-patching process?
- *Proactive service:* This examines proactive system support services, often characterized by remote monitoring/control and management tools.
- *Assigned support team:* This is the dedicated support team for a named customer.
- *Accredited escalation process:* This criterion examines the ISO-9000-accredited escalation procedure.
- *Integration with SAP's escalation process:* This is the integration of the hardware vendor's escalation process with SAP's escalation process for rapid resolution of difficult support problems.
- *Disaster recovery service option:* This is defined as the range of vendor business continuity services available to ensure that an enterprise can consistently recover from a site disaster.
- *Disaster recovery service adoption, Europe:* Does the vendor know how many of its SAP installations in Europe use some form of disaster recovery service?

4.1.6 Consulting: Design and Build

- *Proven SAP methodology:* This criterion refers to a vendor's proven methodology for delivering repeatable consulting services for installing a high-quality SAP infrastructure.
- *Pre-packaged SAP services:* These are the pre-packaged, fixed-price SAP consulting services delivered within this methodology.
- *Information Technology Infrastructure Library (ITIL) problem and change management services:* This criterion examines adherence to ITIL business processes and definitions for IT problem and change management. The ITIL is an initiative developed by the Central Computing and Telecommunications Agency consultancy for the government of the United Kingdom. It offers a set of best practices in 24 service delivery and IT service support areas, including help desk, problem management, change management, software distribution and cost control.

4.1.7 Business Practices

- *Responsiveness:* This refers to the vendor's ability to quickly supply the information requested by Gartner in this survey.
- *Completeness of solution response:* This examines the completeness of the information supplied by the vendor for its SAP solution to the survey requirements.
- *System reliability in the field:* The hardware and software proposed by the vendor is examined in terms of its reputation for reliability, as discerned by Gartner analysts based on client feedback.
- *Support reliability in the field:* Vendor support services are examined in terms of their reputation for quality, as discerned by Gartner analysts from client feedback.

4.2 How We Rated the Vendors

The vendors' scores for each of the 32 criteria are summarized in Figure 5, and are presented in more detail in Figure 6, Figure 7, Figure 8 and Figure 9. These scores are, of course, somewhat subjective. But Gartner has discussed them extensively to reach a consensus among the analysts concerned. The reasons for the scores are given, vendor-by-vendor, in Section 5.

Technology		IBM	Sun	HP
Database Server Availability		45	33	51
Application Server Availability		14	21	25
Disk Storage Subsystem Scores		50	21	63
System Manageability		39	51	42
Services				
System Support		102	84	114
Consulting				
Design and Build		36	31	37
Business Practices				
Total Score for Business Practices		69	42	77
Totals		**355**	**283**	**409**
See Figures 8 and 9 for weightings assigned to each criterion within a section.				

Source: Gartner Research

Figure 5. Summary of Weighted Scores

	Scores: 1-10 (10 = market leader)		
Vendor	**IBM**	**Sun**	**HP**
Operating System	AIX 4	Solaris 7	HP-UX 11
Disk Subsystem	IBM/ESS	A3500	XP512
Technology: Database Server Availability			
Database server clustering options	6	4	6
Clustering adoption for SAP, Europe	5	3	7
Speed of system recoverability	6	6	6
Subtotal	*17*	*13*	*19*
Technology: Application Server Availability			
SAP Enqueue process protection	1	1	6
Hardware availability	5	7	5
Online upgradability	1	6	3
Operating-system availability	6	6	5
Subtotal	*13*	*20*	*19*
Technology: Disk Storage Subsystem			
Availability	8	5	9
Performance	7	3	9
Functionality	7	2	9
Scalability	6	2	9
Subtotal	*28*	*12*	*36*
Technology: System Manageability			
Server manageability	4	5	3
Operating-system partitioning	1	7	2
Cluster management	5	3	5
COD functionality	2	5	6
Storage management	7	2	4
Subtotal	*19*	*22*	*20*

Source: Gartner Research

Figure 6. Basic Scores on Technology

	Scores: 1-10 (10 = market leader)		
Vendor	**IBM**	**Sun**	**HP**
Operating System	AIX 4	Solaris 7	HP-UX 11
Services: System Support			
Response time — standard	5	5	5
Response time — optional	5	5	6
Proactive software patch management	1	1	1
Proactive service	7	3	8
Assigned support team	5	3	7
Accredited escalation process	5	4	6
Integration with SAP's escalation process	6	6	6
Disaster recovery service options	8	7	8
Disaster recovery service adoption, Europe	5	5	6
Subtotal	*47*	*39*	*53*
Consulting: Design/Build			
Proven SAP methodology	7	7	7
Pre-packaged SAP services	4	5	5
ITIL problem and change management services	6	4	6
Subtotal	*17*	*16*	*18*
Business Practices			
Responsiveness	6	3	7
Completeness of solution response	6	5	7
System reliability in the field	7	3	7
Support reliability in the field	6	4	7
Subtotal	*25*	*15*	*28*

Source: Gartner Research

Figure 7. Basic Scores on Services

	Vendor	**IBM**	**Sun**	**HP**
	Operating System	AIX 4	Solaris 7	HP-UX 11
	Disk Subsystem	IBM/ESS	A3500	XP512
	Weighting Factor	**Weighted Scores**		
Technology: Database Server Availability				
Database server clustering options	***3***	18	12	18
Clustering adoption for SAP, Europe	***3***	15	9	21
Speed of system recoverability	***2***	12	12	12
Subtotal		*45*	*33*	*51*
Technology: Application Server Availability				
SAP Enqueue process protection	***2***	2	2	12
Hardware availability	***1***	5	7	5
Online upgradability	***1***	1	6	3
Operating-system availability	***1***	6	6	5
Subtotal		*14*	*21*	*25*
Technology: Disk Storage Subsystem				
Availability	***2***	16	10	18
Performance	***1***	7	3	9
Functionality	***3***	21	6	27
Scalability	***1***	6	2	9
Subtotal		*50*	*21*	*63*
Technology: System Manageability				
Server manageability	***2***	8	10	6
Operating-system partitioning	***3***	3	21	6
Cluster management	***2***	10	6	10
COD functionality	***2***	4	10	12
Storage management	***2***	14	4	8
Subtotal		*39*	*51*	*42*
Note: Weighted scores are the basic scores in Figure 6 multiplied by the weighting factor.				

Source: Gartner Research

Figure 8. Weighted Scores on Technology

	Vendor	IBM	Sun	HP
	Operating System	AIX 4	Solaris 7	HP-UX 11
	Weighting Factor	**Weighted Scores**		
Services: System Support				
Response time — standard	***1***	5	5	5
Response time — optional	***2***	10	10	12
Proactive software patch management	***2***	2	2	2
Proactive service	***3***	21	9	24
Assigned support team	***1***	5	3	7
Accredited escalation process	***2***	10	8	12
Integration with SAP's escalation process	***3***	18	18	18
Disaster recovery service options	***2***	16	14	16
Disaster recovery service adoption, Europe	***3***	15	15	18
Subtotal		*102*	*84*	*114*
Consulting: Design/Build				
Proven SAP methodology	***2***	14	14	14
Pre-packaged SAP services	***1***	4	5	5
ITIL problem and change management services	***3***	18	12	18
Subtotal		*36*	*31*	*37*
Business Practices				
Responsiveness	***2***	12	6	14
Completeness of solution response	***3***	18	15	21
System reliability in the field	***3***	21	9	21
Support reliability in the field	***3***	18	12	21
Subtotal		*69*	*42*	*77*
Note: Weighted scores are the basic scores in Figure 6 multiplied by the weighting factor.				

