



A Collaborative Approach Accelerates the Transition from Silo Management to True Service Management

Restricted Rights Legend

The information contained in this document is confidential and subject to change without notice. No part of this document may be reproduced or disclosed to others without the prior permission of eG Innovations, Inc. eG Innovations, Inc. makes no warranty of any kind with regard to the software and documentation, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose.

Copyright

© Copyright 2007 eG Innovations. All rights reserved.

Introduction

While the worlds of voice, video, and data have become increasingly more integrated under the strategy of convergence, confusion continues in these spheres when discussing the concept of a "service." In the telephony domain, the concept of "service" is well defined. It generally refers to the always-on, utility grade delivery of dial tone for a telephone. In the same way, in the cable TV domain, the service is the programming available on-demand from a cable TV operator.

The concept of "service" in the IT domain is not as clear, or at least not as clearly spelled out. Most users and administrators tend to confuse "applications" with "services." While applications are individual software components, or "silos," like a web server, a database server, an application server, etc., a service refers to functionality that an end user receives. For example, online banking used to check account balances or pay bills is an IT service. Searching Amazon.com and buying a book online is another example of a service.



The Importance of Managing Services

"The loss of a few revenue-producing transactions over a limited time period due to poor web site response can be the least of a retailer's problems. An even bigger problem can be the blow dealt to brand equity."

Networks/Systems Management Newsletter, Network World 1/10/07

As many of the services in the IT domain grow in importance, there is a strong desire to achieve utility-grade reliability and performance through service level agreements (SLAs). Roughly half of all enterprise IT organizations monitor or measure server levels today, or expect to do so soon, according to market research and analysis firm Enterprise Management Associates.

Measuring service performance is one thing -- using the results to initiate prompt, corrective action is another. The challenge for IT organizations, then, is to maintain maximum end-to-end service availability and performance, and this goes well beyond simply measuring service performance. In today's multi-tier environments, a single service involves different silos, so failure in one silo will affect performance in other silos, ultimately compromising the quality of service end-to-end. For example, a failure or slow-down of a database server used in online banking can result in increased load on the application servers. This, in turn, could cause excessive connection queues on the web server, resulting in slow response to users and ultimately even complete denial of service.

It is apparent that measuring and managing service performance in any individual silo does not ensure proper end-to-end continuity.

Yet most enterprises today use silo-based tools for monitoring the performance of applications and services because that is all they know. This white paper, prepared by eG Innovations, Inc., a global provider of service management and IT infrastructure triage solutions, advocates that a new breed of service management software is needed. Such a solution enables a full enterprise-wide view of service availability and performance regardless of infrastructure size, configuration and topology.

Most organizations looking to make a shift from silo to service monitoring and management have typically initiated this paradigm shift from silo to service management in a top-down manner. In such initiatives, the directive to focus on services end-to-end is mostly driven by the C level executives (CEO/CIO), and the implementation is left to the individual silo administrators. Since most organizations are organized as silos, with specific roles and responsibilities for each administrator, an attempt to change the model in which the organization operates (to include cross-silo cooperation and coordination) often faces resistance from the different silo administrators.

Based on real world experiences with deploying the eG Enterprise Suite of service monitoring and management solutions, this white paper suggests that a new **collaborative** approach -- motivated by one or more groups of silo administrators -- has a better chance of adoption in enterprises. As one group begins to see benefits, other groups will begin to show interest and willingness to work with them, resulting in a collaborative approach to managing services delivered by the infrastructure.

In justifying these recommendations, this document examines the problems that IT organizations encounter when trying to apply silo-specific tools and processes to manage services that span multiple tiers of the enterprise infrastructure. The document discusses the features of the eG Enterprise Suite and how the company's Collaborative Enterprise Management™ (CEM) strategy enables IT organizations to successfully implement true end-to-end service management. Two real-world customer examples follow that show how the eG Suite has helped resolve service management issues through this collaborative approach.

