Whitepaper

PremiTech
See IT the way your users do!

Optimizing IT Performance

The case for a new 3rd generation of End User-Based Performance Management solutions

End User-based Performance Management products are the newest generation of solutions for IT performance optimization. Performance Guard delivers the industry's only end-to-end Performance Management solution enabling IT to save significant budget and human resources in service level monitoring, capacity planning and IT performance issue resolution through a real-time holistic view into the entire IT installation's efficiency.

PremiTech
See IT the way your users do!

Industry Analysts such as Forrester Research have identified the need to augment the existing data center management solutions with the newest generation performance management: end-user based performance management.

Introduction

Over the last few years, corporations have been under increased pressure to optimize resource utilization while at the same time adapting more quickly to changes required by the business. This pressure has direct impact on IT organizations faced with more complex installations that need to be implemented, enhanced and managed with fewer allocated resources. Most organizations have invested in up-to-date data center performance management solutions that can be configured to deliver a framework for measuring and automating management tasks and handling network and server alerts.

However, the vast majority of time spent in IT organizations is not used on managing individual servers and networks, but instead on cross platform service-level monitoring, performance optimization, capacity planning and resolving end user issues. For example, it is very difficult for companies that have deployed traditional three-tier infrastructures to detect why a remote-office user is experiencing performance problems only during parts of the day. Such periodic or non-straight forward management issues are among the hardest and thus most time-consuming and expensive to solve.

Since most enterprise management solutions are based on configuring the system for things like device and server alerts, these complex issues are practically impossible to solve using any management architecture that require configuration or modelled simulations to set up a certain measurement point. It would require millions of alternative cross-referencing scenarios to configure every possible combination of measurement points, eliminating the ability to manage periodic issues without an end user-based monitoring solution.

Based on research indicating 80-90% of IT support resources are based on fact finding, Industry Analysts such as Forrester Research has suggested the need to augment the existing data center management solutions with a new generation of end user-based performance management solutions that measure the health of the entire IT infrastructure from servers to networks to end user PCs and devices, and use the measured data to deliver CIO level performance intelligence as well as operational monitoring needs to resolve performance, capacity and configuration constraints and cross system error resolution.

This paper will examine the elements involved in gaining an actual view of end user application performance to quickly detect IT system issues, and how Performance Guard™ from PremiTech can help companies proactively manage their IT infrastructure.

Systems Management Challenges

There are two main issues facing IT and helpdesk managers today: The first issue is centred on obtaining a holistic view of the enterprise IT performance by quickly, often in real-time, acquiring the data needed to accurately identify and solve IT problems. The second issue is centered on the ability to proactively leverage this information to solve issues before they impact the end user - whether it's resolving problems or optimizing the capacity available.

While it's a fact that monitoring and measuring end user application performance will help to ensure IT system availability, the conventional data center-based solutions fall short of delivering the exact information needed to identify and resolve cross-platform system issues quickly. Either the configuration options are missing from the infrastructure management solutions leaving manual processes as options, or the available tools are based on simulations that place even more test based load on the systems and yet still based on a given identified set-up, not the real environment. Simulated data simply does not provide the real picture needed to gain the view that your end user experiences and accurately detect period degradation in response time or trending that results in alerts of capacity problems lying ahead.

Typically, when a system performance problem arises, the IT manager needs to pursue a time-consuming, expensive and tedious process to identify and resolve a given issue. Industry statistics demonstrate that, based on a system of trial-and-error, the majority of the time it takes to solve a given problem is spent trying to find its exact location or cause. This not only frustrates highly skilled help-desk second- and third level support staff, users and customers, but it also costs organizations a significant amount of operational budgets as well as expenses on unnecessary downtime, remediation systems and support. To effectively address this issue, IT managers need the ability to quickly and accurately identify problems at the atomic level of application, network and systems and proactively manage IT systems based on actual information.

Once this intelligence is available, companies can optimize application service quality and availability, while at the same time reduce time and costs to deliver the service. For the IT executive management team, the ability to have this level of insight not only enable very accurate reporting to the business on how the service is rendered, but also delivers the essential planning intelligence needed to efficiently document the reasons for investments needed to proactively run the IT support

End user-based end-to-end performance management give the IT executive management team the ability to very accurately report to the business on how the service is rendered as well as the planning information essential to accurately focus and document where investments are required.

of the organization.

