An Oracle White Paper
March 2010

Managing Service Levels in Distributed Applications

ORACLE

Introduction

Composite applications—such as service-oriented architecture (SOA), cloud, Web 2.0, business process management, and Web services systems—bring a compelling set of opportunities to enterprise IT. Distributed, federated, loosely coupled systems enable efficient business interactions with customers and partners by adjusting easily to rapidly changing business conditions. However, organizations will not be successful in deploying core business processes on these complex systems without the level of control and predictability provided by service-level management (SLM).

Consider, for example, a manufacturer that automatically replenishes its inventory by integrating directly with its suppliers via Web services. Before it can make near-real-time decisions about order fulfillment, the manufacturer requires timely answers from other support services. This raises important quality of service (QoS) issues, such as the following:

- Is the warehouse management service always running and able to provide current inventory levels?
- How long does it take each supplier's system to respond with a promise date?
- Is it possible to minimize business costs by addressing service problems?

To ensure the effective management of high-value business processes in distributed applications, companies need to experiment with SLM early in the SOA adoption lifecycle. They must start with pilots and proofs of concept and, with each success, assemble increasingly complex, business-critical deployments. With composite applications, businesses need to measure service levels for processes that are distributed and federated. If there are issues, IT needs to quickly address them and determine the business impact. Better yet, IT should address service problems before failures occur.

Distributed applications also offer the opportunity to combine system- and application-level information to manage QoS effectively and in a manner that addresses business objectives. Organizations can answer such questions as

- What is the average response time for processing orders from Platinum customers?
- If there are service problems, can the priority of requests be raised for Platinum customers?

Most existing SLM solutions focus on managing devices such as network routers and server equipment. Such solutions are based on binary data; therefore, they are not fully aware of application components such as Web services, let alone the message- or application-level behavior. Furthermore, without understanding the content and context of messages, these management products cannot provide the benefits that come from having a business perspective. As a result, these solutions do not allow organizations to make real-time decisions or proactively alleviate QoS problems based on business objectives.

SLM of composite applications helps enterprises to

- Enforce end-to-end quality objectives that align with business goals
- Effectively plan for growth and reduce business risk
- Set realistic expectations with all stakeholders inside and outside the enterprise
- Justify additional investments

Start Early: Planning and Managing Pilots and Proofs of Concept

Service-level management (SLM) is the continuous and closed-loop process of measuring, reporting, and improving the QoS of systems and applications delivered by the IT or operations group. It includes a proactive, disciplined methodology for establishing acceptable levels of service that address business objectives, processes, and costs.

Pilots and proofs of concept will help form an organization's overall strategy for distributed applications and will highlight the benefits of this emerging technology. Consequently, companies need to evaluate these experiences carefully and, with each success, assemble increasingly complex composite applications. Furthermore, IT needs to quantify and measure success at each turn to justify additional investments in systems such as SOA and cloud.

To this end, SLM can help organizations clearly understand their use of their pilots and can provide visibility into specific performance and availability issues within their systems. Starting early with SLM gives enterprises the requisite experience to get their distributed applications ready for "prime time."

By recording and analyzing service levels over a period of time, the enterprise will have the historical data required to justify deploying additional and more-complex composite applications.

Consider an example. Before a financial services company can expose a claims processing application via service interfaces, it will need to understand the impact the new services will have on its internal systems and the QoS it can expect to provide brokers. Inadequate service levels would directly affect the number of claims the company could process in any given period.

Some of the indicators of the success or failure of these early projects will be performance (response time, throughput) and availability. For example, the company might want to know the following:

- What happens to the response time when a service is available to a small set of trusted partners?
- During business hours, can five 9s availability be ensured while verifying insurance information?
- Which audiences are using the system?

When planning and implementing SLM for distributed applications, there are a number of issues to consider.

Be a Discerning Consumer of Services

Standards and technologies such as XML and Simple Object Access Protocol facilitate the adoption and deployment of services-based systems. Additionally, the Web Services Description Language and Universal Description, Discovery, and Integration standards simplify the definition, publishing, and discovery of services. These standards can speed the use of third-party services, but challenges remain.

- After discovering multiple services that provide similar business functions, an organization will want to select the one that best meets their QoS standards.
- Some enterprises will not risk consuming an outsourced service, even from a trusted partner, until it has clearly demonstrated acceptable service levels over a designated period.

SLM can help separate the proverbial wheat from the chaff among various providers by showing QoS attributes, such as response time, throughput, or availability, which can be the difference between success and failure.