Multi-Tier Infrastructures Complicate Management

When client/server computing was dominant, managing applications and services was relatively easy. When a user complained about performance, it was either due to a problem in the network or with the application the user was accessing. Troubleshooting was easy, since usually no more than two administrators were involved: the network administrator and the specific application administrator. The tools used for diagnosis were mostly specific to each application.

But as IT infrastructures evolved and newer application technologies emerged, it has become common to use multi-tier architectures to develop and deliver today's sophisticated IT services and applications as reliably as possible (**see Figure 1**). In such multi-tier architectures, a single service is provided using a variety of network and application technologies – e.g., a firewall for security, a load balancer to distribute processing across servers, a web server to act as a front-end, a middleware application server to implement the business logic, a database server to host

the data, etc. In this architecture, multiple applications and network devices have to interoperate in order to deliver the service to the end user.

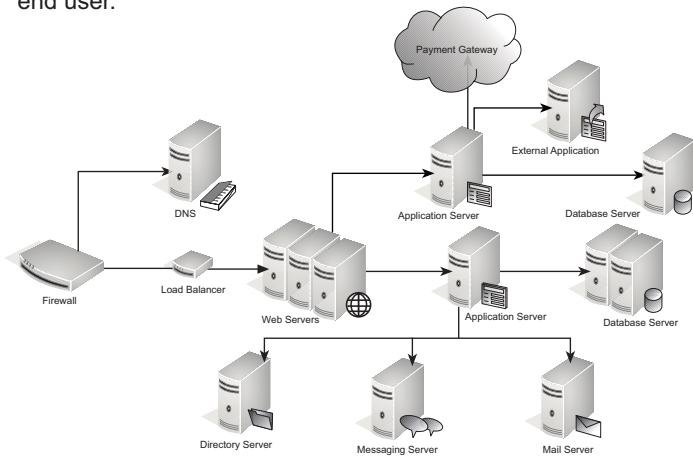


Figure 1 : Business services delivered through a multi-tier infrastructure with web servers, application servers, load balancers, databases, etc.

While multi-tier architectures have many benefits, such as scalability, they also present new challenges in monitoring and management. Since a service involves multiple servers, networks and applications, each of which is usually under a different domain (silos) of control, elaborate communication and coordination is necessary among IT teams to resolve any service performance issues that may arise.

The Role of Service Management

The term service management generally refers to managing service level agreements. Implemented jointly by service providers and end user organizations, SLAs define the degree of service availability and response time – or “quality” -- of a given service both parties agree to accept.

Ensuring that a service is performing up to SLA expectations is the task of service management. Any task that is performed to ensure that a service meets expectations is part of the service management process. This involves such tasks as:

- Routine monitoring of the service
- Diagnosis of problems
- Generating reports to determine if a service is meeting expectations or not
- Planning the capacity of the infrastructure to ensure that the services will perform well in future
- Tracking the configurations of the infrastructure components to determine that they are optimally configured.

The Challenges in Managing Services

There are many challenges in managing IT services:

- **Metrics are taken at the silo level, not at the service level.** One of the first challenges with true service management is defining the service levels appropriately. Most service level management initiatives end up focusing on metrics that relate to one of the silos. For example, a service level agreement can be made for network availability and latency. Another could be in place for web server availability. These are not examples of true service level guarantees, since they deal with individual

silos in isolation instead of how well they work together across multiple tiers.

- **Monitoring is silo-based.** To effectively monitor a service and measure its performance, it is essential to take the end user’s perspective, which is not concerned in the slightest with infrastructure topography or individual silos and domains. All that users care about is whether a service is functioning properly. In other words, can they log-in to online banking to pay bills, check account balances, etc. Also, does service response time meet end user expectations? These are all examples of end user views that span multiple domains. Yet monitoring, which is dependent on gathering and interpreting metrics, traditionally is a silo-specific function.