Become proactive with end user-based performance management

Most data center management systems manage server resources very well. However, for end-to-end management, the IT manager needs the capability to understand the end user experience and detect application availability and performance problems based on the users experience -it is this experience that drive helpdesk requests, the number of alarms and the satisfaction levels of the users that will judge IT's success. .

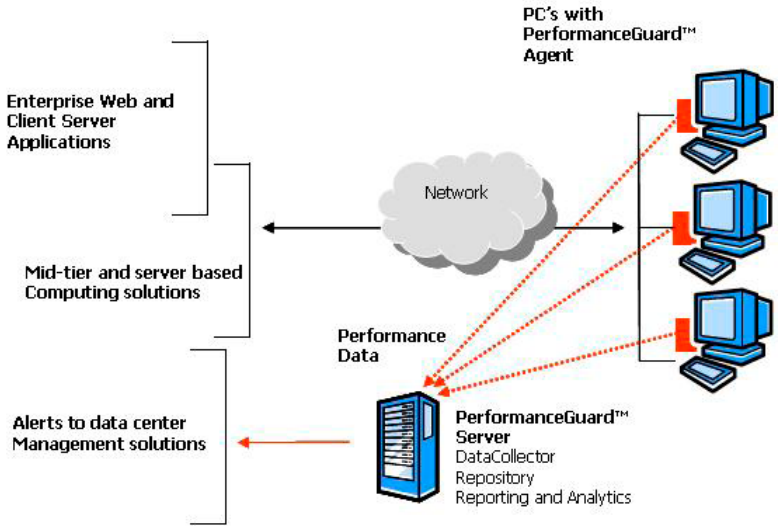
End user performance management can be done either as appliances that reside in the data center or as agents on the individual desktops. Unlike new third generation desktop based agents, these appliances fall short by nature, as they are data center based devices and report data center experience issues.

Passive agents sit on the end user device, measure all local, network, server and application activity and non-intrusively report this to a central IT performance intelligence database server. They deliver an actual, real-time view of the end user experience. IT management can use this intelligence for a multitude of tasks such as service level measurement and resource planning and second- and third level helpdesk support staff can use it to quickly identify problem areas while at the same time use the intelligence to proactively detect a degradation trend in IT service delivery. Passive agents also eliminate the introduction of additional load on the existing IT infrastructure that is typically seen with alternative solutions.

The PremiTech Performance Guard™ solution has garnered industry acceptance as a leading end user-based performance management solution. It is an agent based solution that delivers a complete application, server and network monitoring solution coupled with monitoring of essential end user PC resources. It provides exact measurements of the end user experience including an up-to-the-minute overview of all components and applications touched by any end user.

The information presented by Performance Guard™ enables IT personnel to run performance reports in every possible combination - among agents or groups of agents, locations, applications and components - to quickly identify the root cause of performance issues. In addition to fire-drill resolution, technicians can perform before and after scenarios to determine what periodic problems actually consisted of, how they should be solved, and the impact of the proposed solution. This helps companies ensure they make the right investments in IT infrastructure to deliver the proper performance based on set baselines.

The Performance Guard™ Solution



The Performance Guard™ Architecture consists of an agent on each PC and a collection/repository/reporting server.

The Performance Guard™ Agent

The Performance Guard™ agent is installed on each PC using standard software distribution methods. It passively monitors and collects real end user performance measurements using a light-weight agent installed on individual machines. The agent collects response time and other performance metrics on these machines, it assembles the data into local reports and at predefined time intervals, and then a collection of reports is sent to the Performance Guard™ DataCollector.

Performance Guard™ Server

The Performance Guard™ Server is responsible for inputting, storing, aggregating and displaying all of the vital statistics and performance measurements that are collected by the agents. This is performed by the following subcomponents:

o **The DataCollector** unpacks the reports and inserts the data in the Performance Guard™ Data Repository (database). Communication between the agents and the DataCollector is performed using TCP/IP.

o **The Data Repository** is a standard database server based on either Microsoft SQLServer or Oracle. The DataCollector and the Data Repository are connected using the open JDBC

interface. When the back-end database is an Oracle database, the JDBC connection is implemented as a SQLNet connection.

o **The Display** accesses performance data through a user-friendly Web front-end containing a large set of pre-defined reports on the information stored in the Performance Guard ™ Database, These reports can be presented in a multitude of options such as data tables, graphs, and configuration diagrams.