Starting early with SLM will help differentiate services and facilitate reuse. It will help establish these services and processes and, over time, provide demonstrable QoS. As this happens, organizations will reuse these services for other inter- and intraenterprise business processes and will be better prepared to set realistic expectations between providers and consumers.

- A new division might refuse to switch over to a new customer relationship management service until IT can prove that it has provided acceptable service levels for the past nine months.
- An organization might decide to change its payroll service due to degrading service levels for the accounting department.

Be a Differentiated Provider of Services

Similarly, from the perspective of a service provider, enterprises must deliver QoS metrics and premium services for discrete user populations (for example, the accounting department) by actively managing service levels. This will be a substantial differentiator in competitive business environments and will provide the opportunity for premium pricing.

For example, when the financial services company plans to sell its online payments processing services to customers and potential resellers, it can

- Differentiate its services by providing credible QoS data for its production environments
- Increase revenue by charging a premium for 99.9 percent (as opposed to 99 percent) availability
- Allocate resources effectively by focusing on service use for high-value customers

Develop Awareness of Service Quality

When starting out with pilots and proofs of concept, companies use SLM to monitor the performance and availability of their distributed applications and the constituent components. Pilot projects offer the opportunity for IT to set informal QoS goals for their applications and track against those goals. Here are two examples:

- An organization can determine whether it can maintain an average daily response time of less than eight seconds for all lookups to the user directory service.
- Similarly, it can see if it can provide 99 percent availability from Monday to Friday, 8 a.m. to 5 p.m., for an invoice printing service.

The financial services company from the previous example, in the last subsection, can evaluate data from its pilot project. If it is satisfied with the QoS metrics, it will roll out the claims processing service to its brokers in multiple phases, with each phase including a larger number of external users. With this approach, the insurance company will gradually fine-tune the system, set realistic expectations with its community of brokers, and reduce any potential business risk.

As businesses compile historic service-level data from their early experiences, SLM allows them to achieve a deeper understanding of their environments, analyze system behavior over time, and plan more-effectively for growth. Again, SLM can be the difference between success and failure.

Create Formal and Informal Service-Level Agreements

Service-level agreements (SLAs) are the instruments for defining and enforcing SLM. Businesses use formal and informal SLAs to manage risks effectively.

Informal SLAs are a collection of internal service-level objectives (SLOs) that are valuable for managing QoS, but are not legally binding for the organization. *Formal SLAs* are contractual commitments that

frequently include noncompliance clauses that result in penalties or the loss of accounts. The definition of formal objectives is common for pilot projects. Objectives for a pilot project could be

- Improve response times by 25 percent
- Improve availability from 97 to 99 percent

Therefore, an effective SLM solution will help IT

- Make sure expectations are based on achievable goals
- Prevent service levels from becoming noncompliant with defined objectives
- Manage compliance failures according to the magnitude of the penalties
- Minimize the cost of any noncompliance that occurs

Evaluate Service-Level Objectives and Performance

For each SLA, users create a set of SLOs. With the help of predefined or customizable instrumentation, the management solution accumulates all performance and availability data within the distributed environment. To determine if an objective has been met (for example, an average response time of less than eight seconds) the SLM product evaluates the objective based on

- Measurable data about a service (for example, response time, throughput, availability)
- Performance during specified times (for example, Monday through Friday from 8 a.m. to 5 p.m.)
- Periodic evaluations (for example, every 10 seconds or every eight hours)
- Consumers of the service (for example, Platinum versus Bronze customers)

A key aspect of defining objectives is to set warning thresholds and alerts for compliance failures. In addition, IT should be able to create multiple objectives (for example, for different user roles) for a single application component.

Provide Managers with Reports to Visualize Quality of Service Indicators

Business managers are especially concerned with obtaining periodic reports that give them easy-to-understand views of the service they receive, overall usage, and the business impact of service issues. The management solution should provide easily digestible information to help review the key success indicators of the early projects, evangelize further investments, and set realistic expectations with other stakeholders within the company. See Figure 1.

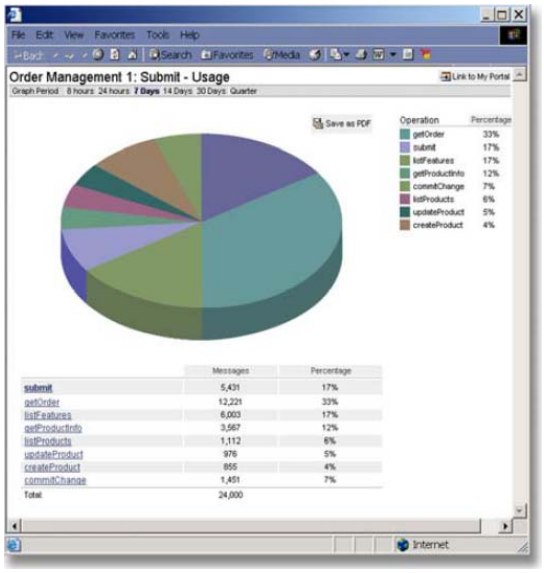

Figure 1. Graphic reports help business managers grasp key QoS indicators.

