



## **WHITE PAPER**

# Application Performance Management

**Application Availability as part of a Performance Management Practice**

## Application Availability as part of a Performance Management Practice

Overview .....	3
A typical Web-based application environment.....	4
Challenges managing availability and performance in a complex environment.....	5
An ideal availability and performance management solution .....	7
The Precise Synthetic Monitoring solution .....	10
Precise Synthetic monitoring Studio .....	10
Precise Synthetic monitoring management console .....	10
Precise agents.....	11
Health checks from multiple geographical locations.....	13
A simple dashboard screen .....	13
Drill-down into slow pages and page subcomponents .....	13
Application availability and performance history .....	13
Support for monitors.....	14
Firsthand resolution to the problem.....	14
Easy installation and configuration .....	14
Conclusion .....	15

### Overview

Enterprise Web-based applications have become an essential component of every business. These applications are mission-critical, executing various organizational business processes. Whether it's a pure Web-based multi-tier application or a front end to an existing ERP or legacy system, the availability of the application and the content of transaction output need to be monitored 24x7, to ensure the reliable and correct operation of the business processes.

One of the biggest challenges facing IT today when it comes to Web-based applications is the requirement to constantly check, 24x7, the availability, reliability, and performance of the applications. In other words, IT must measure the actual service being delivered to the end user and then determine the end-to-end service-level contributors in order to find bottlenecks and adhere to service-level agreements (SLAs).

Precise provides a way to quickly, efficiently, and unobtrusively monitor the availability, reliability, and performance of applications as well as capture the hard-to-get URL-to-SQL application metrics. Precise's synthetic monitoring solution presents these important metrics in a manner that enables crisp communication and rapid proactive or reactive detection, correction, and verification throughout the application lifecycle. Precise's synthetic monitoring solution is part of the Precise product family. It enables users to combine the management of both synthetic and real end-user transactions.

Precise's synthetic monitoring solution allows IT managers to proactively manage the availability and performance of their applications. This paper describes the unique capabilities of the Precise solution and highlights why a combination of synthetic and real end-user transactions must be measured and managed to provide comprehensive application performance management.

### A typical Web-based application environment

Figure 1 depicts a typical and “simple” Web application environment.

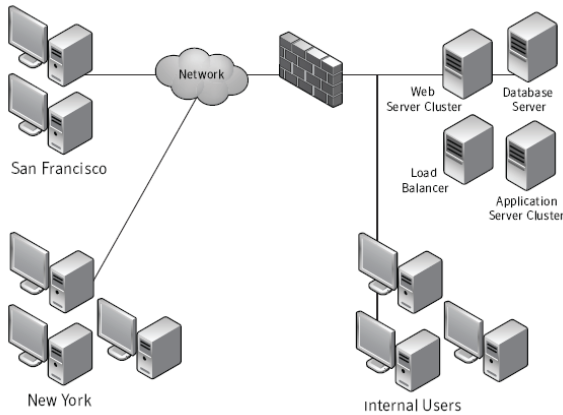


Figure 1 – A typical Web-based application environment

End users located at multiple geographical locations access the application and execute transactions both from inside and outside the firewall. When users execute transactions against the Web-based application, transactions traverse network links, routers, firewalls, load balancers, Web servers, application server clusters, database servers, and storage components. This path causes increased complexity for IT managers and prolongs the time to resolution when problems arise.

There are many potential points where a transaction can fail, for example:

- Network time-outs on a congested network link
- End users who cannot access the application because of a database failure
- Unacceptable response time because of a misconfigured application cluster
- Long delays because of an improper load-balancing policy
- Architectural application problems causing accessibility and performance failures
- Accessible application with incorrect content display, blank pages, wrong pages, etc.

Therefore it is important to test the transaction availability and response time as the end user experiences it, in addition to validating that the content is displayed correctly.

Checking the availability of each component along the transaction path is not good enough, since a component could be up and running, but not functioning properly or interacting with other components properly. For example, an Oracle® database could be available with processes running, but not answering properly to queries. Therefore, to have a “complete” availability assessment, what matters is the end-user point of view. Is the application transaction accessible in a timely way to the end user? Does it display the right information?

