Managing Application Services: A Progressive Approach

Manage the Service, Not the Server

Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303
1 (800) 786.7638
1.512.434.1511

Please
Recycle

Contents

CHAPTER 1

Introduction

Recent years have seen a boom in the scope and breadth of IT infrastructures. The exploding demands of Internet commerce and other network-based business activities are driving requirements for IT infrastructures that are more accessible, more dependable, and more efficient than ever before. At the same time, the competitive nature of the market puts an even greater strain on limited IT budgets. Every technology product, personnel resource, or operational process introduced into the IT infrastructure must deliver maximum value.

As a result of these converging pressures, IT executives are increasingly burdened with finding new approaches to maintaining high levels of service. At the same time, they must find ways to decrease the costs and risks associated with managing large infrastructures.

The IT architecture must support both unpredictable demand as well as rapid rates of change. It must optimize the use of existing resources without requiring redundant investments in specialized equipment, software, or personnel. Above all, it must deliver a very high quality of experience to the users the infrastructure was designed to serve. For users, quality is defined simply in terms of predictable, consistent, application service level availability and performance.

Sun Microsystems, a leading provider of technologies that power the Internet, takes a new approach to managing application service levels throughout the IT infrastructure. Using Sun's advanced solutions for clustering and centralized system management, IT organizations can abstract individual systems into tightly coupled groups of pooled resources, which can then be managed more efficiently through simplified, automated Service Management techniques.

By managing application services, rather than individual servers, IT organizations improve the scalability, availability, and manageability of the entire IT infrastructure. Sun's Service Management solutions help IT professionals create a more fluid, agile infrastructure for delivering highly available, predictable, and sustainable services.

CHAPTER 2

The Net Effect

From wireless to voice-over-IP, streaming video to broadband, the availability of new technologies is spurring demand for increased application availability. More users than ever before are connected to the Net, with more different types of devices. This in turn creates demand for additional applications, and easier, more efficient ways to deliver those applications to a growing number of users. The multiplication of users, services, data, and devices, known as the Net Effect, results in even higher requirements for availability and scalability across the IT infrastructure.

The Net Effect means IT organizations must rethink their approach to designing, deploying, and managing applications as well as the underlying infrastructure. Data centers may now support hundreds of thousands, even millions of users in a single day. Applications are no longer individual pieces of software installed on individual machines, but rather are services accessed as needed by users over a continuously available network.

Managing the entire infrastructure to deliver predictable, scalable, and highly available application services is the next battleground for IT organizations. Sun's approach to Service Management helps abstract the complexities of multitiered server architectures and makes it possible to increase service levels while decreasing costs.

CHAPTER 3

Management Challenges: Scalability, Availability, Complexity

Whether the IT infrastructure supports a global manufacturing enterprise or an information-driven e-retailer, the issues of scalability, availability, and manageability still apply. Efficient management of the proliferation of applications, servers, and storage required for such an infrastructure is a common challenge for all IT professionals.

Achieving Scalability

The challenge of scalability is twofold. As the value of the data center grows, more and more users demand access to its services. Instead of supporting thousands of users, IT managers must now plan to support hundreds of thousands, even millions, when the data center is publicly accessible via the Web. As demand scales, it becomes more difficult to effectively predict the peaks and valleys of usage. Nearly every enterprise with a Web-based presence experiences unforeseen problems when a spike in demand literally brings its service delivery infrastructure to a crashing halt.

To counter the challenge of scalability, many IT departments elect to *overprovision*, that is, to bring more servers online than are truly needed for average loads. Not only is this solution expensive, but it greatly adds to the complexity of the infrastructure and increases the management burden. Therefore, while overprovisioning may result in increased service levels, it also increases costs.

Providing Availability

Just a few short years ago, business ran eight or twelve-hour days, five days a week. IT managers could perform maintenance, test new environments, and change system configurations during the off hours. But the new Net-driven economy has virtually eliminated this. Application services, particularly those accessible via the Web, are used continuously across time zones and geographies. Because these application services are business critical, keeping them continuously available takes precedence over any other IT operation, which makes it difficult to take any system offline for even the smallest amount of time.

To assure near 24x7 uptime, many enterprises moved to a multitiered architecture with redundant or replicated systems at every layer. This structure enables some systems to be offline while maintaining a moderate level of service delivery. Again, the result is a trade-off between maintaining service levels and increasing the complexity and costs of managing the infrastructure.

Managing Complexity

The third — and most difficult — challenge arising from the Net Effect is the resulting complexity of the massively scaled, multitiered infrastructure. Environments can easily grow to hundreds of individual servers, all requiring monitoring, tracking, maintenance, and upgrading for each instance of the operating environment and every application service. While an individual server and application service is simple to manage, a thousand systems are not, even if they are configured identically. The simplest tasks, such as applying an operating system patch or modifying a database, become management nightmares.