Business Insight into Distributed Applications

Done the right way, composite application management can use message content to improve the system's QoS. Organizations can also leverage tags, attributes, and element structure, as well as data from external sources such as Lightweight Directory Access Protocol to provide valuable contextual information.

Leverage Both Content and Context

By leveraging the content and context of the messages exchanged between application components, a management solution can derive rich business insight from distributed processes. When this business knowledge combines with system-level information, it opens up highly efficient and intelligent possibilities for managing service levels. Enterprises can manage services in ways never possible before. For example, the content and context of the messages can provide detailed information about the

- Customer that initiated a service request
- Total dollar value of an order that is being processed
- Appropriate support person for a specific order

Leveraging this content and context has very fundamental and compelling implications for SLM. Enterprises can define SLOs and thresholds for any relevant set of business criteria. For example, they can

- Monitor the performance of the order fulfillment process for all Platinum-level distributors
- Notify the account managers of all customers that are affected by a service failure

- Track the quantity of parts ordered per hour
- Monitor the response times by region
- Define the SLOs by user roles (for example, executives)

With the added insight, the manufacturing company can automatically align service levels with the business objectives.

Enforce Quality of Service According to Business Goals

To intelligently optimize composite applications and reduce risk, organizations must continuously and proactively prioritize use according to business goals. An effective management solution not only measures but also actively enforces end-to-end quality objectives that align with the organization's goals (see Figure 2). To that end, enterprises can use the rich insight provided by services-based systems to optimize performance and availability in a manner that delivers the greatest business value. This is even-more powerful if IT and operations can prevent or mitigate problems before they happen—especially if these decisions reflect business priorities. For example:

- If the performance of the order fulfillment service for the manufacturing company starts degrading, IT should automatically prioritize higher-value orders over lower-value orders
- When a service has been restored, first process orders from the most strategic customers and partners

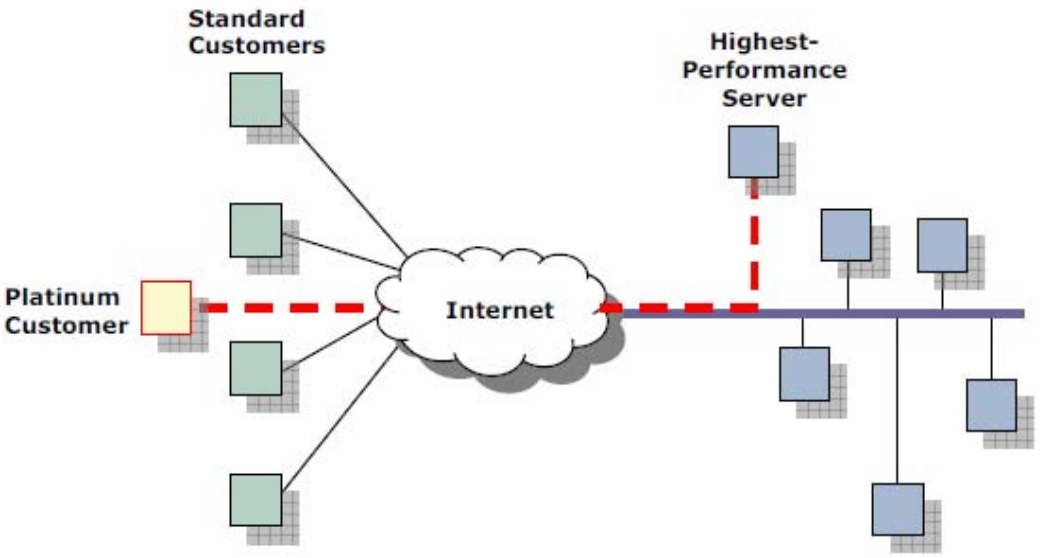

Figure 2. Routing orders from Platinum customers to the highest-performance servers enforces end-to-end quality objectives that align with the organization's goals.