### Challenges managing availability and performance in a complex environment

The management of this “simple” Web-based application, spread across multiple tiers, demands a comprehensive yet straightforward way to check if the application is accessible to all users at different locations as well as ensuring that the right service-level objectives are being met.

Today’s complex, multi-tiered Web-based applications introduce several new challenges, which include:

- Additional tiers and new technologies
- Migration from traditional client/server or mainframe applications
- Inadequate system and application management solutions
- Scalability

The introduction of additional tiers and new technologies (e.g., load balancers, Web servers, application server clusters) adds new challenges. The new environments are distributed and much more complex compared with the simple, traditional client/server two-tier architectures. The multi-tier architecture makes it difficult to pinpoint the source of a problem. These Web applications demand more responsiveness and around-the-clock high performance than their less complex internal predecessor, client/server applications, in order to adhere to SLAs.

Many Web-based applications have migrated from traditional client/server or mainframe applications. The new applications have an expectation of performance. Web-based applications have to perform as well as, if not better than, traditional applications.

Some applications are not managed properly and suffer from inadequate system and application management solutions. Existing monitoring solutions look at each infrastructure component separately, but do not monitor the service being delivered to the end user. Without understanding the end-user experience, the application manager might waste unnecessary resources trying to tune the application. The application manager needs to fight with a vast number of events that seem to be unrelated but are the symptoms of a single failure. For example, if there are database connection errors, multiple events could be generated on the database as well as on adjacent tiers. The same is true for network and other problems.

What happens when the connection to and context of the business is missing? Existing system management tools alert organizations to the failure of infrastructure components but do not provide information on the business impact of such a failure. Monitoring business transactions immediately reflects the impact on the business. For example, by defining the following transaction:

- Log in to an account
- Conduct a few operations such as check account balance, transfer funds, etc.

## Application Availability as part of a Performance Management Practice

- Log out

Running the transaction 24x7 and simulating an end user will indicate if the transaction is available in the case of a component failure. The impact can be seen immediately if a component of the transaction fails, is not accessible, produces the incorrect page and data, or has a slow response time.

Detecting if the application is not available is only the beginning of problem resolution. The next-step analysis is very important to focus quickly on the problem and solve it. Existing monitoring tools do not provide an efficient way to isolate the problem symptom and identify and analyze the root cause. Existing tools either lack the granularity of data required for problem resolution or do not correlate the different events and symptoms to pinpoint the problem.

Scalability is also a challenge. With the introduction of Web-based applications, mid to large accounts have tens to hundreds of Web applications, each running many different transactions. Availability and performance solutions have to scale up in order to monitor the vast amount of Web applications and transactions as well as the individual infrastructure components (e.g., network, Web servers, application servers, databases) that serve these applications.

### An ideal availability and performance management solution

In order to solve the problems identified, a Web-based availability and performance solution should be able to easily define and manage a business transaction service level. In other words, it must be possible to simulate an end-user transaction and run it continuously, 24x7, to check the transaction availability, reliability, and performance. In addition, it must be possible to provide reports and alerts that show the delivered service levels over time.

Typical enterprises consist of a central location and multiple remote geographical locations. An ideal availability and performance management solution should support running transactions from multiple locations, which include:

- From inside the firewall, to test datacenter availability
- From designated multiple geographical locations, to test accessibility of transactions from those key locations

There should be a straightforward and simple way to look at the availability, performance, and reliability of the monitored business transactions from all defined geographical locations. In addition, statistics such as mean time to recovery, mean time between failures, and so on, should be provided.

The solution should verify:

- **Accessibility**—ensuring that transactions are accessible from both inside and outside the firewall.
- **Reliability**—confirming that transactions produce the correct output. Validation for certain objects should be supported to verify generated transaction output.
- **Response time**—response times must fall within the defined service-level objectives (SLOs).

Being able to break down a Web transaction into its pages and each page into its subcomponents is important in analyzing SLA breaches due to long response times. The solution should provide details on the slow pages in a Web transaction as well as be able to focus on the components that cause the delay.