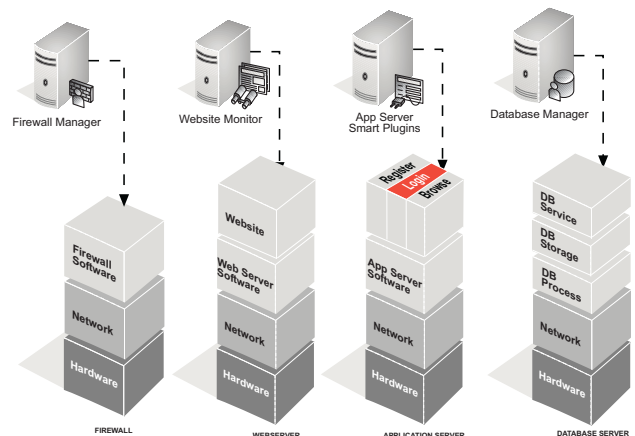


Figure 2 : Silo-based management involves using a different set of solutions for each tier/silo



Service Level Management is

the process of ensuring that IT services are properly linked to business processes and goals, and that there's performance criteria in place to support those goals.

SearchCIO.com, "Service Level Management: Linking IS performance to Business Goals", 9/25/2003

- **Independent views and toolsets are used across the different silos supporting a service.** Application silos typically are organized and managed independently (Figure 2). Therefore, performance is typically measured based on metrics specific to each silo, not taking into account which service(s) they support. Each silo usually has its own tool sets – e.g., database administrators use in-depth database analysis tools, network administrators have packet sniffers, probes, protocol analyzers, etc., while web administrators have log mining tools.

There is little or no coordination or integration across silo management tools. Further complicating the situation is the fact that many silos can be involved in supporting multiple services. For example, an online banking service and a credit

card processing service can use the same back-end database server. The larger an organization, the more silos there are, which intensifies the problem of coordinating across silos.

- **Lack of a coherent end-to-end view.** In most cases, the only people in an IT organization who have a service-oriented, multi-tier view of the infrastructure are the help desk personnel. Unfortunately, these folks do not have access to tools that can give them in-depth views into the real-time status of the silos that are responsible for providing the end-to-end service. At the same time, when a problem happens, the help desk's main task becomes one of coordinating across the different silos. Since there is no common view of the service across the entire organization, the various silo managers and help desk staff engages in finger-pointing and hit-and-miss troubleshooting. This process can last hours or days, during which time the problem and its impact grow more severe.

- **No common understanding of a service topology.** Surprisingly, in many organizations, even a common view of a service topology is lacking. The service topology is rarely documented, and the different silos are not even aware of the exact data flow and dependencies among them. This is one of the main reasons why automatic dependency discovery and mapping technologies have received attention recently. Lack of even basic knowledge of how a service functions means that the help desk and operations personnel spend hours trying to deduce where the root-cause of a problem lies.

The Failure of Top-Down Service Management Initiatives

Service management initiatives traditionally have been driven in a top-down manner. For example, a CIO or CTO puts together a team that is chartered to make sure that IT operations take a service-oriented perspective. Sometimes, the shift to service management is driven by a top-down desire to adopt emerging standards and processes such as ITIL, CoBit, etc.



N-Tier is the New Frontier of IT Operations Management

N-tier distributed infrastructures, whether based on Java Enterprise Edition, .NET, or full-blown Web services varieties, create unprecedented IT operations management challenges.

How does an IT operations staff stay on top of a service that has multiple, often dynamic interdependencies?

FTPOne Special Report: Operations Management, 2004

Experience indicates that such top-down approaches have had limited success. When the service-focused teams are made up from existing silo administrators, it is difficult to get the silo administrators to shift from the tools and perspectives they have used to manage their individual silos. As a result, even within the service-focused team, there is still considerable finger-pointing, rather than pin-pointing (Figures 3 and 4). Even when the service-focused team is assembled without involving silo administrators, there are problems. The service administrators do not have access to tools that can allow them to get a single integrated view of the infrastructure. Consequently, they still remain at the mercy of the silo administrators.