Managing Service Levels for Distributed and Federated Processes

As enterprises gain experience, they migrate to more-mature environments, eventually integrating application components that are distributed and federated across heterogeneous platforms. This facilitates business processes that span multiple divisions within the enterprise or across corporate boundaries to include partners and customers. From end to end, the processes will involve multiple, independently controlled (that is, federated) application components that might rely on other components and processes to complete a task. Typically, each service will interact with numerous others, playing a provider role in some interactions and a consumer role in others. See Figure 3.

Such architectures will also facilitate loosely coupled exchanges between multiple application components. This will allow the enterprise to implement processes that can adjust to business conditions in real time by dynamically selecting the appropriate services. Although this flexibility increases the effectiveness of the process, it also introduces a level of unpredictability into the system.

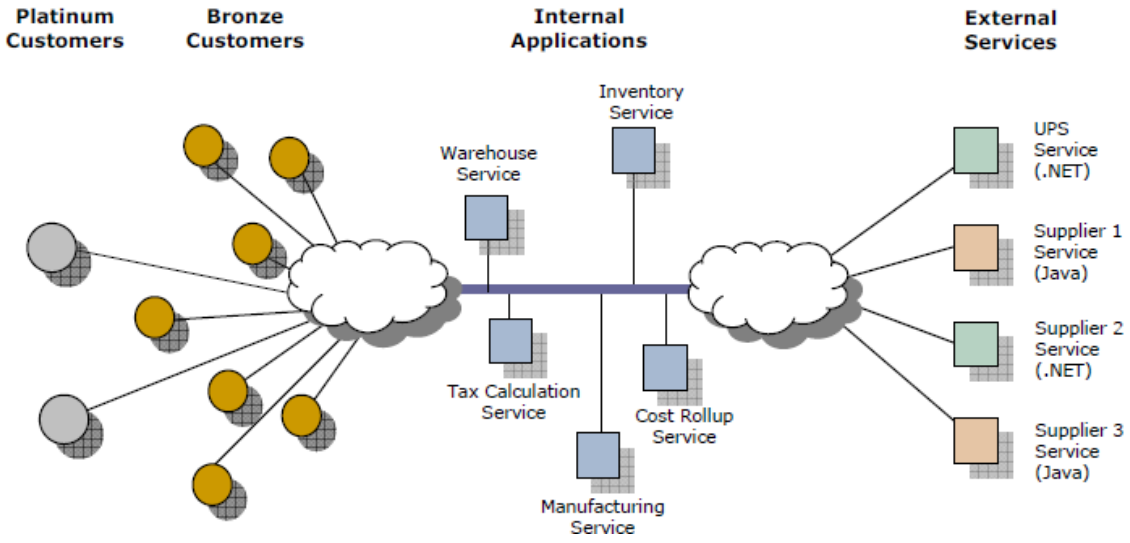

Figure 3. A distributed, federated application for manufacturing and fulfillment facilitates business processes that span multiple divisions within the enterprise or across corporate boundaries.

Consider the example of a company that would like to streamline its manufacturing and fulfillment processes across multiple factories and warehouses, as well as partners' facilities. The company has realized that it is more profitable to outsource the manufacturing of low-margin parts that are not in inventory at the time of order placement. Depending on the available-to-promise inventory levels, the company will either fulfill the order from one of its warehouses or place the order directly with one of its established partners for outsourced manufacturing.

This end-to-end order fulfillment process covers multiple inter- and intracompany services and is relatively flexible. To manage such a process effectively, the enterprise will need to

- Gather service-level metrics from distributed services (for example, inventory, manufacturing, warehouse)
- Monitor services it does not control (for example, response times for a UPS shipping service)
- Make real-time QoS decisions by monitoring each instance of the process (for example, immediately notify IT if the step that explodes the bill of materials goes down)
- Make real-time decisions based on historical data, defined goals, and business objectives (use historical performance numbers to select from a list of tax calculation services)
- Rapidly locate problems within a process and determine their impact
- Pre-emptively address issues before there are compliance failures (for example, if the response time of the cost rollup service is degrading, automatically route the step to a backup service)

As these processes are applied to core business functions within the enterprise, they will include an increasingly diverse set of service providers and consumers. As a result, inter- and intracompany use and associated business risks will rapidly increase. Firms that are implementing distributed and federated processes should take note of the following issues around an SLM solution.

Apply Service-Level Management to Business Processes

Long before a company starts running business-critical distributed applications, it will be required to manage these complex environments effectively by implementing formal or informal SLM to ensure QoS.

Using an SLM solution, the manufacturing company in Figure 3 sets the performance objectives for the entire multistep order fulfillment process. This includes obtaining and aggregating service-level metrics from the order processing services, the inventory services, and the internal or external manufacturing services. The company sets objectives for all critical manufacturing services to be 99.9 percent available during factory hours.