One leading brokerage firm recently reported making nearly 3000 changes to their Web infrastructure every month. At these levels, even the best practices and disciplined techniques of the traditional data center fall short. Numerous changes strain the staff’s ability to complete accurate logging, testing, and tracking procedures, and add incalculable risks to infrastructure operations.

Decreasing risk depends on reducing the complexity of the infrastructure. This is where Sun's approach to Service Management shines. Sun's solutions enable IT organizations to manage resources more efficiently and provide high levels of application services.

CHAPTER 4

Service Management at Work

Sun's Service Management approach works because it applies one fundamental concept to the challenges of achieving scalability, providing availability, and managing complexity: *manage the service, not the server.*

To do this, IT professionals need to apply tools and techniques to help them abstract the individual components of the infrastructure into a dynamic yet simplified whole. With Sun's approach to Service Management, this involves four basic steps:

- Create pools of resources to enable scalable management and services.
- Apply provisioning and change management to enable rapid, flexible deployments.
- Establish procedures and utilize tools that help ensure delivery of predictable, end-to-end service levels.
- Centralize and automate event management to help track and manage resources more effectively.

Creating Pools of Resources

Grouping resources enables systems to be managed as one coherent image. Many IT departments already apply this technique by grouping similar servers into *loosely-coupled* arrays. In this arrangement, each server is physically separate and may even be geographically distant. Application service availability is managed by balancing loads across the servers. This is most often seen at the front-end of a Web infrastructure.

Another alternative is to group servers into *tightly-coupled* clusters, in which there are both physical (hardware) and logical (software) interconnects between each system. The tightly-coupled environment is well-suited for the middle and back-end tiers of the infrastructure because it enables resources to be shared and scaled dynamically, while preserving very high levels of availability.

Creating shared resources abstracts one layer of the complex infrastructure, enabling IT managers to scale the management along with scaling services. Built around Sun™ Cluster 3.0, as well as the Solaris™ Operating Environment and Sun™ server, storage, and network connectivity products, and services, the SunPlex™ system is designed to manage application services for tightly coupled environments, optimizing the availability and scalability of these services. In the near future, Sun will deliver technologies to assist with managing loosely-coupled infrastructures as well.

Apply Provisioning and Change Management

With the rapid pace of Net-driven business, data center infrastructures cannot remain static. Both provisioning and change management are essential for simplifying the environment and enabling rapid deployment and modification. Provisioning allows IT managers to flexibly control how application services are distributed over the available systems. Sun's approach to Service Management tightly integrates provisioning and change management to enable dynamic starting, stopping, and restarting of application services.

Deliver Predictable, End-to-End Service Levels

The effectiveness of Service Management can be measured quantitatively (in terms of system uptime) and qualitatively (in terms of the user's experience.) Availability and predictability directly affect the quality of the user's experience. For example, if an application service such as a Web-based news service is not available, user satisfaction suffers. A major component of Service Management is ensuring that the user has the same — or better — quality of experience all the time, regardless of the time of day.

To achieve this, IT managers must be free to focus on delivering application services rather than devoting their limited resources to the burdens of system administration. Solutions that simplify and automate system and software management tasks enable IT organizations to apply their resources more effectively. In essence, they can optimize the environment to increase service levels and decrease management costs.

Centralize Resource and Event Management

The complexity of today's computing environments requires even more robust methods for tracking, monitoring, and managing all the components of the infrastructure. Service Management takes into account that many factors affect system management and uptime, including the reliability of the hardware and software as well as the resources and procedures used to manage the systems.

Automating and centralizing as many system management tasks as possible removes much of the burden from IT organizations. With technology such as remote monitoring, predictive failure analysis, and automatic recovery, it is possible for systems to intelligently manage themselves, and even alert IT staff to potential problems before service levels are compromised.

Sun's Service Management solutions combine all these elements — shared resources, provisioning and change management, service level management, and centralized event management — to solve the problems of scalability, availability, and complexity facing today's IT customers.

CHAPTER 5

Sun Solutions for Service Management

Sun's new generation of solutions for application service level management include:

- Sun Cluster 3.0 software
- Sun™ Management Center 3.0 software
- Integrated consulting, training, and support services

Table 1 illustrates how each of these solutions applies to the challenges identified by the Service Management vision. Sun is also developing additional Service Management solutions for release in 2001.