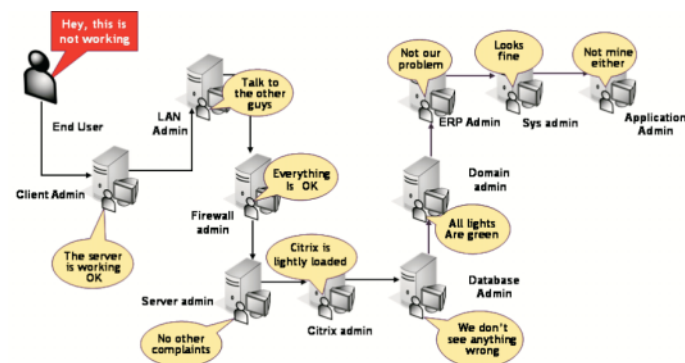


Figure 3 : Infrastructure triage is hard because of the use of different independent toolsets

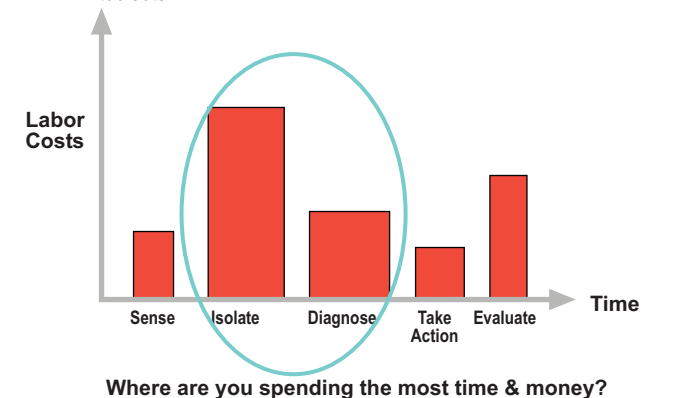


Figure 4 : An analysis of where time and money is being spent the most in the problem resolution process

A Collaborative Approach Accelerates Service Management Adoption

So how can effective service management become widely adopted in an organization? As surprising as it may sound, a collaborative approach could be the best way to shift IT operations from silo to service management. When a user detects a problem with a service, the common complaint always relates to the front-end facing application. For example, when a user sees slow response to a service being accessed through a Citrix Presentation Server, the complaint is that “Citrix is slow” or “Citrix is not responding.” (Figure 5) Likewise when a user has difficulty logging on to an online banking service, the complaint is that the “web site log-in failed.”

The real problem in each of these cases could be elsewhere, but the user complaint always relates to the front-end application because that's all they see. As a result, silo administrators of the

front-end facing applications (i.e., the Citrix administrators, the web administrators) want to prove that the problem is NOT theirs!



Figure 5: Users often blame the front-end applications for problems. Administrators of these silos have to defend themselves and prove that the problem is elsewhere in the infrastructure

But since they have little access to information about the performance of other silos, these administrators spend time trying to confirm that the problem is not in their silo but actually in one of the other silos. Of course, their colleagues in other silos also try to avoid direct responsibility. This “hot potato” scenario only intensifies the seriousness of the problem, and even after hours of finger-pointing, no one really knows the true root cause.

The challenge for the silo front-end administrators is in looking for solutions that go beyond silo monitoring and management. These administrators require tools that can take a holistic, end-to-end perspective, monitor performance across the tiers, and quickly determine in which tier of the infrastructure the root-cause of a problem lies. Once this is determined and proven, the silo expert responsible for this tier can be called in to resolve the problem.

Often, the front-end administrators do not have complete access and visibility into the other infrastructure tiers. Hence, the ideal service monitoring solution must be able to operate effectively in environments with limited visibility across tiers. Once adopted on a wider scale, the same monitoring solution must be capable of providing in-depth views into each tier of the infrastructure. The in-depth views allow for deeper correlation, more proactive problem alerting and faster diagnosis.