Once performance data starts populating the database, IT managers can take advantage of the advanced Web based reporting framework to receive pre-built real-time reports using a Web browser. If more advanced reporting is required, these can be built either using the same browser-based system or by reading the reporting database from advanced reporting tools such as Crystal Reports.

Simple Installation and Set-up

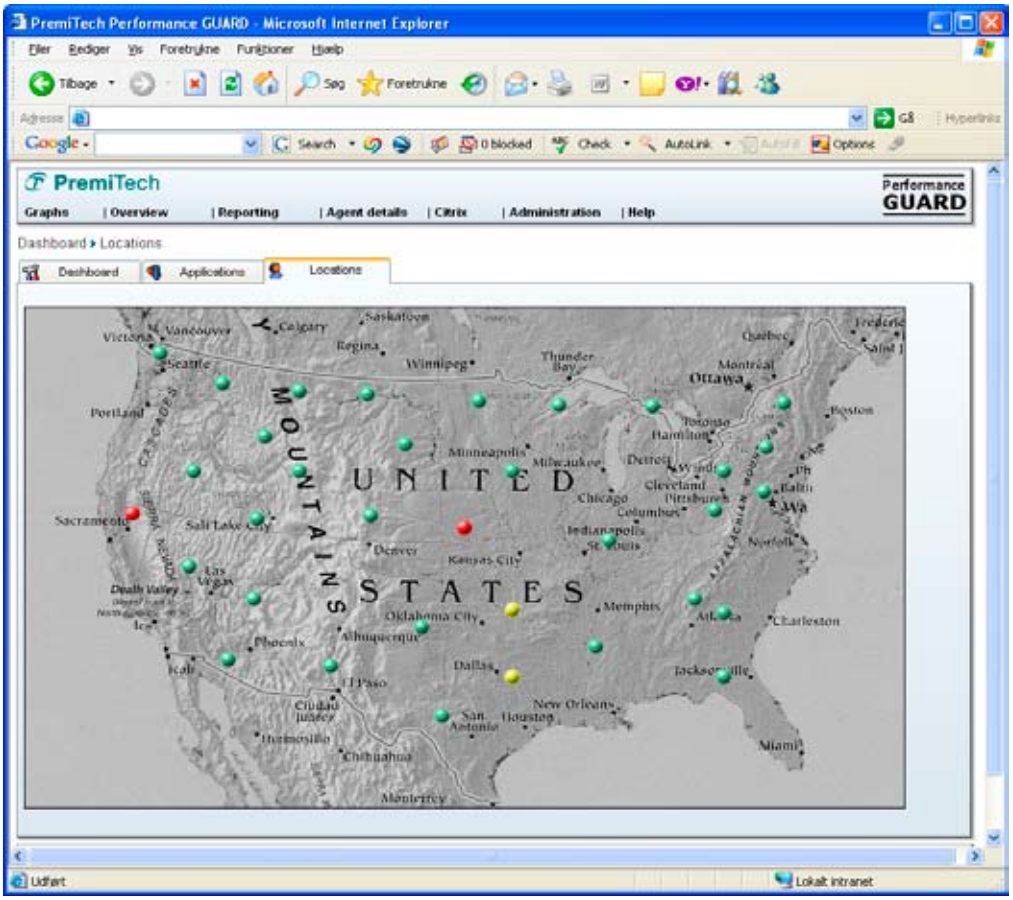
The central Performance Guard™ Server, dedicated to the collection, processing and display of data is installed with a wizard interface. Agents are distributed to all clients using an ordinary software distribution tool or a link to an Install file. The agents install automatically and begin reporting to the Performance Guard™ Server immediately after installation. In standard installations, the agent will consume less than 1% of the end user devices CPU and memory resources. . All maintenance of the agent is centralized and any modifications to the agent configuration are done automatically when the agents report data to the server.

For the first time ever, IT can get its own Executive Dashboards that show the end-to-end performance of all systems and networks over time or in real time as experienced by the users including critical service degregagation information with advanced drill down detailed information only a click away.

Performance Guard™ Measurements Provide Fast Problem Detection

In order to understand the response times reported by Performance Guard™, it is important to understand the fundamentals of how Performance Guard™ measures response time. Response time is measured at the network level, and more specifically, at the TCP/UDP IP level. TCP/IP is the most commonly used protocol today. Only packets that contain application data are considered when Performance Guard™ measures response times. When a client sends a request to a server it sends one or more packets to the server. The server then processes the request and sends one or more packets back to the client.

