Web Performance Management: An Overview

Summary

Web performance management software is evolving from a device-oriented to an application-oriented technology, with increasing automation of both diagnostic and corrective processes.

Table of Contents

- Technology Basics
- Technology Analysis
- Business Use
- Standards
- Selection Guidelines
- Technology Leaders
- Recommended Gartner Research
- Insight

List Of Tables

- Table 1: Web Performance Management
Technology Basics

Traditionally, programmers write applications that run on a single computer using a single language and write code that progresses in a sequential pattern. In this environment, managing performance may not be simple, but at least it involves a finite number of identifiable hardware and software components. When performance falters, the administrator checks a specific number of parameters in the system and storage logs or, failing all else, would blame it on the code.

In the twenty-first century, especially when writing for the World Wide Web, programmers often write applications that consist of a far greater number of components. This newer, more complex architecture has great flexibility, so that companies can mix and match software and hardware components. At the software level, this architecture enables programmers to incorporate pre-established routines with new code (custom written or purchased) and to deploy applications to new platforms. With this architecture, administrators can add capacity and create high availability relatively simply by adding components. At the hardware level, it enables interoperability among disparate hardware and operating system (OS) platforms.

The trade-off for this flexibility is a more complex administrative burden. There are many more components to monitor, more pieces that can go wrong and more connections that can fail. More significant, these systems have so much redundancy that it is often difficult to determine the specific path taken by any one message or transaction, or even the hardware and software components that were utilized in any particular task execution.

At this point in time, most midsize and large companies have performance monitors in place for their mission-critical and large-scale systems. Their established management tools have limitations within a distributed application environment, however. In particular, it is difficult to pinpoint the exact cause of a performance problem. When a problem occurs, the administrator must hunt among many components to find the source of the problem.

Traditional performance monitors are designed to measure individual components. Administrators can easily determine whether machine X (server X or router Y) is doing well or badly today. However, it is less easy to determine whether or not an entire application or system is doing well. As a result, administrators are often not aware of performance problems until complaints arrive from end users. Once these complaints start coming in, the administrators have to deal with the complainers, listening to each one individually and appeasing them, estimating repair time and so on. How much easier their jobs would be if they could avoid the problems altogether, with advance warnings of impending bottlenecks, and repair them before the complaints begin.

Newer Web performance management tools attempt to address these issues. They present to the administrator aggregate views of different infrastructure components working together, tying the business applications to the underlying infrastructure and end-user experience. These tools drill down from end-user issues to specific performance statistics of the underlying applications, systems and network devices that comprise an e-service. They filter the data and correlate performance data to specific applications.

Web performance management products address three problems that network administrators face today:

• Traditional tools for measuring network performance generate a large number of statistics, only a few of which are relevant and useful in any particular situation. The number of generated statistics can be overwhelming in highly distributed, component-based environments.
Web Performance Management: An Overview

- Traditional performance tools deliver component-oriented information, whereas network administrators require that the technical information be correlated with application information.

- The pool of skilled administrators is declining as the most experienced workers reach retirement age or are laid off, or are unavailable because there just aren’t many of them, or are too expensive and younger administrators have a less technical, more business-oriented focus.

Current-generation Web performance management products address these problems by aggregating data from many sources and presenting it to the administrator in a logical manner. This logical presentation can be organized by application, by device type, by system cluster or by geography. In addition, they compare incoming values, pre-defined thresholds and sending notifications, or take some action in response to out-of-bounds values. They also store incoming data in a repository to present historical records and trending reports over time. In addition, they may have the ability to juxtapose several patterns on a single graph for correlation reports.

Performance Measurements

Today’s Web performance measurement products gather information from multiple sources. They may receive alerts from management platforms, such as Tivoli or Unicenter. They may read the log files of these platforms. Some products detect events and generate their own alerts.