Source: Gartner Research

Figure 9. Weighted Scores on Services

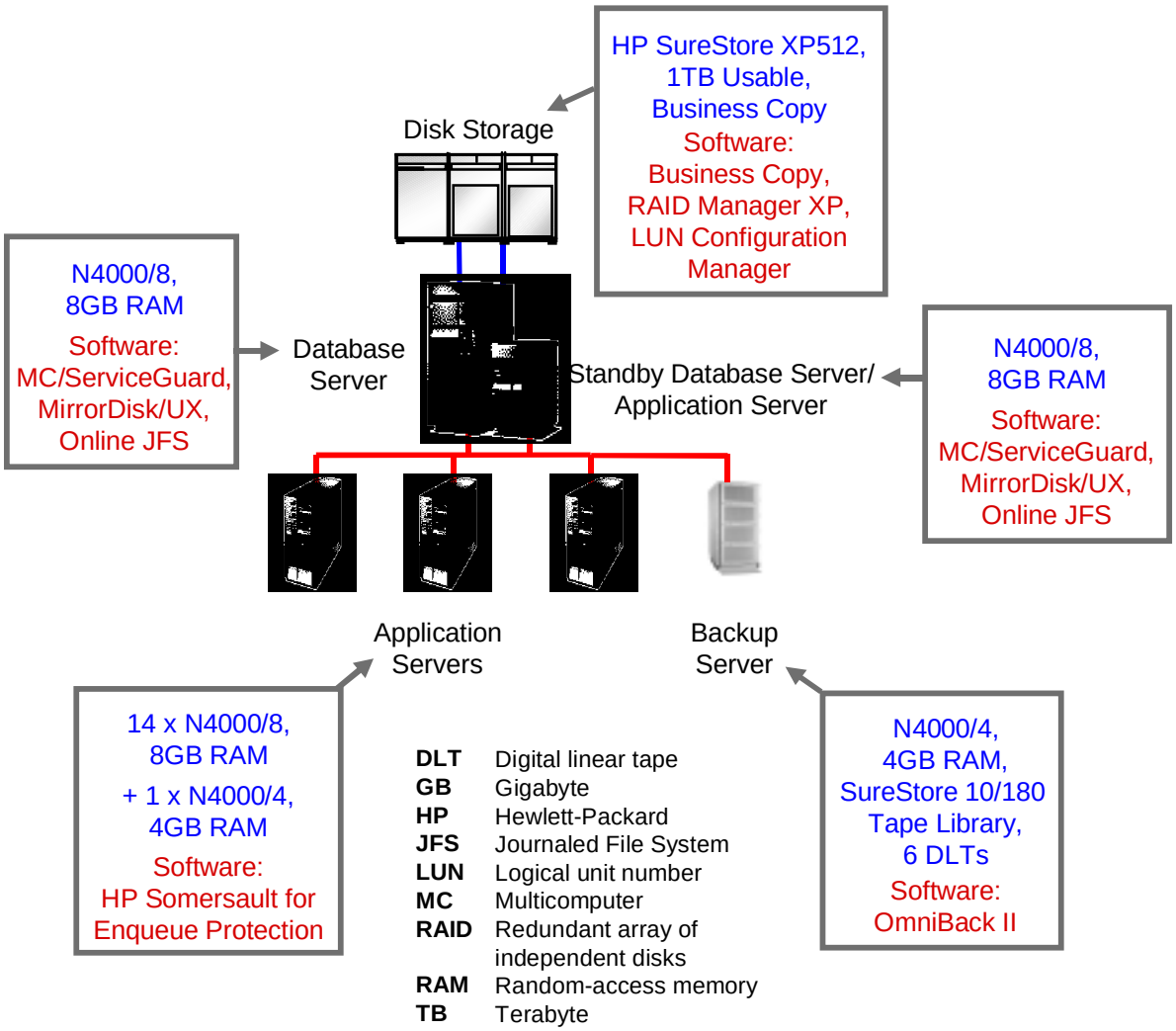
5.0 The Ratings in Detail

For each of the three vendors selected for this stage of our research program, we went through the criteria and assigned scores on a scale of zero to 10. The results were as follows.

5.1 Hewlett-Packard

HP proposed the system configuration shown in Figure 10. The proposal contained a total of 18 N-Class servers from its HP9000 range: a main database server with a failover database and application server, 15 additional application servers, and a backup server. For storage, HP specified its own XP512 enterprise storage disk technology.

This system configuration and the services supporting it achieved the highest total score in this survey. Overall, HP scores (unweighted) totaled 193 out of 320. These scores derived from the following considerations.



Source: HP and Gartner Research

Figure 10. HP's Configuration

5.1.1 HP's Technology: Database Server Availability

Database Server Clustering Options: HP uses MC/ServiceGuard as the clustering software to automate failover to an application server, which acts as a standby database server. It also uses MirrorDisk/UX and Online Journaled File System (JFS). The HP SAP Competence Center in Walldorf, Germany, maintains scripts for using MC/ServiceGuard with SAP systems. There are several clustering options, which are complex but well-proven. *Score: 6 out of 10.*

Clustering Adoption for SAP, Europe: The above database server clustering options for SAP production systems are widely adopted by HP customers in Europe. At the time of the survey, there were over 450 such customers using these options. At least 50 percent of large HP database servers (i.e., not just SAP systems) implement MC/ServiceGuard. Mission-critical operation of HP9000 systems has been a long-time HP strength. *Score: 7 out of 10*.

Speed of System Recoverability: Gartner believes that, among the three hardware vendors, there is little difference in the speed of recovering both the system and the SAP application itself in the event of a failover of the database server using the clustering software preferred by the vendor. *Score: 6 out of 10*.

5.1.2 HP's Technology: Application Server Availability

SAP Enqueue Process Protection: HP included its Somersault product to protect the Enqueue and Message processes as a systemwide single point of failure. Somersault is unique to HP. It provides process-mirroring system software in memory to protect the Enqueue process from failure of the R/3 application server running it (for further details, see Appendix A). If this server fails, a paired application server resumes the Enqueue process responsibility, so that batch and online processing is uninterrupted. R/3 kernel patches may be applied on a live system if Somersault is installed. The Message process is installed on the database server and protected by its clustering. Somersault is bundled with HP consulting services to design, build and test it. Somersault currently only works with R/3 release 4.6 and very few customers are using it.

HP's offering was based on 16 application servers, providing plenty of room for load balancing and redundancy between servers. The potential drawback is that management tasks are significantly increased. It also means that the application servers make up a significant part of the overall cost of the installation. *Score: 6 out of 10*.

Hardware Availability: N-Class hardware has the usual redundant and hot-swap components, such as power supplies, cooling fans and PCI adapters. However, redundant network interface controllers (NICs) are not supported on the application servers as configured. MC/ServiceGuard does support failover and alternate pathing for I/O controllers, but this was not specified on the application servers.

Surprisingly, the N-Class doesn't have hot-swap system disks either. Each server supports a pair of mirrored system disks with HP's MirrorDisk/UX software to manage them (add-on software again). However, MirrorDisk/UX does support redundant disk controllers with load balancing.

Other hardware-based high-availability features include memory scrubbing, ECC protection on memory and data paths. Memory chip de-allocation is supported through Dynamic Memory Resilience (DMR) in HP-UX 11.0. DMR scans memory for single-bit errors with the goal of de-allocating memory pages before double-bit errors occur. Page de-allocation attempts to allow the system to keep running even though an uncorrectable error may have occurred in the user space. In this case, only the affected jobs are terminated, though in terms of the overall aim of application availability it is pretty catastrophic. If the system is re-booted or an auto-restart takes place, the system automatically re-boots around de-allocated memory.

As well as detecting failing memory, the system's Event Monitoring Service (EMS) can also potentially detect a failing CPU and de-allocate it (unless it's the "monarch" processor on which the operating-system kernel processes are running). The Dynamic Processor Resilience (DPR) feature in HP-UX 11.0 migrates all processes currently scheduled on that processor to active processors. If a processor crashes, DPR similarly de-allocates it and the system will re-boot around it. Such features contribute to the overall reliability of the server: It keeps running but with degraded performance. Our biggest criticism of the HP servers is that, although the high-availability functionality is generally good, many of the features are supported through add-on options, which add to the cost of the system as specified. *Score: 5 out of 10*.

Online Upgradability: N-Class servers are single, nonpartitionable servers. If CPU boards, operating systems or new versions of the application need to be installed, the server must be taken down. This is not catastrophic, because the multiple applications servers and R/3 load balancing mean that the application keeps running, although performance will be degraded. With eight CPUs on board, the N-Class servers are at full capacity. Adding capacity will disrupt operations. If faster processors are required, for example, the servers will have to be taken down one by one — there is no smooth, nondisruptive way

to add capacity or upgrade the existing servers. The time required for physical upgrades will impact service levels. Ideally, enterprises should perform scheduled maintenance when there's not too much pressure and the workload can be balanced among the servers. However, managing such maintenance is complicated because administrators will have to deal with so many servers. This is both time-consuming and prone to error. *Score: 3 out of 10*.

Operating-System Availability: Administrators will have to go through a similar process, server by server, if new applications or patches to system software have to be tested and implemented. HP-UX will support live patch updates, but these should be tested before installation on a live machine. HP-UX supports dynamically loadable kernel modules so that, for instance, device drivers can be added and modified online. However, concurrent operating systems cannot be supported on a single server. To update the operating system, the server has to be taken down. However, HP is developing virtual partitioning for HP-UX. This will enable administrators to set up separate test and production software stacks on the same server. *Score: 5 out of 10*.

5.1.3 HP's Technology: Disk Storage Subsystem

Availability: HP offered its XP512 storage subsystem, which is based on Hitachi Data Systems' (HDS') Lightning 9900. It features full redundancy, "call home" capability, nondisruptive upgrades and maintenance, and online microcode updates. Data cache is fully mirrored for writes. Control parameters are stored in separate mirrored control memories. The XP512 supports a wide selection of data replication options, including HP SureStore E Business Copy XP, Continuous Access XP and the asynchronous extension. *Score: 9 out of 10*.

Performance: There is now enough client feedback to compare the XP512 with EMC's 8000 series and IBM's ESS F20. The XP512 generally delivers what was promised and, in most situations, the performance of any of these systems will be more than adequate. *Score: 9 out of 10*.

Functionality: HP has improved the usability of the XP512 and differentiated it from HDS' Lightning 9900 through a combination of microcode and software developments. These software enhancements include CommandView, development of a Web-based control program, a consistent look and feel for all management functions (which simplifies training), graphical status monitoring, visualized host/device/file/LUN mapping, OpenView integration, path failover load balancing software (AutoPath XP), and certification and integration with software packages from companies such as Oracle and Veritas. *Score: 9 out of 10*.

Scalability: The XP512 was specified with one terabyte, but can be configured with up to 24 terabytes of disk and up to 32 gigabytes of user data cache. XP512 packaging allows users to install an initial footprint and add expansion frames as the subsystem grows. This avoids the costs and operational disruptions associated with box swaps and encourages subsystem growth. *Score: 9 out of 10*.

5.1.4 HP's Technology: System Manageability

Server Manageability: HP's configuration of some 18 servers would take some managing. System management involves people, processes and tools. HP's management tools have been criticized for being poorly integrated, but that's improved recently. The ServiceControl suite of products for HP-UX covers three areas: workload management, fault management and configuration management. ServiceControl now includes a single-point, multisystem management tool called ServiceControl Manager. By integrating tools such as the Ignite/UX system image tool, the GUI-based System Administration Management tool and the EMS tools with ServiceControl Manager, HP has been able to improve HP-UX's multisystem manageability.