TABLE 1 Sun Service Management solutions enable IT customers to *manage the service, not the server.*

Requirement	Service Management Steps	Sun Solution	Key Feature
Scale the management	Create shared resources	Sun Cluster 3.0	Global Devices, Global File Service, Global Network Service, and Scalable Services
Deploy rapidly, flexibly	Provisioning and change management	Sun Cluster 3.0	Application Services: define, provision, change
Deliver predictable, end-to-end service levels	Service level management	Sun Cluster 3.0	Failover Service
Track and manage resources and events	Centralized event management	Sun Management Center 3.0	Tracking, monitoring, fault and event management
		Sun™ Remote Services	Service notification

Sun Cluster 3.0 Software

Sun Cluster 3.0 is Sun's next generation clustering product, delivering an easy-to-use, continuously available, scalable clustering solution. Tightly integrated with the Solaris 8 Operating Environment, the Sun Cluster 3.0 solution offers a number of industry-leading features:

- Global Devices, Global File Service, Global Network Service, and Scalable Services enable continuous file and network availability.
- Sun Cluster 3.0 is the first solution to offer a Global File Service for standard UNIX® file systems.
- Sun Cluster 3.0 is the first and only cluster solution that offers scalability — within a system as well as across multiple systems — for applications, making it easier for IT managers to deliver consistent levels of service during periods of peak demand.
- Failover Services increase the flexibility of the infrastructure and significantly increasing application service availability.
- The Global Network Service with built-in load balancing, along with Scalable Services, enable IT managers to truly *manage the service, not the server.*
- The Sun Cluster 3.0 API, software developers kit (SDK), and agent builder are available to assist developers in creating highly available and scalable application services.

Sun Management Center 3.0 Software

Sun Management Center 3.0 software, an enterprise-ready systems management solution, is a powerful tool for managing up to thousands of Sun systems, the Solaris Operating Environment, and applications running on the Solaris platform. The new features of Sun Management Center 3.0 focus on helping IT professionals improve service levels and decrease costs by using scalable, adaptable, tightly integrated management tools:

- Web-based interface provides access to management information using secure SSL communication.
- Enhanced, proactive event and alarm management, and tight integration with Sun Remote Services, provide early notification of potential service problems.
- Enterprise-ready functionality, such as grouping and group operations, command-line interface utilities, and module configuration propagation enable IT staffs to manage above the level of the individual systems.
- Integrated with leading third-party tools — such as Computing Associates Unicenter TNG, HP OpenView VantagePoint Operations, and Tivoli TES — to address enterprise-wide, heterogeneous environments.

- Comprehensive online diagnostics and predictive failure analysis identifies hardware faults and warns of problems before they occur.
- A new licensing model provides the Sun Management Center 3.0 basic package free to all customers for an unlimited number of nodes. Advanced System Monitoring and Premier Management Applications packages are available with a licensing fee.

Consulting, Training, and Support Solutions

Most unplanned outages are due to errors caused by people and process breakdowns rather than failed equipment. So focusing simply on hardware uptime does not ensure that the entire infrastructure can meet required levels of service availability. In addition to proven technologies for Service Management, Sun also helps customers meet their service level goals through its world-class services organization.

Sun provides a full portfolio of integrated consulting, training, and support solutions to assist customers with applying the Service Management approach to their existing or planned network infrastructures:

- SunReady™ Availability Assessment for Sun Cluster is a procedural review designed to assess and rate a customer's data center capabilities in relation to their desired service levels.
- Sun Cluster Application Readiness Service offers the essential implementation and project management services required to configure new installations of Sun Cluster software as well as establish and test basic operations.
- Sun Cluster Data Center Readiness Service provides additional testing, procedural development, and documentation services to further support cluster operations within the data center.
- Comprehensive training courses ensure optimized implementation of Sun Cluster 3.0 and Sun Management Center 3.0 implementation.
- Sun Remote Services 2.1, the latest addition to Sun's remote monitoring solution set, includes real-time reporting features and Web view capabilities, and is an integral part of both the SunSpectrum Platinum™ and SunSpectrum Gold™ support packages.

CHAPTER 6

Conclusion

The explosive demand for application services makes it imperative that IT organizations develop robust, flexible infrastructures that can easily absorb the impact of continual change. The key to designing, deploying, managing, and sustaining such high-value environments lies in a comprehensive approach to application service level management.

Founded on the concept The Network is the Computer™, Sun continues to deliver the technologies and services necessary to fully harness the power of the Internet. With its new Service Management solutions, Sun helps IT organizations move from managing individual environments to managing the services they need to deliver.

CHAPTER 7

Resources for Further Information

To learn more about Sun's new approach to Service Management, please visit the following Web-based resources:

- Sun Cluster 3.0: www.sun.com/clusters
- Sun Management Center 3.0: www.sun.com/sunmanagementcenter
- Consulting, Training, Support: www.sun.com/service/solutions/sunclusters

Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303

1 (800) 786.7638
1.512.434.1511

http://www.sun.com/solaris

December 2000