Additionally, the basis of the target objectives could be the results of monitoring the production environment over time; for example, the company might realize that it takes an average of 72 hours to completely fulfill the order for a specific inventory line and set that value as the objective.

The manufacturer can monitor the response times of all its different suppliers and for each part select, in real time, the supplier that provides the best performance.

Ideally, the SLM solutions for these dynamic environments are flexible enough to observe these processes without having to control them. This becomes especially critical if an enterprise has multiple processes it would like to track or numerous process engines within the organization.

One way to do this effectively is to identify the operations and correlate all messages that are part of the process. In this manner, the SLM solution tracks the end-to-end service levels.

Identify Service Problems Early and Determine the Impact

IT and operations personnel spend an inordinate amount of time trying to pinpoint bottlenecks in distributed multistep processes. They need an IT solution that enables easy identification of which operations or services are problematic. Furthermore, they need to determine the impact of services that fail by understanding the business processes affected by a service going down or rejected customer requests.

Additionally, by bridging system- and application-level information (as discussed in the previous subsection), companies can identify services that cause processes to perform poorly and can more-clearly understand the business impact service levels have on revenue and productivity. Once the management system understands the operations and messages for that part of the process, it can use this information to create effective diagnostics and impact analysis.

For example, if a sales tax calculation service fails, the IT group would want to immediately know the answers to the following:

- Which business processes are affected?
- Which customers are trying to access the service?
- Are there any potential penalties for SLA noncompliance?

Span Multiple Technical Architectures

A successful SLM solution must address the key architectural issues within a complex scenario of distributed and federated processes:

- Measurements are accumulated from distributed locations and aggregated to provide an end-to-end view.
- The distribution of the management architecture itself easily scales as the logical network of managed application components grows. A distributed architecture allows IT to add measurement points and increase capacity as needed.
- In federated environments, the management system is noninvasive, allowing the enterprise to interact with services it does not control. (Noninvasive means it should not require modification of the managed services or the messages transmitted across the wire.)
- The SLM system provides remote monitoring capabilities to manage services hosted by other entities.
- In complex environments, the management solution cannot be platform or application specific, because business processes cross multiple platforms that might have their own unique way of ensuring QoS. Even so, it is wise to optimize the management solution for each platform where it will perform the role of a cross-platform intermediary (for example, Java or Microsoft .NET).
- The management solution leverages system-level information that is collected by traditional management products (such as IBM Tivoli or HP OpenView) to enhance service-level metrics and provide a cohesive perspective below the services interface.

Close the Loop by Analyzing Historical Information

As the IT team manages its processes over time, historical knowledge will give them added insight—including the ability to identify bottlenecks, fine-tune SLOs, and observe changes in system health. Enterprises can build historical insight into these distributed applications and use this insight to understand clearly the business impact of varying service levels.

Continuing with the manufacturing company example provided earlier, what if IT discovers that whenever the tax calculation service goes down, it affects the order fulfillment process? Because they can track the average dollar volume and number of orders processed per hour, they can better understand the revenue and productivity lost due to the service failure. If multiple processes are affected, IT will also clearly understand the total impact on the business.

Armed with service-level metrics derived from system- and application-level knowledge, IT can fine-tune their systems over time to align them closely with the business objectives of the enterprise.

Conclusion

Enterprises will not gain the full benefits offered by distributed applications without first achieving the level of control and predictability provided by SLM. This white paper has covered three primary issues with respect to SLM for composite applications:

- Companies need to experiment with SLM early in their adoption of composite applications to ensure that they can effectively manage high-value business processes in distributed environments. By starting with pilots and proofs of concept, companies will be in a position to assemble increasingly complex and business-critical services-based systems.
- Enterprises must take advantage of the opportunity to combine system- and application-level information to manage QoS in a manner that addresses business objectives.
- To run core business processes on SOAs, organizations will be required to effectively manage these complex environments by implementing formal or informal SLM. With SLM, enterprises can clearly understand the business impact service levels have on revenue and productivity, and easily diagnose and mitigate service problems within business processes.

  SLM for distributed applications provides clear benefits by helping enterprises

  - Enforce end-to-end quality objectives that align with business goals
  - Effectively plan for growth and reduce business risk
  - Set realistic expectations with all stakeholders inside and outside the enterprise
  - Justify additional investments

SLM can be the difference between the success and failure of distributed applications, such as SOA and cloud-based systems.

Managing Service Levels in Distributed Applications
March 2010

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

SOFTWARE. HARDWARE. COMPLETE.