## Application Availability as part of a Performance Management Practice

There should be support for monitoring the availability of underlying application and infrastructure components. For example, availability checks for J2EE™ application servers and SQL checks for the database should be conducted. There are many infrastructure and application components that should be supported. In order to identify application availability problems, an initial mandatory monitor list should include the following:

- HTTP URLs checks
- J2EE EJB™ checks
- Database checks (ODBC/JDBC)
- Network ICMP Ping
- Enterprise applications
- UNIX and Microsoft Windows® OS CPU, memory, disk metrics
- FTP
- LDAP
- SMTP
- POP3
- IMAP4
- DNS
- RADIUS
- Generic TCP/IP
- MAPI
- Citrix
- SOAP
- Generic shell

The ideal solution should be scalable to support many instances of the monitored list defined above. Scalability means being able to run many transactions against multiple different applications and infrastructure components as well as being able to support high scheduling requirements (i.e., small enough intervals between each run of a transaction).

There should be firsthand resolution for the “suspected” component causing a transaction failure. Operations managers want to know which tier is causing the problem—whether it is the network, Web server tier, application server cluster, or database. A simple check of each component can tell if the component is available. Such a check can pinpoint a component failure, which enables quick resolution and a shorter MTTR (mean time to repair). For more complex problems, such as reduced response times and performance bottlenecks, a deep drill-down is required.

Additional or “next-step” analysis must be provided in order to isolate the definitive root cause. To solve performance issues and adhere to SLAs, very detailed data should be collected from multiple tiers along the transaction path. Granular data should be collected at the Web server tier, application server cluster, and database and storage components. Then this data should be correlated to enable response-time segmentation across tiers. When the problematic component is identified, an in-depth analysis of this component should be provided. This means enabling drill-down to the root cause of the problem. For example, in J2EE servers, an in-depth monitoring of the different JSP™, EJB, and JDBC™ connections down to the method level should be provided. In the case of databases, an in-depth analysis of SQL statements and objects should be provided.

The solution should be durable enough to cope with failures like losing connection between an agent (POP) located at a remote geographical location and the management console. The agent should

## **Application Availability as part of a Performance Management Practice**

continue to work independently to test the application availability and, when the connection is recovered, to sync up with the management console.

The solution should be easy to install and configure. Configuration means an easy way to define all the monitored components, including application transactions and infrastructure components, as well as being able to integrate with management platforms.

## **Application Availability as part of a Performance Management Practice**

### **The Precise Synthetic Monitoring solution**

Precise's synthetic monitoring solution provides application availability monitoring using synthetic transactions. With its intuitive Web-based management interface, Precise's synthetic monitoring solution makes it easy to define and record end-user transactions. Focusing on Web-based applications, the solution monitors the availability, performance, and reliability of key business processes. Application availability and performance can be monitored 24x7. Routine application health checks can be scheduled as needed. Availability by unique geographic location can be tested and measured.

Precise's synthetic monitoring solution provides an authoring environment, called Studio, for business transaction recording. Using Precise, IT managers can record a business transaction against a Web-based application. Validations could be asserted to verify existence of different objects on the Web page. Then this recorded transaction would be used in the execution environment.

The execution environment includes:

- A focal management console that collects the data from multiple distributed agents. The management console visualizes the application status at the different locations.
- Agents, located at multiple geographical locations, that run the prerecorded transaction. The agents verify the accessibility, content, and response times of a transaction.

### **Precise Synthetic monitoring Studio**

Studio is a script-less transaction recording engine that enables:

- Definition of a business transaction by virtually working in a Web browser-like interface while conducting usual operations against a Web site. Any user that knows how to work with Internet Explorer is able to record a business transaction.
- Addition of assertions to validate the content of the page, such as specific text objects or graphical components. This is done during the transaction definition phase.
- Visual testing of the recorded transaction, with the ability to see how it is played back before it is deployed.
- Recorded transactions to be deployed directly to the management console, which in turn populates the recorded transactions to agents for execution.
- The environment to maintain recorded transactions when the Web application is changed or updated.