Performance Guard™ measured response time is the time elapsed between the last sent client request-packet until the first reply-packet is received from the server.



An agent aggregates response time measurements based on the server and TCP port on which the server and client communicate. For example, response times for all communication with a specific Web server within a single report period will include the following information (which will be sent to the DataCollector):

- o Accumulated response time
- o Number of connections
- o Number of trains sent and received (The term train is used when multiple packets are sent in one direction before any packets are sent in the opposite directions).
- o Number of bytes sent and received

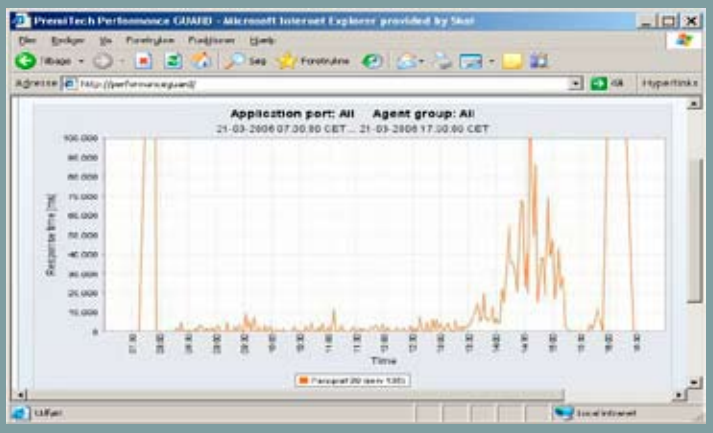
The response time for the combination of <agent, server, service> is then calculated by the back end as an accumulated response time divided by the number of received trains. In order to display response time from measurements taken on multiple clients, it is necessary to aggregate the data further. In this case, the response time concerning a group of agents and a specific <server, service> is calculated as the sum of accumulated response times divided by the sum of received trains for all agents within the group.

These metrics become the foundation of all aggregated data which is utilized to deliver the corresponding graphs, comparisons, and other related measurements.

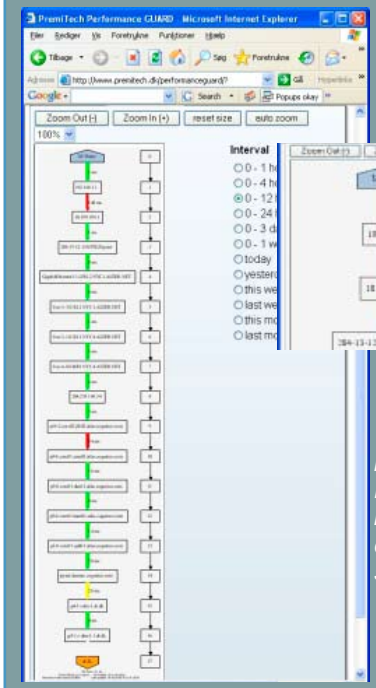
When SAP Performance issue is unexpected - a real life example

Example: Resolve SAP performance issues
Identifying a specific performance issue for an application such as SAP delivered via Citrix over a WAN has so many variables that it is very difficult to identify. Often, all silo's will report systems ok, yet the end-user experiences poor performance. End-user performance management is the only way to quickly identify the performance bottleneck and resolve it.

Illustration 1 shows a performance degregation from a set of SAP users in New York experiencing poor performance on the SAP system in Germany during afternoon peak hours.



The end-to-end monitoring diagram indicates two performance hot-spots.



Unexpectedly, the first problem is response between the end-users network segment and the internet.

The second »red« problem area is the tranmission time between New York and Europe. The delay is due to the distance, not capacity, and as such cannot be solved by more bandwidth.

It is the combination of the two delays that made SAP grind to a halt. However, the default solution by adding more expensive SAP servers or capacity over the atlantic will be very expensive and still not solve the issue. Better local network capacity in New York will reduce the response time to an acceptable level solve the problem at a much lower cost.

This information is used by IT for fast error detection, proactive problem-solving on the component level, and ongoing system performance tuning.