Web performance management products employ some or all of the following approaches to gathering data:

- **Collect data from established sources.** They may read the log files of systems or DBMSs, alert files from device monitors, data collected into EMS systems. This information can be used to discern performance patterns over time and can be used for scheduling issues. In addition, this information can be used to determine the frequency and severity of specific malfunctions.

- **Display real-time status data.** This type of data collection is often used for Root Cause Analysis. This application typically provides a digital cockpit or dashboard that graphically displays the status of applications and color-coding alerts and that provides the ability to drill down to increasingly detailed information to pinpoint the cause of a malfunction.

- **Measure performance of transactions.** Whether live or simulated transactions are used, this form of data measures end-to-end transactions to mimic the end-user experience. The management software initiates user transactions from multiple agents residing on remote machines. This is an end-to-end test exercising all levels of multtier applications. The software executes test scripts provided by the administrator. Scripts can run automatically at pre-defined intervals (for example, every 15 minutes). These simulations provide a way for the administrator to be alerted to potential problems before end users become aware of them. However, they often are not capable of pinpointing the root cause of the problem; this will require sleuthing using other tools. Furthermore, when the test is a simulation, the administrator gets insight into only a subset of the total end-user experience.

Application Performance Management (APM)

APM products diagnose the “health” of a system by comparing performance data against pre-defined thresholds. Once a problem is detected, the policy engine performs intelligent policy-based decisions to diagnose and resolve problems. All offer a choice of media, such as fax, phone, e-mail and pager, to notify a human of the problem and perhaps include some diagnostic information and expert advice. Some can also initiate built-in actions, such as the ability to disable or enable servers, schedule traffic away from bottlenecks and failures or restart a service (that is, an application). Some will initiate customized routines
Web Performance Management: An Overview

built with a scripting language. Some products can even initiate application-specific or DBMS-specific actions, such as increasing a database connection pool size.

All of these products provide a screen-based interface through which the administrator defines rules-based policies for automatic response to problems. These definitions map events with appropriate responses.

Service-Level Management (SLM)

Traditional network-monitoring tools generated statistics about specific devices. Administrators today, however, need to understand performance in terms of end users and the applications that they access. Once administrators are able to obtain metrics in this form, they can deal more effectively with quality of service (QOS) issues. They can plan service-level objectives and enter into service-level agreements with their users to create greater satisfaction and productivity for their users. This style of network management is called “service-level management.”

In the phrase “service-level management,” the term “service” refers to an automated function that is expressed in business terms that, from a technical point of view, may consist of many different technology components. In this context, a “service” is an automated user function—a software component that “serves” a function to a requesting client. The “service” is a black box—a complete function that is served up as a unit. The specific technical implementation details—location of the server and disk file, language of the application code and so on—all are transparent to the client.

To implement service-level management, administrators need first to find a reliable way to gather performance metrics. They also define the applications in terms of its component hardware and software resources. Once the environment is defined in this way, the administrator goes on to identify problems that might cause performance problems and to specify policies for corrective actions. Once these maps, policies and rules are created, the administrator will have the means to initiate service-level objectives (SLOs) and service-level agreements (SLAs) with clients.

SLM software is still in the early phases of development. Recent releases of Web management software have the ability to consolidate log data from Web servers, application servers and network layers into a repository. For example, they can consolidate log data from the Web server and the Web application server with data from network and system tools into a single database. Similarly, they can consolidate data from multiple sites. Ideally, users should have the capability to drill down from the transaction layer into the network and system layers for the purpose of problem determination. This capability will likely become available in future releases of software.

Technology Analysis

Web performance management is an evolving technology layered on top of traditional network management methodology. It has two primary goals. The first is to simplify network management so that even complex multisite, multitiered componentized environments can be controlled. The second is to replace device-oriented and component-oriented performance metrics with application-oriented metrics that more directly reflect the end-user experience.