### **Precise Synthetic monitoring management console**

The management console is the focal point for configuring, defining, and viewing the status of the Web applications.

## Application Availability as part of a Performance Management Practice

A typical scenario starts with the definition of a Web application and associating the Web application with prerecorded Studio transactions. Then the management console will send the prerecorded transaction to the different agents for execution. Agents simulate the real user activity and check for accessibility, content validation, and performance metrics such as page performance, object performance, HTTP errors, and more.

Precise's synthetic monitoring solution supports policies that define key behaviors such as:

- Alerting and action policies to be taken in case of a failure.
- Service-level agreements (SLAs) for page load time and transaction time. The SLA for page load time can also be defined automatically based on history data.
- Scheduling policy—when and how often to execute transactions.
- Storage policies—the granularity of the data kept and when to keep it.

Precise's synthetic monitoring solution supports firsthand resolution to isolate the problem. Given that the application is not accessible or SLAs are being breached, Precise's synthetic monitoring solution determines the most probable component that is causing the problem. This is done by checking the availability of each infrastructure component involved in the transaction execution, the network accessibility, and the Web server, application server, and database. When the problematic component is identified, only one correlated alert is sent, eliminating the flood of alerts to the management console caused by a single failure.

### Precise agents

Precise's synthetic monitoring solution agents actually do the work. They execute the prerecorded transactions according to the times defined in the scheduling policy. Each transaction is checked for accessibility, content validation, and performance metrics such as page load times and HTTP errors. Results are communicated back to the management console.

Agents can run in a clustered mode when increased scalability is needed to support numerous Web applications and transactions. In this mode a "master" agent, the management console, distributes the work to other members in the cluster.

Once applications, transactions, and policies are defined in the management console, it is very easy to bring up new agents at different geographical locations. Once the agent is installed and running, the definitions and policies are populated from the management console.

Precise's synthetic monitoring solution can be used to define and easily manage business transaction SLAs. This feature makes it possible to do the following:

- Define a business application or group of related business applications
- Define the service-level objectives (SLOs) for those business transactions

## **Application Availability as part of a Performance Management Practice**

- Create different policies regarding the characteristics of business transaction execution, such as how often they run, how much result data to keep, retry policy, and so on
- Monitor 24x7 the availability and performance of those business applications and receive notification when an SLA breach occurs

## **Application Availability as part of a Performance Management Practice**

### **Health checks from multiple geographical locations**

Availability, reliability, and performance health checks should be conducted from multiple geographical locations as well as from inside and outside the firewall. Checks inside the firewall make sure that the application is available first to local users, and then that remote locations can access the application. Therefore, checks from different locations confirm that the application is accessible to users from that location. The default configuration is to run all checks from all locations. Through the use of policies, the product also offers flexibility to run certain transactions from certain locations or from all locations.

### **A simple dashboard screen**

A simple dashboard screen shows the application availability and performance status of multiple geographical locations.

In addition, the a status screen can show the Web applications' availability and performance at each of the geographical locations.

Clicking on a potentially problematic application will display the status of the application's transactions, thereby providing more details on the problem, whether it's a network or protocol level error, content validation error, or a performance problem.

### **Drill-down into slow pages and page subcomponents**

Precise's synthetic monitoring solution provides the means to break a Web transaction into its pages and focus on the slow pages.

The server infrastructure is a complex multi-tier environment, composed of several tiers such as a Web server tier, J2EE application tier, and a database tier.

For more insight into the page subcomponents that cause the long response time, Precise breaks the problematic page into sub subcomponents, such as images or jsp's. To get to the root cause of the problem requires the use of other Precise components, such as Precise for SAP, or Precise for J2EE, which are described in separate white papers.

### **Application availability and performance history**

In order to analyze the system performance and availability data over time, Precise's synthetic monitoring solution provides the Activity screen where both the performance and the availability data are presented by day or by week.

Performance and availability information is available either for an application or for a geographical location. The data may also be filtered by a specific geographical location or a specific application. The activity data is calculated according to the performance SLA definitions and the availability SLA

## **Application Availability as part of a Performance Management Practice**

definitions that were applied for the applications. Alternatively it is possible to view the worst error information.