Quickly Isolate Local Resource Issues and Problematic Processes

The Performance Guard™ Agent captures the following local (PC) metrics:

Resource Metrics

CPU Usage: percentage of CPU time not spent running idle

Free Physical Memory: Amount of physical memory available for allocation

Free Paging File: Amount of paging file space available for allocation

Virtual Memory: Amount of virtual memory available for allocation

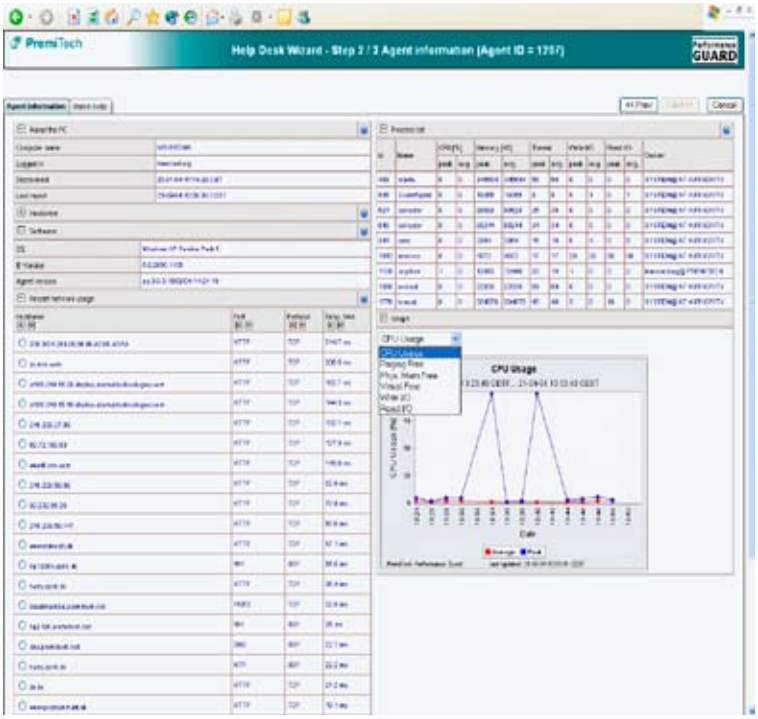
Local Process Metrics

CPU Usage: Percentage of available CPU time used for a certain process

Memory Usage: Number of bytes that a process has allocated that can not be shared with other processes

Thread Count: Number of operating system threads used by a process

Handle Count: Number of operating system (Windows) handles used by a process



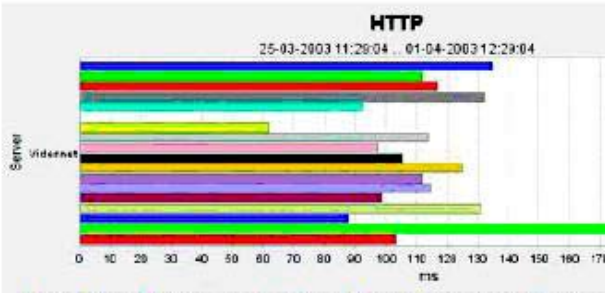
Values for these metrics are sampled at regular (adjustable) intervals. The maximum value and average value is reported for each sampled interval timeframe.

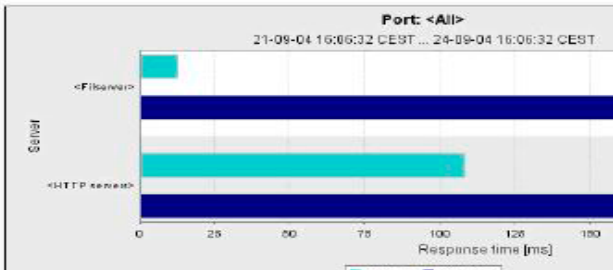
Windows has many idle processes running that are of no interest; therefore, a limit may be set on the number of processes monitored by Performance Guard™. This can be performed on the CPU, memory and I/O level, or the exact processes to be monitored can be specified.

Component Groupings Key to Problem Resolution

The most important configuration requirement in maintaining the Performance Guard™ system is in the grouping of agents. Data aggregation is performed at the group level. Agents can be grouped by more than one dimension. For that reason, an agent can belong to multiple groups that represent a location, job function and/or application usage.