The technology for implementing the first goal is more mature. The gathering of performance metrics has been an essential technology for a long time. Event-driven technology, with generation of automatic responses and alerts, is a widespread phenomenon that is a high priority for system vendors, DBMS vendors, virtually and industry-wide phenomenon in the race toward a self-managing operating environment. At the present time, the technology is mature enough to support automatic alerts to a human when a problem occurs, with perhaps some expert advice on how to handle the problem. It is probably too
Web Performance Management: An Overview

soon to expect automatic repair of the problem without at least the approval of a human. With the collective intelligence of so many industry segments participating, however, one can expect a large measure of this goal to be achieved within the next five to seven years.

SLM, however, is somewhat less mature. Though the phrase QOS has been around for about five years, nevertheless, SLM is still in the early stages of development. The industry has not even established a standard approach to measuring performance at the service level. Similarly, early adopters are still experimenting with different types of SLAs to see how they can be appropriately designed. At a technology level, some companies have incorporated a repository into their products that stores information from multiple sources, including both hardware and software such as application servers. Automatic drill-down is not yet available, however; the administrator must write custom queries to correlate information about the various relevant components. Some Web management products have begun to provide user interfaces for entering information about components and the relationships between them; this is by no means a finished effort, however.

Business Use

Web performance management products are designed for multisite, multitiered component-based environments where traditional device-oriented performance management tools are not sufficient for managing the network environment. These products are most applicable to Web servers but are also appropriate for other component-oriented application environments.

These products, and especially SLM, are particularly important in business-to-consumer (B2C) applications where the Web presentation essentially replaces the displays and the personnel of a physical retail store. An unpleasant experience can discourage a customer from returning. Important for internal applications as well, a reflection of the company image is important to productivity and morale.

Standards

The Distributed Management Task Force (DMTF) is the important standards organization for management at this point in time. Its Web-Based Enterprise Management (WBEM) and Common Information Model (CIM) specifications form the basis of additional work coming out of the OASIS protocol management task force and the ongoing work of the Internet Engineering Task Force (IETF). These specifications are intended to provide a single set of semantics so that management tools of all sorts will be able to write to a common repository and share data with one another.

The DMTF was created to accelerate the adoption of management standards for distributed desktop, network, enterprise and Internet environments. It took over the WBEM standard from BMC Software, Cisco Systems, Compaq, Intel and Microsoft to provide an open forum for further development of this common standard for managing all systems, regardless of instrumentation type. The WBEM standard includes a data model, the CIM standard; an encoding specification, xmlCIM Encoding Specification; and a transport mechanism, CIM Operations over HTTP.

The CIM specification is the language and methodology for describing management data. The CIM schema includes models for Systems, Applications, Networks (LAN) and Devices. The CIM schema will enable applications from different developers on different platforms to describe management data in a standard format so that it can be shared among a variety of management applications.

The xmlCIM Encoding Specification defines XML elements, written in Document Type Definition (DTD), which can be used to represent CIM classes and instances. The CIM Operations over HTTP specification defines a mapping of CIM operations onto HTTP that allows implementations of CIM to interoperate in an open, standardized manner and completes the technologies that support WBEM.
Web Performance Management: An Overview

Several organizations have built on this DMTF core. The Open Group provides an open source reference implementation, called Pegasus, of the standard. In addition, IBM has developed a version for Linux called the “Standards-Based Linux Instrumentation for Manageability,” or “SBLIM” (pronounced “sublime”). In addition, Web management committees in IETF and OASIS are considering adopting WBEM and CIM standards for their own work.

Selection Guidelines

<table>
<thead>
<tr>
<th>Table 1: Web Performance Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
</tr>
<tr>
<td>Supported Platforms</td>
</tr>
<tr>
<td>Intelligent Management</td>
</tr>
<tr>
<td>SLM</td>
</tr>
<tr>
<td>Simulation</td>
</tr>
<tr>
<td>Easy Administration</td>
</tr>
</tbody>
</table>

Technology Leaders

Several dozen vendors of APM products market their wares alongside major electronics manufacturing services’ (EMS’) platforms, like HP OpenView and IBM Tivoli.