ServiceControl Manager enables multiple servers to be managed from a central point — a crucial feature in this case. Its role-based security, where the concept of roles is used to authorize access to specific tools on specific nodes, is designed to reduce human error (which Gartner estimates accounts for about 40 percent of unplanned system downtime; see Section 1.1). This means that senior administrators can have greater authorization than junior administrators or operators. Similarly, integrated software distribution (SD/UX), the ability to keep a "golden image" for all servers in a cluster (Ignite/UX), and the ability to retain and compare system configuration details (System Configuration Repository) are all designed to improve availability, by reducing the time to bring systems back to a stable state and configuration. HP-UX 11 also features a kernel-logging feature, which enables support engineers to identify the root causes of system problems more easily.

One downside of the HP-UX tools, however, is that most are add-on products — they're not included as part of the base operating system. For this application, HP has specified additional system management tools such as GlancePlus, MirrorDisk/UX and Online JFS. Licenses for each server push up the cost of the overall solution. The ServiceControl suite and ServiceControl Manager software may be very effective, but like most aspects of HP-UX, they're expensive. Given that they automate key functions, make it easier to bring systems back to a stable state, and limit access to authorized personnel, we'd advise clients to make those investments. *Score: 3 out of 10*.

Operating-System Partitioning: N-Class servers do not support hard system partitions. HP introduced "virtual partitioning" (vPars) into the HP-UX 11i version of its operating system in 4Q01. HP's vPars will be able to support separate individual instances of the operating system and application stack on the same server-like hard partitions. They will also allow more flexibility in being able to dynamically reallocate resources (CPUs and memory) to different partitions — i.e., they'll enable resources to be removed or added to running partitions without having to stop them. The vPars capability is new and unproven, however. Until it's available, stable and proven, we cannot evaluate it in our comparisons here. Until then, hard partitioning with dynamic reconfiguration capabilities is the best that can be hoped for in the Unix world. *Score: 2 out of 10*.

Cluster Management: General manageability of the individual nodes in an HP-UX cluster can now be handled from the central point provided by ServiceControl Manager. Scripting and management of specific failover cluster services using HP's MC/ServiceGuard remains complex, but HP's SAP Competence Center can provide scripts for MC/ServiceGuard and SAP. Again, the investment may be long and complex to set up, but overall, we believe it's worth it. *Score: 5 out of 10*.

COD Functionality: Unlike IBM and, until recently, Sun, HP provides COD on its midrange servers. This could be significant as a nondisruptive means of adding CPU capacity, providing that the server is not already fully utilized. However, on the configuration specified here, there isn't much additional capacity available. All the servers are N-Class and only support up to eight CPUs and 32GB. In its proposed configuration, HP suggested that the application servers had to be eight-way machines, but the database servers only had to be six-way machines, leaving some room for a little extra capacity. *Score: 6 out of 10*.

Storage Management: The functionality of HP's logical volume manager isn't as sophisticated as some. In addition, HP's OmniBack products for backup and recovery are not as sophisticated as those from Veritas and Tivoli. *Score: 4 out of 10*.

5.1.5 HP's Services: System Support

Response Time — Standard: Like IBM and Sun, HP offers a standard support service response time of between four and eight hours from receipt of a call. *Score: 5 out of 10*.

Response Time — Optional: Like IBM and Sun, HP offers a standard support service response time of between four and eight hours for response from receipt of a call. Similarly, HP also offers optional, enhanced response times, most often characterized by a "24x7x4" approach, meaning service would be available 24 hours a day, seven days a week, with a four-hour response time. HP was rated slightly higher than others due to its Critical Systems Support (CSS) program, which offers an immediate response and a call-to-repair commitment of less than six hours. *Score: 6 out of 10*.

Proactive Software Patch Management: Like all hardware vendors, HP does not offer comprehensive software patch management as a standard support feature. It must be purchased as an additional item. *Score: 1 out of 10*.

Proactive Service: HP, through its relatively expensive CSS program, offered the most comprehensive proactive service suite, including proactive notification, change planning and its High Availability Observatory service. This service provides a secure connection to HP's mission-critical support center, providing automated system diagnostics and configuration information collection on supported hardware, software and network devices. *Score: 8 out of 10*.

Assigned Support Team: HP scored highly, at 7 out of 10, through its CSS program, which provides an assigned local account support engineer and an HP response center account advocate, both trained in SAP R/3. Additionally, HP customer engineers maintain a profile and knowledge of the customer's R/3 environment. *Score: 7 out of 10*.

Accredited Escalation Process: HP's mission-critical escalation processes are ISO-9000-registered. In addition, HP's escalation procedures enable local HP management to coordinate problem escalation, ensuring quick resolution and response organization. *Score: 6 out of 10*.

Integration With SAP's Escalation Process: HP is an SAP global services partner and hardware partner. Accordingly, HP has a high degree of integration with SAP's escalation processes. *Score: 6 out of 10*.

Disaster Recovery Service Options: Although out of the scope of this detailed analysis, we credited HP for offering complete disaster recovery outsourcing services plus specific expertise related to SAP recovery implementations, including using data replication technologies to shorten recovery times in the event of a disaster (e.g., Oracle Standby Database and disk-based remote copy). *Score: 8 out of 10*.

Disaster Recovery Service Adoption, Europe: A significant number of HP's SAP customers across Europe utilize these disaster recovery services. At the time of the survey there were at least 180 such customers with HP. *Score: 6 out of 10.*

5.1.6 HP's Services: Consulting Services

Proven SAP Methodology: All three respondents scored equally (at 7 out of 10) on this criterion, with no discernable differentiation in the methodology used to professionally design, build and commission the server infrastructure for an SAP project. All vendors will accept turnkey responsibility for this critical project activity. The organization that delivers these services is called Hewlett-Packard Consulting. *Score: 7 out of 10*.

Pre-packaged SAP Services: All three respondents had a wide range of pre-packaged SAP services, but HP scored slightly higher than the others because of the range of standard, fixed-price, packaged SAP services described by standard documentation. *Score: 5 out of 10*.

ITIL Problem and Change Management Services: HP scored 6 out of 10 because it has practices and services based on ITIL methodology (the use of which is considered to be a best practice in problem and change management), plus service management in general. HP uses this methodology when implementing SAP solutions. *Score: 6 out of 10*.

5.1.7 HP's Business Practices

Responsiveness: HP was generally quick to respond to queries and helpful in assisting Gartner in completing this survey. *Score: 7 out of 10*.

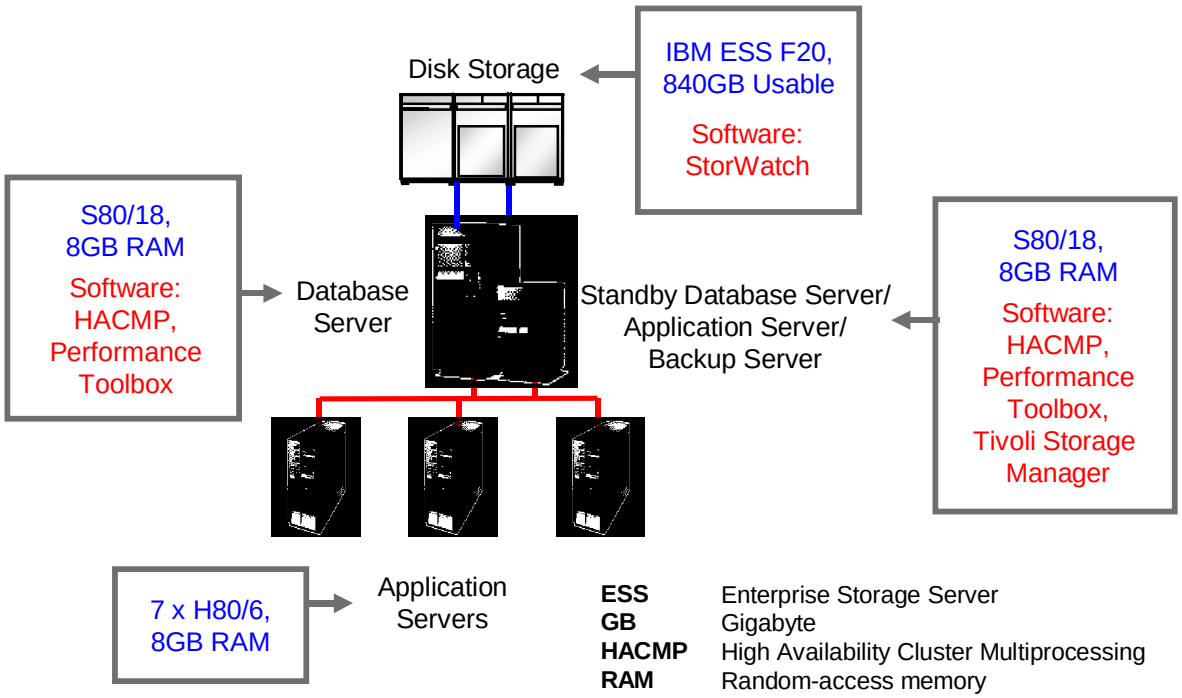
Completeness of Solution Response: HP provided extensive details on all SAP HAS model elements within a single integrated proposal, with full backup information where appropriate. *Score: 7 out of 10*.

System Reliability in the Field: Based on extensive feedback to Gartner analysts from clients around the world, HP servers score well on system reliability in the field. *Score: 7 out of 10*.

Support Reliability in the Field: Worldwide Gartner client feedback on HP's support services is also very good. *Score: 7 out of 10*.

5.2 IBM

IBM proposed the configuration shown in Figure 11. The system is based on S/80 and H/80 Class servers from the RS/6000 (pSeries) range. For storage, IBM proposed its own F20 ESS enterprise storage disk technology and 3583 Ultrium tape libraries containing two Linear Tape Open tape drives.