Once an agent has been installed, Performance Guard™ will automatically detect all servers and services that agents are accessing. To assist in the process of defining which servers and services to monitor in greater detail, Performance Guard™ will list all accessed servers and services ranked





by activity level. From this list, IT can select the appropriate servers and services to monitor. Once the known servers and services have been selected, IT can build the required groupings. Like agents, both servers and services can be assigned to more than one group, but by default they are grouped by location.

As an example, viewing the mean response times for a set

of servers and agent groups in a given time period and for a given service provides visibility into whether some groups of agents have better response time than others. If the groups of agents are geographically separated, there could be a network problem with one of the groups.

Real-time Data Output for High Visibility into IT Systems

Once the groupings have been defined, reports generated, and input entered into the repository, information can be extracted to optimize IT systems. Data can be generated (in any combination of agent, server, or service groupings) for a given time-period, and displayed using graphs that reflect information, such as:

o Response Time
o Account Histogram (%)
o Request per Second
o Active Agents
o Sent Bytes per Second
o Sent Packets per Second
o Received Bytes per Second
o Connections per Second
o Connections Reset (%)

By comparing all agents (and response times) against two applications, it becomes obvious where technical resources need to concentrate their efforts to correct the problem.

Custom Reports Display Collected Data for Real-Time or Historical Reporting

Custom reports are a collection of graphs that provide users with an overview of the service delivered by either a specific application or a number of applications. All graph types known to the Performance Guard™ display can be added to a report. The report is HTML-based and can be accessed via a standard Web browser. The administrator may customize the appearance of the report - font, background, color, etc.

Application Ping Feature Measures Application Availability

Even though Performance Guard™ is a passive end user monitoring tool, there are situations where the ability to turn the agent into an »active« agent may be beneficial (e.g. measuring application availability). By using the application »ping« feature it is possible to dictate when and how often an agent (or group of agents) actively »ping« a server for availability and response time.

There are 3 types of active pings to choose from. One is a »standard« ping testing the server, the second is an HTTP-ping testing a Web server if one is present on the server in question and the third is trace pings that uniquely enable tracking of the response from every single node in the network. It is unique to third generation solutions that agents can perform either as passive or active roles driven purely by the IT manager's needs.

Trace Ping is Critical

In many scenarios, understanding application, server and network performance is critical. However, networks often represent a mix of lines all being »black holes« to most IT managers. Thus, a trace ping enable pin-pointing network issues down to the individual node in a carrier network - this without having to ask anyone for logs or data. In the SAP example, the trace ping feature identify the path to the right resolution for a fraction the cost. These types of problems are often attacked with very expensive, yet non-effective, solutions - an expense that often deliver end user performance management returns in weeks or days, as opposed to months.

Ensuring Compliance with Service Level Agreements

A Service Level Agreement (SLA) report is a collection of SLA items, such as minimum response time criteria. An SLA item can be either an active item - a »ping« from the Performance Guard™ agent against a predefined application - or a passive item that is a »normal« passive response time measurement. The SLA item contains threshold values for response time and availability. If these thresholds are exceeded, the SLA item is considered failed.

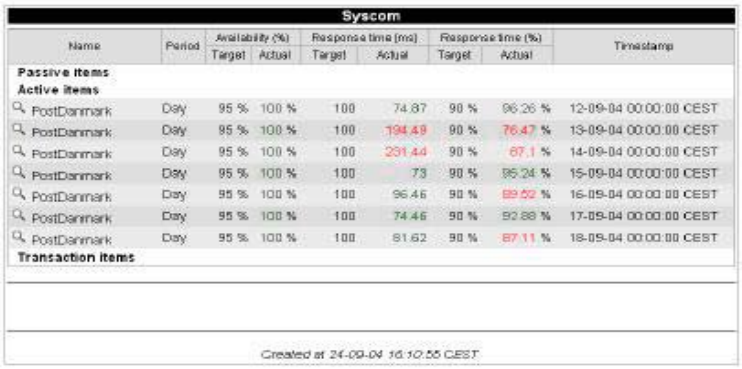
Alarms Sound When Response Time Slows

Alarms can be configured and defined as a point in time where the associated baseline alarm threshold has been exceeded. The alarm itself is defined on the Agent, Server, Service grouping and is sampled once every minute by the backend database. If the threshold is exceeded, the alarm is registered and its status is displayed on the initial Performance Guard™ Display page and/or can be configured to be sent as email. When an alarm comes from an end user it allows for rapid identification of the problem source and proper prioritization of the issue in the IT manager's responsibilities.