With an effective service monitoring and management solution in place, the front-end administrators can prove where the real problems lie. By doing so, they can reduce the finger-pointing and effectively provide quantified information that highlights where the infrastructure bottlenecks are.

Faced with hard evidence about the performance of their silos, the other administrators are now motivated to get involved in the troubleshooting process. Often, once this approach is proven to be correct and they start seeing the benefits of holistic, integrated monitoring, the different silo administrators begin to rely more and more on the service monitoring solution as a way to first determine which silo a problem relates to. In this way, the collaborative approach has a greater chance of adoption and long-term implementation than the typical top-down scenario.

Clearly, a service management solution has to be owned by one of the groups of the IT operations team. Since the silo administrators have roles and responsibilities within their silo, the task of monitoring and management across silos must ultimately lie with a separate



Barriers to Effective Performance Management

The three key barriers to effective communication in managing application performance are:

- **Organizational territoriality**
- **Lack of effective processes for working together**
- **Lack of effective tools to diagnose problems and support collaboration**

*Enterprise Management Associates webinar
"Managing Application Performance Over the Network:
An Assessment of IT Issues and Priorities" January 2007*

service monitoring team. Since most organizations do not have such a team in place, the help desk or support desk could be entrusted in the interim with the additional responsibility of overseeing the availability and performance of critical business services.

In any case, they already have this responsibility, albeit in a reactive manner (i.e., after customers complain). Rather than relying on the help desk for reactive support and therefore providing a very limited view into the performance of the IT infrastructure, IT organizations can enhance their efficiency by empowering the help desk to perform their current functions better. For example, if help desk personnel have a real-time view into the performance of each infrastructure silo, as well as the results of the root-cause diagnosis, they could more easily determine which silo is responsible for a problem and just call the administrator(s) of that silo to take steps to restore the operation of the critical services.

An Ideal Service Management Solution for Multi-Tier Infrastructures

For the collaborative approach to accelerate the service management process, it is critical that IT operations staff have access to the right monitoring and management solution for a multi-tier infrastructure. Key requirements of such solution include:

- **Measuring service performance end-to-end** in terms of availability and response times for different transactions, so that service performance degradations can be determined right away. Monitoring should be performed end-to-end, rather than relying on network pings, so that more accurate metrics of service performance can be obtained.
- **Providing an in-depth insight into the performance of specific tiers**, like the front-end applications (Citrix, web, terminal services, etc.) so that silo administrators of these applications are comfortable using the monitoring solution for day-to-day operations.

- Ability to provide insight into the performance of other tiers.** Administrators using the monitoring solution should be able to correlate service performance with the status of the individual tiers that are responsible for the service delivery. Since each of the tiers is probably under a different domain of control, the solution must be able to operate with strict security safeguards across all tiers. Monitoring across tiers may have to be done in an agentless manner because of the access restrictions across tiers. The extent of monitoring that can be achieved is also limited by the access controls imposed in the different tiers. For example, a server administrator may not have access to the statistics collected at the network layer. SNMP access will probably not be provided to the server or application administrator. Yet, the monitoring solution must be able to provide the server and application administrators with a view into the network performance, so they can understand where a problem may be occurring.
- Automating the analysis of the infrastructure.** To minimize the burden on the service management staff and to allow them to focus on the key bottleneck areas, the monitoring solution should automate routine tasks. Setting thresholds for each and every metric is a time consuming exercise. An ideal service management solution has pre-defined thresholds set to industry-standard best practices, providing a simple, out-of-the-box configuration. Furthermore, the monitoring solution should have the ability to automatically determine baselines for each metric, which ensures that administrators do not have to spend endless hours manually setting baselines. As a result, the monitoring solution can automatically detect anomalies and highlight just the areas of the infrastructure that need attention.