Source: IBM and Gartner Research

Figure 11. IBM's Configuration

This system configuration and the services supporting it achieved a reasonable score in this survey. Overall, IBM scored (unweighted) 166 out of 320.

5.2.1 IBM's Technology: Database Server Availability

Database Server Clustering Options: IBM used its High Availability Cluster Multiprocessing (HACMP) product to automate failover to an application server, which acts as a standby database server and backup server. Like all clustering solutions, these are complex but well-proven. *Score: 6 out of 10.*

Clustering Adoption for SAP, Europe: The above database server clustering options for SAP production systems are adopted quite extensively by IBM customers in Europe. At the time of the survey, IBM did not know how many of its SAP customers were using these options. *Score: 5 out of 10.*

Speed of System Recoverability: Gartner believes that there is little difference among the three vendors in the speed of recovering both the system and the SAP application itself in the event of a failover of the database server, using the clustering software preferred by the hardware vendor. *Score: 6 out of 10.*

5.2.2 IBM's Technology: Application Server Availability

SAP Enqueue Process Protection: IBM has no means of protecting the critical Enqueue and Message processes, other than to include them within the database server cluster. This can result in lost transactions in the event of a server failure. *Score: 1 out of 10.*

Hardware Availability: Like HP, IBM also proposed a multiserver solution for the application servers. It proposed seven RS/6000 H80 servers. IBM's AIX 4.3.x operating system includes Workload Manager (WLM) software, so multiple instances of the SAP application servers can be run on a single server and the workload can be balanced to guarantee appropriate resource levels for different applications. The rest of the system consists of a pair of clustered IBM RS/6000 S80 servers for the database server, with a 2105-F20 disk subsystem, attached via Fibre Channel. All the servers run IBM's AIX Unix operating system (AIX 4.3.3).

The IBM H80 servers are midrange, rack-mounted servers. They support up to six 500MHz processors, and IBM proposed these maximum configurations of processors. The H80s would each be configured with 8GB of memory. High-availability features include redundant NICs for both server-to-server and client attachment, and an uninterruptible power supply for the application servers. Neither of these are standard offerings, but power and network connections are more susceptible to failure, tampering or problems outside the jurisdiction of system administrators than other aspects of the system. Therefore, we advise clients to consider these as worthwhile investments and to include them in specifications for the system. Clients should note that automation of the failover between NICs requires additional HACMP software to be installed on the servers which, like HP, IBM did not propose.

A key feature of the RS/6000 servers is the integral service processor and its role in monitoring hardware components for both hard and soft failures. The system monitors hardware and operating-system parameters against predefined performance limits and is claimed to offer protection against hard and soft failures of CPUs, memory I/O buses, power supplies, internal and external disks, disk adapters, networks, network adapters, cooling units and input power for all components. IBM claims the system will attempt to de-allocate failing CPUs gracefully. If a CPU fails, it is logged and taken down for an automatic restart. Multibit memory errors can be detected, but only single-bit memory errors are corrected. Again, the failed memory chips are logged and omitted from the re-boot process. The system can also be set to do a "reduced power on self test and re-boot" routine to speed up recovery. *Score: 5 out of 10.*

Online Upgradability: H80 Class RS/6000 servers are single, nonpartitionable servers. If CPU boards, operating systems or new versions of the application need to be installed, the server has to be taken down. This is not catastrophic as the multiple applications servers and R/3 load balancing mean that the

application keeps running, although performance will be degraded. With six CPUs on board, the H80 Class servers are at full capacity. Adding capacity will disrupt operations. If faster processors are required, for example, the servers will have to be taken down one by one — there is no smooth, nondisruptive way to add capacity or upgrade the existing servers. The time required for physical upgrades will impact service levels. Ideally, enterprises should perform scheduled maintenance when there's not too much pressure and the workload can be balanced among the servers. However, managing such maintenance is complicated because administrators will have to deal with so many servers. This is both time-consuming and prone to error. *Score: 1 out of 10.*

Operating-System Availability: AIX is a solid operating system with a good record for reliability. IBM did not specify any special system software in addition to AIX 4.3.3 for the application servers. The operating system itself carries out system monitoring and "trend" data collection. AIX also incorporates disk mirroring and a journal file system as standard. AIX 4.3.x supports dynamically loadable kernel modules for the online addition and modification of drivers. The diagnostic and monitoring features are reasonably comprehensive and combine with the onboard service processor to handle remote offline diagnostics. AIX tends to have more features built into the operating system as standard, and is consequently less expensive than HP-UX.

Like the midrange HP servers, the RS/6000s do not support hard partitions or dynamic reconfiguration capabilities so, apart from being able to hot-add Peripheral Component Interconnect (PCI) controllers, hardware upgrades cannot be undertaken without taking the server down. However, software patches can be applied online, although some may require a re-boot. Concurrent operating-system levels can also be loaded onto different disks and loaded at the next re-boot. If the system needs to be rolled back to the previous operating system, a second re-boot will take it back to the original operating system. *Score: 6 out of 10.*

5.2.3 IBM's Technology: Disk Storage Subsystem

Availability: IBM offered its ESS subsystem, the first in the Seascape family of cross-platform products. This uses RS/6000 cluster technology with an operating system based on AIX. The cluster consists of two symmetric multiprocessing (SMP) units with four processors each. On the back end, ESS uses Serial Storage Architecture loops to connect the disk drives. ESS is not a breakthrough in disk technology, but a solid subsystem that should be evaluated using conventional criteria. It descends from Versatile Storage Server, and past experience suggests that the hardware will be solid but that nondisruptive hardware maintenance and microcode updates must be market-validated. In case of a microprocessor failure, the ESS will perform internal "cluster failover" and will block access to data until it has completed this process. After recovery, the host connections of the failing "side" will be lost. *Score: 8 out of 10.*

Performance: For a few months, ESS had a leadership position in performance, only to lose it to EMC's Symmetrix 8000 in April 2000, which lost it two months later to Hitachi's Lightning 9900 (comparable to HP's XP512) product. *Score: 7 out of 10.*

Functionality: IBM launched ESS toward the end of 1999 without many important features. By the end of 2000, however, IBM could deliver native Fibre Channel support for AIX, Windows NT/2000 and NUMA-Q platforms, S/390 FlashCopy, Peer-to-Peer Remote Copy, and FlashCopy support for major Unix and Windows NT platforms. ESS is a viable system in high-availability, heterogeneous environments. *Score: 7 out of 10.*

Scalability: The ESS can be configured with up to 11.2 terabytes of disk storage and is upgradable in groups of eight 36-gigabyte drives. *Score: 6 out of 10.*

5.2.4 IBM's Technology: System Manageability

Server Manageability: Like HP, IBM proposed multiple application servers. But its seven servers would be less onerous to manage than HP's 16.

IBM was one of the first Unix vendors to add a graphical front end to its system management software. The System Management Interface Tool provides both X-Windows and Java-based GUIs, as well as a CLI. IBM also scores above HP and Sun because features such as disk mirroring and Web-based remote system management of multiple machines are native to AIX. Other AIX strengths include automatic device configuration at boot time, the ability to install new versions of the operating system and maintenance fixes while the current operating system is running, and centralized installation administration via the Network Installation Manager. However, AIX 4.3.3 cannot match all of the fault and configuration management features for managing multiple systems that are now available with HP's ServiceControl Manager. To maintain its lead in future evaluations, IBM will have to improve AIX's multisystem management capabilities. Another objection is that AIX does not support the full set of Unix System V commands. Gartner advises clients to use tools that assist with multisystem installation and deployment, asset tracking, tracking of operating-system releases, configurations, and patch management. *Score: 4 out of 10.*

Operating-System Partitioning: IBM does not support partitioning on its high-end SMP servers. IBM's AIX operating system does include WLM software, which offers policy-based CPU and memory prioritization. WLM gives administrators greater control over the allocation of CPUs and memory to processes. It is equivalent to HP's Process Resource Manager (PRM) and Sun's Solaris Resource Manager. However, all workload management tools are not yet well-proven. AIX's WLM is a no-cost feature of the operating system, whereas HP's PRM is an add-on with associated cost. In the SAP configuration reviewed here, WLM could be used to ensure that workloads on the application and database servers are not starved of resources at critical periods, such as during batch jobs. *Score: 1 out of 10.*

Cluster Management: IBM's HACMP clustering software is well-respected and widely adopted within the IBM AIX SAP user base. IBM has added many manageability features to HACMP over the years, such as automatic replication of configurations across the cluster and SNMP support. *Score: 5 out of 10*

COD Functionality: IBM does not offer COD on its midrange H80 servers, although it is now available on the high-end S80 and pSeries 680 SMP servers that would typically be used for the database servers in an R/3 solution. IBM's original proposal specified 18-way RS/6000 S80 database servers with 32GB of memory. The S80 and pSeries 680 both support a maximum of 24 CPUs. COD configurations are available, however, which mean that a 24-way system can be bought with only 18 CPUs activated and paid for. Additional CPUs can subsequently be activated and paid for as capacity requirements increase. These upgrades take place without the delays imposed by having to call out a system engineer and have system boards physically installed, so there's an advantage in availability terms. However, there will be a price premium to pay for COD functionality.