Advanced Environments such as Citrix Presentation Server and VMWare

Often, larger IT organizations take advantage of thin clients such as Citrix. Third-generation solutions like Performance Guard have added modules to uniquely understand the resource utilization not only on the PCs having the Citrix ICA client installed, but also the resource utilization on the Citrix

Server farm. Especially in Citrix or VMware installations, it is often very problematic to identify the source for performance issues. End user performance management solutions almost instantly measure the performance at each level of the installation and pinpoint the areas of constraint - whether it is the source application, the Citrix server, the network or the end user client. Until Performance Guard™, these scenarios often



Syscom

Name	Period	Availability (%)		Response time (ms)		Response time (%)		Timestamp
		Target	Actual	Target	Actual	Target	Actual	
Passive items								
Active items								
PostDanmark	Day	95 %	100 %	100	74.87	90 %	96.26 %	12-09-04 00:00:00 CEST
PostDanmark	Day	95 %	100 %	100	194.49	90 %	76.47 %	13-09-04 00:00:00 CEST
PostDanmark	Day	95 %	100 %	100	231.44	90 %	67.1 %	14-09-04 00:00:00 CEST
PostDanmark	Day	95 %	100 %	100	73	90 %	95.24 %	15-09-04 00:00:00 CEST
PostDanmark	Day	95 %	100 %	100	96.46	90 %	89.52 %	16-09-04 00:00:00 CEST
PostDanmark	Day	95 %	100 %	100	74.46	90 %	92.88 %	17-09-04 00:00:00 CEST
PostDanmark	Day	95 %	100 %	100	81.62	90 %	87.11 %	18-09-04 00:00:00 CEST
Transaction items								

Created at 24-09-04 16:10:55 CEST

presented IT with no way of gaining insight into the complete deployment.

End user performance Management in the Enterprise Management Framework

As said early on, end user performance management complements traditional data-center solutions. Thus, integration at some level is frequently required. The ability to send alerts with detailed information to enterprise management systems such as Microsoft Operations Manager, IBM Tivoli or HP OpenView is done using standards based SNMP formats. The alerts can contain detailed information and links to the appropriate reports thereby rapidly accelerating the problem resolution by data-center second- and third-level helpdesk staff. Just like it needs to fit into enterprise management frameworks, it needs to fit into desktop frameworks. Agent deployment is easily done through existing installation solutions. It comes either as an MSI package or as a self contained exe file that can be distributed through scripts, email or a url link.

The Performance Guard™ Advantage: Quick Problem Identification and Proactive Systems Management

The Performance Guard™ approach to application performance monitoring, including real-time measurements, alerts and reports, can instantly show trends and problems that are not yet detectable by monitoring the infrastructure. This provides the first ever CIO dashboard-level end-to-end view into IT performance. It supplements infrastructure management tools to show a complete picture of health of an application from the end user level and deliver a true productivity view of services received by the users. With this new view, IT departments can optimize system performance, increase their productivity and the productivity of their users, and reduce expenses on infra-

End User based Performance Management: See IT the way your users do. What you can learn will surprise you.

PremiTech
See IT the way your users do!

United States Headquarters:

PremiTech Inc.
33 Wood Ave South,
Suite 600
Iselin, NJ 08830
Phone: (732) 744 3419
Fax: (847) 330 6361

EMEA Headquarters:

Lautruphøj 1-3
DK-2750 Ballerup
Denmark

Phone: +45 702 772 08
Fax: +45 702 772 09

www.premitech.com

About PremiTech

PremiTech is the market leader in end-user based systems performance management. PremiTech's Performance Guard™ solution is the only solution of its kind to deliver real-time data on IT system performance based on the critical end-users' perspective. This actual performance information allows an organization's help-desk and IT departments to provide proactive service before problems occur, and to quickly determine the root cause of a problem when it does occur. In addition to these immediate benefits, Performance Guard delivers the information foundation for real-time capacity planning across servers, networks and end-user devices. Premitech's Performance Guard is installed at more than 2,000 locations in more than 100 countries, including customers such as HSBC, GE, CIBC, eTrade, Costco, TDC, NESA, and is recognized for its leadership by strategic partners like Citrix, IBM, Gotham, DynTek and others. Performance Guard was rated »Excellent« in a 2005 product review by InfoWorld. To learn more, visit www.premitech.com.