The Importance of Monitoring & Management

On average, 76% of firms' IT budgets go to ongoing operations and maintenance, as opposed to new investments.

Forrester Research "2005 Enterprise IT Outlook: Business Technographics North America"

- Ability to provide service-specific views.** By depicting the entire data flow of a business service, the ideal monitoring solution allows the service management staff to visualize the data flow between the infrastructure tiers supporting the service. The service management staff can better understand cause-effect relationships within an infrastructure and determine in which silo/s the possible cause of a problem exists.
- Abstracting the details of each silo.** The service management staff is unlikely to be experts at each of the silos. As far as possible, in order to enable them to assess performance across the tiers, the monitoring solution must use a common view to represent the status of each of the tiers. Doing so reduces the learning curve and makes it easier for service management staff to effectively manage the different tiers of the infrastructure.

The eG Collaborative Enterprise Management Suite: the Ideal Solution for Service Management and Infrastructure Triage

The eG Enterprise Suite from eG Innovations meets all of these requirements for end-to-end service management. Administrators have different options to measure service performance and availability. The measurements can be made by emulation of user activity and mouse clicks, or by observing a users working on their desktops. Metrics of service performance can also be collected from the front-end web servers and used to determine whether the service performance is within acceptable bounds.

The performance of the service can be time-correlated with metrics collected from all tiers of the infrastructure. In-depth real-time monitoring for Citrix, web, terminal services, etc. ensures that any problems with the front-end resources can be detected quickly. Custom models are available for most common enterprise applications, including J2EE middle tiers, messaging servers, active directory, and back-end databases like Oracle, Microsoft SQL, Sybase, DB2, etc. As a result, front-end performance and backend performance can be compared over the same time window.

Besides monitoring using software agents deployed on the servers, administrators can also monitor servers and applications in an agentless manner – i.e., without deploying agents. This capability can be useful in environments where administrators of one silo may have limited access to other silos or domains.

Using time-tested statistical quality control techniques, the eG Enterprise Suite automatically baselines each and every metric that it collects. The baselines are computed based on historical data, and take into account the time-varying behavior of metrics during a day.

To make it simple to triage problems across the different silos, the eG Enterprise Suite allows administrators to configure service topology graphs that capture the dependencies between the network, server, and applications involved in supporting each service (**Figure 6**). From the service topology representation, administrators can comprehend the cause-effect relationships among different infrastructure components, and quickly determine when a problem happens as well as which silo is responsible.

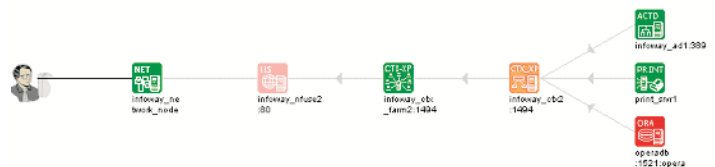


Figure 6: A service topology representing the applications and network devices supporting a service and the inter-dependencies between them. In this example, a user connects through a network router to a web front-end. In turn, the front-end allows the user access to Citrix servers that host enterprise applications. Active directory is used for user authentication, while an Oracle database hosts the data accessed by the applications hosted on the Citrix servers.

Finally, eG Enterprise uses a unique hierarchical model to represent the state of each network device, server or application it monitors (**Figure 7**). The use of a consistent representation for every infrastructure component shortens the learning curve for service managers. They do not have to monitor and manage each infrastructure component differently.

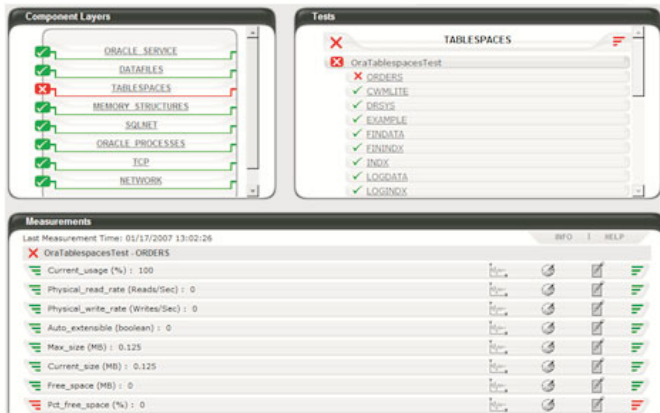


Figure 7: eG Enterprise uses a hierarchical model to represent the status of each infrastructure component being managed. The layer model clearly highlights which layer a problem pertains to.