Meanwhile, the seven H80 application servers do not offer COD and are fully configured with six CPUs each to meet the capacity requirements of our SAP installation. Additional capacity for the application servers would have to be added by "scaling out" with additional servers. While R/3 can handle the load balancing across the additional servers, installing, configuring and maintaining them would be a further management burden. *Score: 2 out of 10.*

Storage Management: IBM's Tivoli Storage Manager software is considered one of the best in the industry in terms of backup and recovery functionality. *Score: 7 out of 10.*

5.2.5 IBM's Services: System Support

Response Time — Standard: Like HP and Sun, IBM offers a standard support service response time of between four and eight hours from receipt of a call. *Score: 5 out of 10.*

Response Time — Optional: All of the vendors offered enhanced response times as an optional extra, most often characterized by a 24x7x4 approach. IBM was rated slightly lower than HP, whose CSS program offers an immediate response and a call-to-repair commitment of less than six hours. *Score: 5 out of 10.*

Proactive Software Patch Management: Like all the hardware vendors, IBM does not offer comprehensive software patch management as a standard support feature. It must be purchased as an additional item. *Score: 1 out of 10*

Proactive Service: IBM provides dedicated support teams and in some cases on-site support engineers. However the service is not quite as comprehensive as HP's. Hence, IBM was rated 7 out of 10 in this important attribute. *Score: 7 out of 10.*

Assigned Support Team: IBM provides assigned support teams but does not organize these around an end-user's mission-critical computing environment with the clarity or dedication of HP's approach. *Score: 5 out of 10.*

Accredited Escalation Process: IBM is ISO-9000-registered and was rated 5 out of 10 for its accredited escalation procedure. *Score: 5 out of 10.*

Integration With SAP's Escalation Process: IBM is an SAP global services partner and hardware partner and has a high degree of integration with SAP's escalation processes. *Score: 6 out of 10.*

Disaster Recovery Service Options: Although out of the scope of this detailed analysis, we credited IBM for offering complete disaster recovery outsourcing services plus specific expertise related to SAP recovery implementations, including using data replication technologies to shorten recovery times in the event of a disaster (e.g., Oracle Standby Database and disk-based remote copy). *Score: 8 out of 10.*

Disaster Recovery Service Adoption, Europe: At the time of the survey IBM could not say how many of IBM's SAP customers in Europe used its extensive disaster recovery services. *Score: 5 out of 10*

5.2.6 IBM's Services: Consulting Services

Proven SAP Methodology: All three respondents scored equally well in this section, with no discernable differentiation in the methodology used to professionally design, build and commission the server infrastructure for an SAP project. All vendors will accept turnkey responsibility for this critical project activity. The organization that delivers these services is called IBM Global Services. *Score: 7 out of 10.*

Pre-packaged SAP Services: All respondents had a wide range of pre-packaged SAP services, with IBM scoring 4 out of 10 in this section because of the lower range of standard, fixed-price, packaged SAP services apparently offered and described by standard documentation. *Score: 4 out of 10.*

ITIL Problem and Change Management Services: IBM was rated 6 out of 10 because it has practices and services based on ITIL methodology, the use of which is considered a best practice in problem and change management, plus service management in general. IBM uses this methodology when implementing SAP solutions. *Score: 6 out of 10.*

5.2.7 IBM's Business Practices

Responsiveness: IBM was generally quick to respond to queries and helpful in assisting Gartner in completing this survey. *Score: 6 out of 10.*

Completeness of Solution Response: IBM responded with a reasonably comprehensive solution response. *Score: 6 out of 10.*

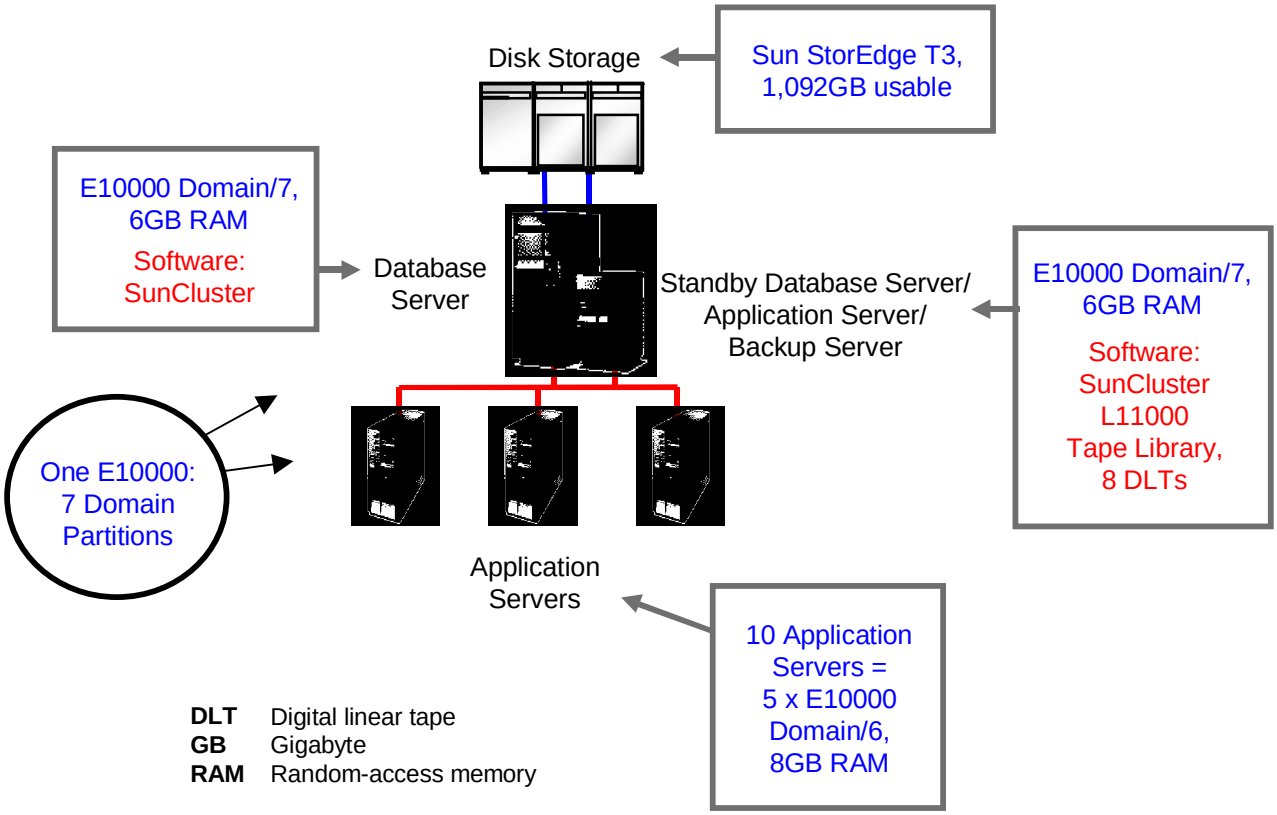
System Reliability in the Field: Based on extensive feedback received by Gartner analysts from clients around the world, IBM servers score well on system reliability in the field. The vertically integrated operating-system stack (including file system and volume manager), plus IBM's test procedures, contribute to overall system reliability. S*core: 7 out of 10.*

Support Reliability in the Field: Worldwide Gartner client feedback on IBM support services is also good. *Score: 6 out of 10.*

European Users: At the time of the study, IBM did not specify the number of R/3 customer installations in Europe using this clustering solution. But Gartner is aware of many such installations. *Score: 5 out of 10.*

5.3 Sun Microsystems

Sun proposed the system configuration shown in Figure 6. The proposal contained a single E10000 partitioned into one R/3 database server domain and ten logical R/3 application servers, which are configured into five physical domains. Sun proposed its own StorEdge T3 enterprise storage technology.



DLT Digital linear tape
GB Gigabyte
RAM Random-access memory

Source: Sun and Gartner Research

Figure 12. Sun's Configuration

This system configuration and the services supporting it achieved the lowest score in this survey. Overall, Sun scored (unweighted) 137 out of 320.

With a single E10000 server, Sun can support a single database/update server, with two system boards containing seven CPUs and a total of 6GB of memory, and five application server domains, each also consisting of two system boards with seven CPUs, but with 8GB of memory. All these logical servers are contained in a single E10000 footprint. Sun recommended that, to isolate failures, *external* network connections should be used to connect the servers. Each domain is equipped with a quad-Ethernet interface card and separate Fast Ethernet networks for internal communications between the database and application server domains, and between the SAP servers and SAP front ends and administrators.

Sun argued that consolidating the SAP application on a single server increased flexibility and improved manageability. While we agree with the benefit of increased flexibility, we do not agree that a single server with multiple partitions is any easier to manage than an equal number of single systems. Administrators could use its Dynamic Reconfiguration feature to reconfigure partitions to imitate different possible combinations of servers: many small four-way partitions, each running application servers, combined with a large database partition; or midsize application servers, like the eight-way midrange servers offered by HP and IBM, and a midsize database partition.

5.3.1 Sun's Technology: Database Server Availability

Database Server Clustering Options: Sun offered its SunCluster clustering software to automate failover to an application server, which acts as a standby database server and backup server. SunCluster for SAP software includes monitoring agents to check the health of the software stack "above" the hardware and operating system. Consequently, there is some monitoring of the application and application-related problems. This means that administrators may be alerted to application problems faster than would be the case if only the hardware and operating system are monitored. Responses may be automated with pre-written scripts for common application problems. However, SunCluster software (v.2.2) is less mature than either HP's or IBM's failover software (see Appendix A), so Sun scores just under its rivals. *Score: 4 out of 10.*

Clustering Adoption for SAP, Europe: At the time of the survey, Sun did not know how many of its SAP customers were using either the SunCluster or Veritas clustering software for SAP production systems in Europe. Hence, it is difficult to determine how widely adopted these options are in practice. *Score: 3 out of 10.*

Speed of System Recoverability: Gartner believes there is little difference among the hardware vendors in the speed of recovering both the system and the SAP application itself in the event of a failover of the database server, using the clustering software preferred by the hardware vendor. *Score: 6 out of 10.*

5.3.2 Sun's Technology: Application Server Availability

SAP Enqueue Process Protection: Sun has no means of protecting the Enqueue and Message processes, other than to include them within the database server cluster. *Score: 1 out of 10.*

Hardware Availability: The availability benefits with the E10000, as proposed, come mainly from the ability to reconfigure domains in the event of hardware problems. If a processor in one application server domain fails, then only that domain goes down — each domain is isolated in both hardware and software terms. In itself, however, partitioning offers no major availability benefits. It's no different from the HP and IBM solutions with their individual midrange servers — there are no additional benefits in terms of automated failovers between partitions with the system as specified.