Clicking on a potentially problematic application from a specific day or week will display the performance and availability information of the application's transactions in that time period, providing more details on the SLA compliance, and the percent of successful runs.

Clicking on a potentially problematic monitor availability indicator will display the availability graph throughout this time period. Clicking on a potentially problematic monitor performance indicator will display a graph with the performance information of this monitor from each of the geographical locations.

Clicking on each one of the graph points will lead to the transaction analysis and to the page analysis graphs.

### **Support for monitors**

Precise's synthetic monitoring solution supports a rich set of monitors, which are divided into two main groups. The first is Web HTTP transactions, and the second is infrastructure monitors such as J2EE EJB checks (WebSphere and WebLogic), database checks (ODBC/JDBC), Network ICMP Ping, enterprise applications, UNIX and Microsoft Windows OS CPU, memory, disk metrics, FTP, LDAP, SMTP, POP3, IMAP4, RADIUS, DNS, Generic TCP/IP, MAPI, Citrix, SOAP, and Generic shell.

### **Firsthand resolution to the problem**

Having a large number of monitors enables Precise's synthetic monitoring solution to check both the synthetic end-user transaction as well as different infrastructure components along the transaction path. In doing so, Precise's synthetic monitoring solution provides a firsthand problem diagnostics tool. For example, when an application is unavailable, Precise's synthetic monitoring solution can issue a check to the database component, then to the J2EE component, and to the Web and network components. This provides a health check of the different components when a failure occurs. For a deeper analysis, the additional parts of Precise are required.

### **Easy installation and configuration**

Precise's synthetic monitoring solution comes with a set of policies that are easy to define and apply to all running applications and transactions. It is built in an object-oriented fashion, so a new transaction or application added to the policy group automatically inherits the attributes and characteristics of the defined policy. Policies exist for different kinds of transaction execution behavior such as scheduling, storage results, assertions, retries, reporting, alerting, integration with Enterprise Systems Management platforms, and more. Moreover, installing a new agent (POP) at a remote location is very simple. When a new agent is deployed at a remote location, it inherits all the applications, transactions, and policy definitions and starts to run immediately.

### Conclusion

Web-based applications are an essential component of every business. They are distributed, involve multiple tiers, and employ new technologies. Managing the availability of Web-based applications is challenging. Verifying that the right SLAs are being met is even more difficult.

Precise's synthetic monitoring solution is an application performance and availability management solution and a component of the broader Precise TPM solution. Precise's synthetic monitoring solution facilitates checking application transaction availability, both from inside and outside the firewall and from multiple geographical locations. The solution simulates an end user, mimicking end-user activity and testing for accessibility, content, and performance of Web applications.

Precise's synthetic monitoring solution allows IT organizations to take a proactive approach to application availability. Critical business transactions and service levels can be profiled and routinely monitored. SLA metrics can be easily established and proactively managed. As service levels are not met or as the performance of key business transactions begin to deteriorate, IT staffs are alerted of the problem before end users are impacted. In order to drill down to the root cause of the problem, additional components of Precise should be used. Using Precise helps companies manage service levels, contain costs, maximize the efficiency of the operational investment, and improve the end-user experience.

## Application Availability as part of a Performance Management Practice

### About Precise Software Solutions

#### Commitment, Focus, Experience

We created the Application Performance Management market and this is all we do.

For over 15 years Precise Software Solutions has helped our Global 2000 customers manage business performance in complex, heterogeneous environments, assuring availability and business continuity. Precise offers a complete solution – from discovery through ongoing management – that allows our customers to focus on their core business. We offer the broadest platform support, in terms of enterprise application, operating system, database, and development environment coverage. Precise is the solution of choice for IT as an organization-wide standard for application management.

Visit our Web site

[www.precise.com](http://www.precise.com)

To speak with a Product Specialist in the U.S.

Call toll-free 1 (877) 845 1886. For specific country offices and contact numbers, please visit our Web site.