Examples of the Collaborative Service Management Approach

Pin-pointing the Problem Silo

This section presents real-world examples of how the collaborative approach helped accelerate IT organizations to adopt a true service management capability. The first example involves a mission-critical telecommuting project. Employees of this organization were to begin working from home, using Citrix technologies for remote access to the corporate network to perform their day to day tasks.

Following the initial weeks of deploying this technology, the IT organization had been receiving complaints from users that Citrix was slow, and that Citrix would freeze all suddenly and they would not even see their keystrokes echoed back. The Citrix administrators could find nothing wrong with the Citrix server configurations or capacity. Since users were complaining about Citrix, the Citrix administrators of this silo were feeling pressure from management to do something. They even tried changing the server hardware, the client software, and the Citrix software, all to no effect.

The eG Enterprise solution was deployed in this infrastructure to see where the problems were. An analysis of the Citrix servers indicated nothing abnormal. CPU usage was within limits. Free memory on the servers was over 500 MB. There was no unusual session activity or network traffic on the servers, and the session traffic was well balanced across the servers.

Figure 8 illustrates the TCP retransmissions seen on the different servers of the infrastructure, and how eG Enterprise helped triage this infrastructure. From the Citrix servers, eG agents tracked the TCP traffic activity. One of the critical metrics monitored was the TCP packet retransmissions. In a typical network, TCP packet retransmissions should be low. In this case, however, eG Enterprise discovered and alerted about excessive TCP retransmissions.



Figure 8 : Diagnosis of a performance problem in a Citrix farm. The percentage of TCP retransmissions from different servers are compared over time. Servers xp1 and xp2 that clients were connecting to had excessive (over 20%) retransmissions indicating that a network problem was causing the retransmissions

By reviewing the data collected on network health, the Citrix team could immediately see that the problem was in the network silo, and they involved the network administrators to resolve the issue. Having seen the benefits of end-to-end monitoring from a service

Slowdown in Service Management Adoption

Analyst firm Forrester indicates that only 30% of companies aim to implement full business service management by 2010 and only 12% see themselves having achieved fully integrated management by this date.

Computerworld, Dec 4, 2006

The two Citrix servers in the farm were showing a very high percentage of TCP retransmissions.

Prior studies had indicated that acceptable response time and throughput for TCP traffic can be obtained as long as retransmissions remain below 5%. In this case, the retransmissions were close to 20% at peak times, which can have a severe impact on the performance seen by end users. This is because TCP uses an exponential window opening and closing mechanism. If the network is healthy, the window of data in transmission increases exponentially. However, if there is significant packet loss on the wire, the data window contracts exponentially. This, coupled with timeout mechanisms implemented in TCP, ensures that if a network is in poor health, response times and throughput for data traffic is poor. This evidence of poor network performance was especially surprising given that the servers were hosted in the same data center and high speed interconnects were used to connect users to these servers.

perspective, this organization has continued to use and see benefits from eG Enterprise.

This example has highlighted that, by using limited visibility into another silo, a service monitoring solution can help service managers or silo administrators quickly determine where the cause of a problem lies.

Creating a Service Desk for Effective Service Management

The second example is a Fortune 100 financial institution. This enterprise was using a variety of silo management tools, yet the performance of their web services were not meeting the service quality they wished to guarantee to their customers. As with any financial institution, every minute of slow down or down time costs thousands of dollars in lost revenue. The web front-end was often thought to be the bottleneck in the infrastructure.