As for recovery, Solaris has a dynamic reconfiguration utility, which allows an administrator to reconfigure a domain and bring it back up minus any failed processor. This is like the HP and IBM midrange servers,

which can be set to auto-restart in the event of a crash and automatically de-allocate failed processors or memory. However, Sun administrators have to write a script to automate the process of de-configuring the failed processor in the manner this is handled for the HP and IBM midrange servers. The same is true for memory failures.

The major benefit that dynamic reconfiguration brings is the ability to "hot swap" system boards. (System boards include up to four processors, memory and I/O slots.) System boards can be logically detached from an online partition, powered down, replaced and reattached to the partition. By using alternate pathing, connectivity to devices from the de-configured system board can be re-routed via I/O controllers on other system boards and so, although performance is degraded, the domain remains available. In the configuration suggested for the application here, each partition supports two system boards. In the case of a CPU or memory failure, it would be possible to restart the partition with the failed processor or memory chip de-configured. When a service engineer arrived with a replacement board, it would then be possible to hot swap the faulty system board as described above, without having to take down and restart the partition. Although degraded, some degree of partition availability could be maintained throughout the process.

To minimize performance degradation after a processor or memory failure, Sun recommended that the system should be converted to a COD system and that all system boards should be fully populated. This means that each of the domains has two system boards each with four CPUs: eight CPUs per domain instead of either six (for the application servers) or seven (for the database servers). With this configuration, in the case of a CPU failure, the failing domain can be brought back at 100 percent performance. There are no additional time delays for the installation of a replacement system board, because the replacement CPU is already installed in the system. This would be an advantage compared with the HP and IBM application servers, but it has to weighed up against the likelihood of such an occurrence.

Processor failures are relatively rare, but catastrophic when they do occur. Ironically, Sun servers have suffered more than their share of reliability problems. During 2000, a number of Sun E10000 users reported processor failures due to an e-cache error (see Appendix A). Sun subsequently provided various explanations, software patches and workarounds to try and correct this highly disruptive fault. This problem highlighted the immaturity of Sun's support processes at the time. Other customers with this potential problem were apparently not proactively alerted by Sun. We understand the reliability problem is now resolved. If customers experience further problems of this nature, software patches and modified CPU and memory boards are available from Sun. In this case, COD looks like an expensive solution to a vendor-specific problem and it doesn't prevent the problem from occurring.

Returning to the E10000's high-availability features, hot-swap system boards and dynamic partitions aside, the Sun server isn't that different from the HP and IBM midrange servers. Despite its high-end configuration, the E10000 has ECC main memory system buses, but only parity protection on internal and external caches. On this class of server, we'd have expected end-to-end ECC on all data paths. Failing memory is detected but not de-configured. The administrator must de-configure memory prior to a re-boot. Redundant I/O paths are supported either via alternate pathing or via the Veritas Volume Manager (VxVM). However, alternate pathing and the VxVM utilities cannot both be supported. Alternate pathing supports failover to standby controllers — additional high-availability software isn't required. It is automated for Small Computer Systems Interface (SCSI) controllers, but has to be managed manually by the administrator for network NICs. The system proactively monitors cooling, power, SCSI controllers, NICs, memory and CPU, although as already noted, when failures are detected recovery isn't always automated.

Overall, we've rated the E10000 application servers slightly ahead of HP and IBM, because dynamic reconfiguration and alternate pathing capabilities mean that spare capacity can more quickly be brought into play to replace capacity lost through component failures or failures within a partition. *Score: 6 out of 10.*

Online Upgradability: Despite the support for partitions and dynamic reconfiguration, the E10000 will not support processors of different clock speeds. If you want to upgrade an E10000 with faster processors, you have to replace all the processors at the same time, with the same speed of processor — a costly and time-consuming undertaking. Nevertheless, the ability to "hot add" system boards and COD processors to partitions means the E10000 again scores slightly ahead of the HP and IBM midrange servers. *Score: 6 out of 10.*

Operating-System Availability: The Solaris operating system supports "live" operating-system upgrades — i.e., the ability to build a new upgraded operating environment image, while the system is still running with a full application load. The new operating-system image is loaded to a separate disk partition. Re-booting then brings the system or partition back up with the new version. Again, this can be tested in spare or development partitions and then applied to production partitions on identical hardware.

Software patches can be installed or removed from a running Solaris system, and customers with a Sun Spectrum support contract can download the PatchDiag tool to determine and track the system's patch status. Patches that affect the kernel or device drivers will require a system re-boot in order to take effect.

Solaris 7 and 8 feature dynamically loadable kernel modules, so new device drivers can be added without a re-boot. The operating system supports the Veritas file system, which allows backup, defragmentation and other administrative tasks to take place without interrupting work in a partition or on the system as a whole.

The ability to test patches and maintenance level releases in a partition without affecting the rest of the system lifted Sun's scoring ahead of the competition on this criterion. *Score: 6 out of 10.*

5.3.3 Sun's Technology: Disk Storage Subsystem

Availability: For storage, Sun offered its StorEdge T3, which was introduced in June 2000. T3 is based on development by MaxStrat, a company acquired by Sun in 1999. This system is a midrange disk subsystem and is not a breakthrough in disk technology. The midrange, modular disk subsystems cannot compete against the scalability, performance, availability and functionality of the enterprise subsystems. Sun lags behind HP and IBM in providing robust, enterprise disk subsystems.

Features of the T3 include automatic failover and "mirrored" cache if two controller units are connected, "hot pluggable" components, N+1 power and integrated battery. The disk arrays use a RAID 5 technique with nine drives in a RAID group: eight data and one parity; or seven data, one "hot" spare and one parity. The relatively large RAID group will create heavy overhead in a data rebuild in the event of a disk drive failure. RAID 0 and RAID 1 are also supported. *Score: 5 out of 10.*

Performance: Sun says a single controller has a bandwidth of 93MB/sec and can support up to 4,400 I/O operations per second, or "IOPS" (100-percent read in 8KB blocks). In a production environment, actual throughput will be lower. For example, a RAID 5 array with 18.2GB disks can sustain 1,680 random-read and 620 random-write IOPS. As yet, Gartner has little performance data or client feedback, and enterprises should carefully check performance claims and request written guarantees. *Score: 4 out of 10.*

Functionality: The T3 has a 10Base-T Ethernet connection, which can be used by storage management software. It also supports Veritas Volume Manager. The subsystem supports segmented Fibre Channel,

which allows static switching only. Veritas Volume Manager provides dynamic multipathing to allow path redundancy. Remote mirroring is supported through StorEdge Network Data Replicator, while StorEdge Instant Image provides point-in-time copy. Both are software solutions, which cause host overhead and are generally less advanced in comparison with the enterprise subsystems' hardware functions. *Score: 2 out of 10.*

Scalability: The T3 is based on modular sets of desktop drawers (up to eight can be installed in a rack, leaving space for two switches) with 163GB or 329GB RAID 0 capacity per module. Each array group contains nine disk drives (18.2GB or 36.4GB at 10,000 RPM) connected to two FC-AL loops, and has a place for a controller card with a 256MB cache. An array with a controller unit can support one expansion drawer (to be increased to three). Two arrays with controllers can be connected in a redundant configuration. The T3 is unproven for handling very large data volumes compared to the competition. *Score: 2 out of 10*

5.3.4 Sun's Technology: System Manageability

Server Manageability: The key value-add in Sun's proposal is the fact that the complete SAP solution runs on a single server. The proposition is simple: One big server will be easier to manage than 18 smaller ones. But the big server still has to be partitioned into seven logical servers, and those partitions must be managed separately. Each partition has to have its own boot disk and network connection, and must be loaded with its own instance of the operating system. If there are any advantages, they come from greater flexibility — the ability to reconfigure partitions to suit changing business needs. With most SAP configurations, however, maintaining stability is likely to be a greater need. Furthermore, there's no evidence from E10000 users that consolidating applications on a single footprint actually saves them any management costs. Any savings that come from physically consolidating servers in a centralized data center could equally well be achieved using small rack-mounted servers, physically located together.

Gartner's Application Server Evaluation Model rates Sun lower than both IBM and HP on manageability. Solaris' management tools are immature and do not match the functionality of IBM AIX's and HP-UX's management tools. Sun's Solaris Resource Manager software, for example, which provides policy-based workload management on a single server or within a single partition, is a recent addition to Solaris and still in relatively limited use in the field. Its also an optional cost feature, unlike IBM's WLM software, which is included as part of AIX. The Resource Manager software is not integrated into the Solaris Management Center GUI.

However, Sun does have one manageability advantage over HP and IBM: Solaris is Unix SVR4-compliant. It provides a full set of Unix System V commands for use via its CLI, and this endears it to many Unix administrators — IBM's and HP's operating systems are not fully System V compliant.