Innovations in Web performance management technology are coming out of some startup companies, like Empirix, ProactiveNet and Resonate. Another leader, Mercury Interactive, is somewhat more mature, with a 12-year history. Each of these vendors claims differentiation based on one or more of the innovative technologies described in this report.

Empirix

1430 Main St.
Waltham, MA 02451, U.S.A.
Tel: +1 781 993 8500
Fax: +1 781 993 8600

Established in September 2000 as a spin-off from Teradyne, the privately held Empirix offers a range of testing and monitoring products for Web, voice and network applications. It supports three Business Lines: Contact Center Test, Communications Infrastructure Test and Web Applications Test. The Web Applications Line consists of OneSight, an APM product. It is designed to monitor Web sites for performance bottlenecks and to help determine the root cause of the bottlenecks. FarSight is a hosted
Web Performance Management: An Overview

Web application monitoring service that monitors Web-site performance from a global network in real time. Partners include ATG, BEA, BroadVision, Fujitsu, IBM, Intel, Microsoft and Oracle.

**Mercury Interactive**

1325 Borregas Avenue
Sunnyvale, CA 94089, U.S.A.
Tel: +1 408 822 5200
Fax: +1 408 822 5300

Founded in 1989, Mercury Interactive highlights its relationships with business application vendors (SAP, Oracle, PeopleSoft and so on) and system integrators (EDS, CSC and others) in delivering their solutions. Mercury Interactive is a publicly held company with $400 million in annual revenue and more than 1,800 employees. Its Topaz product is a suite of integrated APM products that are designed to monitor Web sites for performance bottlenecks and to help in determining the root cause of performance problems.

**ProactiveNet**

2041 Mission College Blvd., Suite 260
Santa Clara, CA 95054, U.S.A.
Tel: +1 877 277 6686
Tel: +1 408 935 6800
Fax: +1 408 935 6888

Founded in 1998, the privately held ProactiveNet’s first product, an early warning network management solution, featured a decision support system for network management. The ProactiveNet Enterprise Solution is an APM product that is designed to simplify performance management for system and network administrators by helping to identify the root cause for slowdown or failure of applications. ProactiveNet has established strategic partnerships with BEA Systems, IBM, Keynote Systems, Rational Software and Sun Microsystems.

**Resonate**

385 Moffett Park Drive
Sunnyvale, CA 94089, U.S.A.
Tel: +1 408 548 5500
Fax: +1 408 822 5679

Founded in 1995, Resonate closely identifies its brand with the concept of SLM solutions. Its Resonate Commander product is an APM product that proactively locates, diagnoses and resolves outages and performance bottlenecks. The publicly held Resonate, with revenue of $10.9 million in 2002, partners with BEA Systems, IBM and Sun Microsystems.

**Recommended Gartner Research**

Vendor Rating: Empirix Emphasizes Customer Satisfaction, VDR-19-5149

Vendor Rating: Mercury Interactive, BDR-18-7561
Web Performance Management: An Overview

Mercury’s Growth Strategy Comes with Challenges, FT-18-4258

Mergers and Acquisitions Increase in the Software Industry, COM-19-1690

Insight

Web performance management software is evolving from a device-oriented technology to an application-oriented technology, with increasing automation of both diagnostic and corrective processes. The administrator is refocusing attention from equipment to the end-user experience. As automated forms of marketing and sales become major sources of revenue for corporations, the end-user interface becomes not simply a means to productivity but a strategic initiative for corporations. For these reasons, the need for Web performance management technology is very important to many enterprises.

Hedy Alban began her IT career as a programmer, reconstructing a relational database structure by analyzing hex dumps. Today, she brings to her customers almost two decades of experience in programming, system administration, research analysis and publication development. In addition to managing several publications targeting IT professionals, she has managed the delivery of information services to IBM, Hewlett-Packard, Digital Equipment, GTE, Software AG, AT&T, Encore Computer and others. Hedy is a member of the OASIS Management Protocol Technical Committee.