Being a large enterprise with tens of thousands of servers, they had several silos for each domain of expertise (web, middle tier, databases, single sign-on, networking and load balancing, security, etc.). The presence of different silos, each with its own domains of responsibility, meant that there was little coordination across silos, and invariably, problem resolution entailed hours of bridge calls involving all the silo experts. The help desk had no access to the monitoring tools and was adding little value to the troubleshooting process.



Need for a Service Management Champion

For a service management initiative to succeed, appoint an “Organizational Effectiveness Coordinator”. His or her job is to help everyone improve processes without either taking away others’ authority or assuming others’ accountabilities.

*CIO Magazine, March 2005 -
Beneath the Buzz - “ITIL is a powerful tool, but holds pitfalls in store for those who get obsessed with it.”*

Part of the exercise with deploying eG Enterprise was to determine if the web tier was indeed a bottleneck. At the same time, separate help desk views were created in eG Enterprise to provide authorization for help desk personnel to view the status of the monitored infrastructure. The results of this exercise were dramatic (Figure 9). Since the help desk now had real-time access to the status of the critical services, they could detect problems before they affected customers. Further, by looking at the service topology maps, they could detect which silo was responsible for a problem and contact the right administrator.

The net results of this exercise were:

- A documented 30% reduction in time-to-fix for non-complex incidents – this translated into nearly 5,000 fewer customers affected by a problem;
- An order-of-magnitude improvement in time-to-isolate and fix complex incidents;
- Significant acceleration of the shift towards an end-to-end, service oriented management capability with fewer personnel needed on bridge calls to troubleshoot problems;

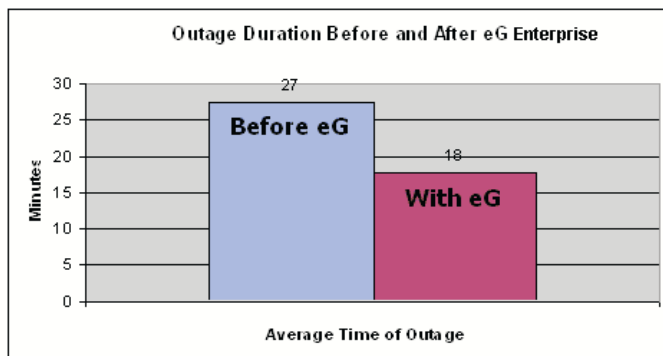


Figure 9: Benefits seen by one organization from deploying the eG Enterprise service monitoring solution in their n-tier infrastructure

Summary

The use of multi-tier architectures has increased the complexity of IT infrastructures, and no single administrator can claim to be an expert in all the technologies involved in delivering a service to the end user. As a result, most organizations have resorted to silo-based monitoring and management, with specialized tools being used for each silo. A consequence of the silo-oriented management approach is that no administrator has the service or the end-user view of the infrastructure. Consequently, problems go undetected, and even if a problem is detected, it takes a long time to determine what is the root cause of the problem.

This paper has discussed why top-down initiatives for service management tend to fail and why a collaborative approach, driven by one or more of the silo groups of an organization, can be a more effective way to accelerate the shift from silo-based to service-oriented monitoring and management. This conclusion is based on direct customer experiences using the eG Enterprise Suite.



SINGAPORE

eG Innovations Pte Ltd
33A Tanjong Pagar Road
Singapore 088456
Ph : (65) 6423 0928
Fax : (65) 6423 1744

USA

eG Innovations, Inc.
33, Wood Ave, South
Suite 600, Iselin
New Jersey 08830
USA
Ph: (866) 526 6700

INDIA

eG Innovations Pvt Ltd
2, Murali Street, Mahalingapuram
Chennai 600 034
India
Ph : (91) 44 2817 2801 / 2817 2799
Fax : (91) 44 28179041

Email : sales@eginnovations.com Web : www.eginnovations.com