Overall, though, Sun's system management tools lack the comprehensiveness of its competitors. Once again, however, it's the superior flexibility provided by its Dynamic System Domains that boosts Sun's manageability score. *Score: 5 out of 10.*

Operating-System Partitioning: Sun is the only vendor able to offer partitioning, and this has allowed Sun to present the E10000 as an alternative to a more traditional SAP configuration of multiple application servers. Sun's Dynamic System Domains are more advanced than any other Unix vendor's and also give it an edge with features such as hot-swap system boards. *Score: 7 out of 10.*

Cluster Management: Sun's cluster option is based on SunCluster 2.0. Gartner rates the SunCluster software as inferior to HP MC/ServiceGuard and IBM HACMP. It isn't as mature or as proven. Many Sun customers in fact opt for the Veritas clustering software instead. Sun announced an enhanced version of

its SunCluster software at the start of 2001. SunCluster 3.0 has a global file system for easier centralized management and in that respect is in advance of its competitors. It now needs the customers and references to prove itself against Veritas. Good clustering can be achieved on the Sun platform, but not yet with Sun. *Score: 3 out of 10.*

COD Functionality: Sun is the leader for COD functionality because the E10000 is an enterprise-class server and Sun took the lead in offering COD on an enterprise-class Unix server. COD is meant to provide additional capacity for applications, such as Web servers, where it is difficult to predict capacity requirements. It has minor benefits in terms of serviceability for applications, such as ERP, that have more predictable capacity requirements. The major benefit is that when additional capacity is finally needed, it can be added without having to bring the server down. It can also serve to bring spare processors online to replace failed CPUs or system boards — again, without the delay caused by physical replacement. *Score: 6 out of 10.*

Storage Management: Sun's storage management offering, like its overall storage offering, is considered inferior to those from the other vendors. *Score: 2 out of 10.*

5.3.5 Sun's Services: System Support

Response Time — Standard: Like HP and IBM, Sun offers a standard support service response time of between four and eight hours from receipt of a call. *Score: 5 out of 10.*

Response Time — Optional: Again like IBM and HP, Sun offers an optional enhanced response time, most often characterized by a 24x7x4 approach. *Score: 5 out of 10.*

Proactive Software Patch Management: Like all hardware vendors, Sun does not offer comprehensive software patch management as a standard support feature. It must be purchased as an additional item. *Score: 1 out of 10.*

Proactive Service: Sun provides a somewhat limited proactive-service offering. The Sun Remote Services (SRS) program is limited to monitoring Sun server and storage products. The current version of SRS, version 2.0, now stable after a number of bug fixes, is due for an update. It has been at its current revision level for over 18 months. Sun says the next revision, version 2.1, due out shortly, will include new features such as Web view reporting. *Score: 3 out of 10.*

Assigned Support Team: Sun provides dedicated support teams and, in some cases, on-site support engineers. However, the service is not as comprehensive as HP's. *Score: 3 out of 10.*

Accredited Escalation Process: Sun is ISO-9000-registered. *Score: 4 out of 10.*

Integration With SAP's Escalation Process: Sun is an SAP global services partner and hardware partner, and has a high degree of integration with SAP's escalation processes. *Score: 6 out of 10.*

Disaster Recovery Service Options: Although out of the scope of this detailed analysis, we ranked Sun slightly lower in the area of disaster recovery services because Sun does not have its own disaster recovery service business unit. However, it does offer business continuity and disaster recovery consulting services, and has alliances with Comdisco, IBM and SunGard for disaster recovery outsourcing requirements. *Score: 7 out of 10.*

Disaster Recovery Service Adoption, Europe: At the time of the survey, Sun could not say how many of their European customers use these disaster recovery services. *Score: 5 out of 10.*

5.3.6 Sun's Services: Consulting Services

Proven SAP Methodology: All three respondents scored equally well in this section, with no discernable differentiation in the methodology used to professionally design, build and commission the server infrastructure for an SAP project. All vendors will accept turnkey responsibility for this critical project activity. The organization that delivers these services is called Sun Professional Services. *Score: 7 out of 10.*

Pre-packaged SAP Services: All respondents had a wide range of pre-packaged SAP services, but Sun had a smaller range of standard, fixed-price, packaged SAP services described by data sheets. *Score: 5 out of 10*.

ITIL Problem and Change Management Services: Sun does not usually offer services based on ITIL methodology, the use of which is considered to be a best practice in the problem/change management, and in service management in general. *Score: 4 out of 10.*

5.3.7 Sun's Business Practices

Responsiveness: Sun proved to be very slow in responding to simple requests for this survey. *Score: 3 out of 10.*

Completeness of Solution Response: Sun's response was less comprehensive and required considerable follow-up effort. *Score: 5 out of 10*

System Reliability in the Field: Gartner has received extensive feedback from its clients around the world of poor E10000 system reliability due to the e-cache problems. *Score: 3 out of 10.*

Support Reliability in the Field: Based on client feedback to Gartner analysts, the Sun support organization does not enjoy a good reputation for providing high-quality support in comparison to more-established competitors like HP and IBM. *Score: 4 out of 10.*

6.0 Conclusions

At the time of the survey, all three vendors struggled to present a seamless solution for a fully integrated SAP HAS project. The main reason was that they are structured so that different internal organizations deliver computer systems, consulting services and support services. The vendors sometimes try to mask this by leading any proposal with a single sales representative. In reality, enterprises need to address many points of contact at each vendor to get detailed answers about various parts of the overall solution.

The complexities of designing, planning and "costing out" large-scale enterprise applications require specialized and dedicated resources to respond appropriately and in a timely manner. All three vendors seemed to lack the people, resources and processes necessary to deliver.

HP has always had a strong focus on selling and supporting mission-critical Unix systems. HP's "Five Nines — Five Minutes" marketing message exemplifies this approach. This message is more hype than reality, but it does communicate HP's direction. HP developed specific consulting and support services for SAP projects several years ago. These can be very proactive in nature, albeit with a high price tag for many users. HP's server and disk storage technologies are mature and have a good reputation in the field for reliability. Many of HP's earliest large-scale, mission-critical SAP projects ended up with EMC storage for reliability reasons (and HP sold EMC disk systems until 1999), but the XP range has already ended this trend. HP is more able than its competitors to discuss the level of overall SAP application availability that can be expected, or in some cases guaranteed, from a given SAP HAS configuration. Not

surprisingly, HP scored consistently higher than its competitors in this survey, with an overall weighted score of 409.

Like HP, IBM often claims superiority in the Unix world for its mission-critical systems. HACMP is very well-established. IBM has a very good reputation for reliability of its pSeries (RS/6000) server products in the field. IBM's disk storage strategy embraces the newer ESS product range, which is clearly less established in the market. EMC has been commonly installed in IBM RS/6000 SAP projects. Like HP, IBM also has very well-established consulting and support organizations. IBM scored well in this survey, coming second with an overall weighted score of 355.

Sun is the clear leader in general-purpose Unix computing and it offers the most sophisticated partitioning capability with its Solaris domain technology. From a clustering perspective, Sun is behind its competitors with the SunCluster product. The Veritas third-party cluster product is frequently installed in SAP projects. Unfortunately, Sun's E10000 server family, while a very successful product, has not enjoyed the same level of reliability as those from HP and IBM. Sun has ambitious plans in the disk storage market, but the StorEdge products do not currently compare with those of EMC, HP or IBM. EMC has been very widely installed in SAP projects using Sun platforms. Sun's consulting and support organizations are growing rapidly, but its enterprise application integration competency, particularly with regard to SAP services, is much less mature than those of HP and IBM. Sun's proposal for an SAP HAS was heavily predicated around the strengths of its technology architecture. Gartner believes high-availability solutions are best delivered through an integrated set of tightly bound technologies, services and processes — an approach that was absent from Sun's proposition. Sun did not score well in this survey, with an overall weighted score of 283.

All three vendors have introduced new systems during 2001, and Sun has adopted Hitachi as its main storage partner. All three would probably submit different configurations to an invitation to tender made at the end of 2001. The new systems are outlined in Appendix B.

7.0 Recommendations

From this SAP HAS research study, three main recommendations emerge for enterprises planning to build a large-scale, mission-critical, SAP infrastructure. They are:

- Address IT process issues first. This is not the main focus of this report.
- Put the right questions to hardware vendors.
- Identify a complete SAP HAS solution.

7.1 Addressing Business and Process Issues

The logical starting point in any mission-critical SAP project is with the business itself. It is imperative to know how much it will cost the business if the SAP system should go down. This cost of downtime should be used to justify investments to enhance the infrastructure to reduce both planned and unplanned downtime, so that optimum SLAs can be met.

Enterprises providing the SAP system infrastructure for themselves should invest in maturing their core IT management processes of:

- Configuration management
- Change management

- Problem management
- Performance management and capacity planning
- SLA reporting

If the SAP system infrastructure will be outsourced, the enterprise must check that its service provider is as efficient as it claims to be at these management processes. If it is not, any problems will impact the enterprise's business in just the same way as if the enterprise had provided its own system.

7.2 Asking the Right Questions

Enterprises should always ask hardware vendors the right questions but, surprisingly, this does not always happen. This is often the case regardless of whether the enterprise is selecting a new hardware vendor or utilizing an existing one. Some of the questions to ask are as follows:

7.2.1 Technology Questions

- Where are the single points of failure in the proposed system?
- What evidence do you have of actual failure rates for the proposed servers and storage in the field?
- Where are the reference sites for other customers with exactly this proposed configuration?

7.2.2 Consulting Services Questions

- How appropriate are the backgrounds, skills and experiences of all proposed vendor consultants?
- Are the proposed consultants SAP-certified for Unix and your database?
- Exactly what system documentation will be provided?
- How will skills transfer be enabled from vendor consultants to the enterprise's staff?
- What quality assurance mechanisms are built into the project deliverables?

7.2.3 Support Services Questions

- What is the promised interval from call time to respond time for a hardware or software problem at any time throughout the week?
- What are the time periods needed to fix problems?
- How proactive is the vendor in predicting and identifying problems before they occur?

7.2.4 General Questions

- What level of operating system, database, SAP Basis and SAP application availability will be achieved by the proposed solution?
- How cost-effective are the technology options offered by the hardware vendor, i.e., in terms of the claimed improvement in overall availability of the SAP application?
- How many customers are using this solution now, and is it really proven?
- If the system is absolutely mission-critical, is there a senior manager in the hardware vendor who is willing to take overall responsibility for the complete SAP HAS solution, bearing in mind that multiple vendor organizations will be involved?

7.3 Identifying Everything That a Complete System Requires

Enterprises should use the SAP HAS model defined in this *Strategic Analysis Report* to plan for all elements of hardware, software, consulting services and support services that will be needed for integration into a complete system for the required infrastructure. Consider all design, build, test and operational phases of deployment of that infrastructure. Utilize the SAP HAS vendor criteria to evaluate hardware vendors not profiled in this *Strategic Analysis Report.*

Appendix A: Related Research

Server Hardware

Enterprise Guide to the Application Server Evaluation Model (R-10-0234)

Selecting the Best R/3 Server (KA-03-2449)

Selecting the Best R/3 Server: Additional Factors (Ed Thompson, Nov 98)

64-bit SAP — Can R/3 Users Ignore It? (TU-09-0862)

64-bit SAP Hardware Platforms: Ready or Not? (DF-09-9208)

Hardware Reliability Problems With Sun UE Servers (SPA-09-5788)

Storage Hardware

A Storage System Evaluation Model (COM-11-9285)

System Availability

Making Smart Investments to Reduce Downtime (TG-07-4033)

HP's Somersault Aims to Increase R/3 Availability (E-09-5068)

Sun's Full Moon Finally Begins to Appear (E-12-6742)

Availability: How Do Your Application Services Stack Up? (SPA-12-8280)

Oracle Database Reorganization Concepts (TU-11-6934)

Sizing SAP Systems

Sizing SAP Systems: Art or Science? (TU-08-0030)

Sizing SAP Systems: Who Should Be Involved? (DF-08-6331)

Appendix B:

Survey Updates

Technology never stays still and each vendor has made some changes to its technology capabilities and offerings since responding to the survey. The main points are given here. But enterprises should always check the features and availability close to the time of evaluation.

In September and October 2001, all three system vendors (HP, IBM and Sun) announced new servers. HP announced a new midrange server, the rp8400, which supports up to 16 processors and two hard partitions. IBM announced a new high-end server, the pSeries 690, which features unusual dual-core Power 4 processors. Sun announced the follow-up to the E10000, the Sun Fire 15000.

HP's Superdome and rp8400

When responding to this survey, HP did not have a large partitionable server as an alternative to multiple application servers. Since then, however, HP has announced Superdome. Superdome supports up to 64 processors and up to 16 partitions — more than enough capacity to support the workload of the six application servers in the original proposal. So, in the case of large SAP R/3 implementations, we would expect to see Superdomes proposed for both application servers and database servers.

As yet, however, HP's hard partitions ("nPartitions," as HP calls them) cannot be dynamically reconfigured, so they are less flexible than Sun's Dynamic System Domains feature. Hardware cannot be added to or removed from a partition "on-the-fly." But HP is introducing a more dynamic form of partitioning into the HP-UX 11i operating system. vPars will provide software-based partitioning with the capability of dynamically moving resources between vPars either under operator control or under the control of scripts — in response to changing workload requirements, for example. vPars and nPars would not in themselves ensure higher availability compared with multiple midrange servers, however. But they would make for a more manageable system and that could lead to better availability.

The HP server rp8400 is a baby version of the Superdome, and its modular architecture, with cells (CPU/memory boards) and a crossbar backplane, is based on the Superdome design. The big difference between this and previous HP midrange Unix servers, such as the N-Class, is that the rp8400 supports up to 16 processors (compared with eight for the N-Class). In keeping with the trend of adding high-end features to midrange systems, the rp8400 also supports partitioning. The rp8400 supports two hard partitions and in future will support virtual partitions. Typically, the hard partitions could be used to create separate, isolated partitions for test and production. The software-based vPars could be used to create multiple virtual servers within a hard partition. In the SAP R/3 context, for example, a production partition in an rp8400 could be used to run multiple application servers.

The rp8400 is relatively compact: Two will fit in a single rack. Its performance is good (140,000 tpmC), but prices are more than double those of the N-Class. In the SAP R/3 application, the rp8400 would make for a more scalable database server, and the number of application servers could be halved.

IBM's eServer pSeries 690

The pSeries 690 (p690) represents a new generation of IBM Unix servers. It supports IBM's dual-core Power 4 processors. The dual-core processors consist of two microprocessors, shared Level 2 cache memory, and switching logic and interfaces on a single chip. These two-way SMP chips are in turn mounted on a specially devised multichip module (MCM) which supports up to four Power 4 chips. Up to four MCMs can then be supported on the p690's backplane to create a 32-way system. The very dense packaging of the two-way processors, MCMs and crossbar backplane results in short, high-speed

interconnects between the component modules. IBM claims very high performance for the p690 — approximately 1.8 times that of its predecessor.

The p690's other key innovation is the inclusion of logical partitions (LPARs) for the first time on an IBM Unix server. Originally developed for IBM mainframes, LPAR technology promises greater flexibility and more granularity (i.e., smaller partitions) than Sun's Dynamic System Domains. The main benefit of partitions or domains is the ability to consolidate workloads on a single platform — Sun has used this feature extensively, as in the proposal considered in this study to host a complete SAP R/3 application on a single, partitioned server. IBM's introduction of LPARs on the p690 is initially limited, however: they can't be used dynamically with this first release — in other words, the server has to be brought down to change partition configurations. Consequently, we would advise enterprises to limit use of the p690 to back-end database serving for high-availability SAP R/3 applications. IBM has provided the plumbing for LPARs, but wait for a future release for it to become really useful.

IBM has also enhanced the midrange servers proposed as application servers in our study. The RS/6000 H70 and H80 have been followed by the eServer pSeries 660 (p660) Model 6H1. This has faster PowerPC RS64-IV processors, though it still supports up to six processors in the same rack-mount chassis as the RS/6000 H70 and H80. The Model 6H1 offers good performance (over 57,000 tpmC) compared with other low-end/midrange servers, but looks "long in the tooth" compared with more recent designs.

Sun's Sun Fire 15000

The Sun Fire 15000 is Sun's follow-up to the E10000. Also known as StarCat, it is more scalable than the E10000 (up to 76 CPUs and 576GB of memory, compared with 64 processors and 64GB of memory for the E10000), has a new interconnect and supports a new generation of Scalable Processor Architecture (SPARC) processors — UltraSPARC III. The E10000's key features, Dynamic System Domains and Dynamic Reconfiguration, are of course supported and, like the E10000, the Sun Fire 15000 is intended for use as a large consolidation server. Just as Sun proposed the E10000 as a single platform on which to consolidate a complete SAP R/3 workload, using its domain capability to host the individual applications and database servers, so the Sun Fire 15000 is likely to be used in a similar way.

The Sun Fire 15000 is due to be generally available in December 2001, although some features such as Dynamic Reconfiguration will be delayed (until mid-2002), while Sun completes testing with the Solaris operating system. StarCat is only supported in Solaris 8 and above, so customers on an earlier operating-system release will have to plan migrations to the newer release. Sun will continue to sell the E10000 for customers not yet ready to move and while StarCat completes its testing and reaches maturity.

Sun calls StarCat "the Unix mainframe" as a reference to its mainframe-like availability, manageability and scalability. On paper, at least, it appears to outscale IBM's z900 mainframes. In the real world, however, proof of StarCat's scalability and availability is inconclusive: Sun has yet to provide any benchmarks that would enable comparisons with other high-end Unix servers or, indeed, IBM mainframes. There is no doubt, however, that StarCat's predecessor, the E10000, has been extremely successful in the market (with over 5,000 sold) and has offered customers a number of mainframe-like features that were previously unavailable on Unix platforms. Those features (e.g., partitioning, the ability to do concurrent repairs or upgrades, and support for huge amounts of memory and I/O) — together with support from leading independent software vendors — provide a firm base on which to build. The one doubt lingering in many clients' minds, however, is the loss of credibility that Sun suffered over the reliability of its servers after the problems with processor failures on the E10000. The onus is now on Sun to prove that it can build and support systems with mainframe-like reliability and availability.

Appendix C:

Acronym Key

24x7	24 hours a day, seven days a week
24x7x4	24 hours a day, seven days a week, with four-hour response time
AIX	Advanced Interactive Executive
CLI	Command line interface
COD	Capacity on demand
CPU	Central processing unit
CSS	Critical Systems Support
DLT	Digital linear tape
DMR	Dynamic Memory Resilience
DPR	Dynamic Processor Resilience
ECC	Error-correcting code
EMS	Event Monitoring Service
ERP	Enterprise resource planning
ESS	Enterprise Storage Server
GB	Gigabyte
GUI	Graphical user interface
HACMP	High Availability Cluster Multiprocessing
HAS	High-Availability System
HDS	Hitachi Data Systems
HP	Hewlett-Packard
I/O	Input/output
IOPS	I/O operations per second
IT	Information technology
ITIL	Information Technology Infrastructure Library
JFS	Journaled File System
LPAR	Logical partition
LUN	Logical unit number
MCM	Multichip module
NIC	Network interface controller

PCI	Peripheral Component Interconnect
PRM	Process Resource Manager
RAID	Redundant array of independent disks
RAM	Random-access memory
SCSI	Small Computer Systems Interface
SLA	Service-level agreement
SMP	Symmetric multiprocessing
SNMP	Simple Network Management Protocol
SPARC	Scalable Processor Architecture
SRS	Sun Remote Services
TB	Terabyte
VxVM	Veritas Volume Manager
WLM	Workload Manager