



Solutions Brief:

Ensure the Availability and Performance of Your Critical Applications Using the Genius of Splunk

Background

Since web-based business applications are increasingly linked to revenue and productivity, enterprise IT departments feel an enormous pressure to develop applications quickly, maintain them at high levels of reliability and constantly monitor and improve the quality of service to end users.

Most web applications in today's enterprise IT infrastructures are SOA-based — composed of business application modules packaged into interoperable services that often interact with legacy backend systems. When virtualization is introduced into the environment, it decouples these applications from the hardware and makes monitoring even more complex. In such a dynamic environment, controlling the performance and availability of business services depends on having visibility into SOA components and the many layers of application software, operating systems, and virtualization and hardware infrastructures.

IBM WebSphere is a leading software platform for SOA environments that enables dynamic, interconnected business processes and delivers highly effective application infrastructures. These complex infrastructures span many organizational and datacenter silos, making detection of issues and the discovery of a root cause arduous, expensive and time-consuming.

Splunk software indexes data from any application, operating system, server or network device, enabling you to search and analyze billions of events across your IT infrastructure from one location in real time. Even with dynamic and complex WebSphere-based environments, Splunk allows almost instantaneous visibility into every element of the enterprise datacenter. This visibility lets you troubleshoot application problems quickly, fine-tune applications for optimal performance and availability, avoid service degradation or outages through proactive alerting and deliver cross-silo, end-to-end monitoring at a low cost.

Troubleshooting Today's Distributed Application Stack is Chaotic

A 2009 Gartner survey of CIOs indicated that linking IT strategies and plans to business priorities is a top concern for IT leaders¹. Customer-facing, web-based applications are increasingly becoming a competitive differentiator for businesses. Because these applications directly impact business results, IT organizations are focusing on delivering high service levels for both internal and external users. Yet there are several factors that make maintaining and improving service quality complex and difficult.

Many of these applications have evolved over the years. As a result they've developed into a jumbled and layered assortment of technologies. And now, organizations are moving to service-oriented architectures (SOA), which are based on sharing reusable business service components throughout an enterprise. This sharing of components between many different business segments makes it harder to assure service levels or ascertain the root cause of a problem. Transactions performed in a SOA infrastructure traverse many components, including application servers, databases, message buses and backend transaction processing systems. Virtualized environments, where applications are not assigned to specific hardware, present another dimension of complexity. Correlating server, storage, network and hypervisor issues against application behavior within virtual environments becomes fraught with difficulty.

Enterprise IT departments tasked with meeting and maintaining performance and availability service levels face the challenge of coordinating across many in-house silos and even external technical support organizations. Pinpointing problems and root cause analysis usually burns up thousands of man-hours in "war room" calls involving representatives from technology areas spread across the enterprise and beyond.

Macy's online retail application, comprised of infrastructure components including Apache web servers, F5 load balancers/firewalls, IBM WebSphere application servers and clustered Oracle databases, is a perfect example of this dynamic. The company websites macys.com and bloomingdales.com generate a significant portion of overall revenue. In the past, when incidents arose the support team had an exceedingly difficult time pinpointing the specific cause of the problem. Not only did the team involve representatives for each IT functional area, they had no way to troubleshoot from the source and no one team had visibility of the complete picture. Typically it took between 24 and 48 hours before they could determine where an issue occurred, and getting the logs to developers to find and fix the code consumed a few more hours. In general resolving problems "the old way" took the Macy's melded support team approximately multiple days.

¹ Gartner Research: "Align Enterprise Architecture to the Top 2009 CIO priorities" ID: G00169317

Lack of Illumination from Traditional, “Red Light/Green Light” Tools

Traditional monitoring tools for IBM WebSphere and other app server environments are point-products designed to address specific aspects of performance and availability management. Several of these tools identify and resolve specific component-level issues reasonably well, so why do IT departments still spend hours, days or even weeks performing these tasks?

- **Lack of cross-tier visibility:** Since traditional approaches to solving specific monitoring problems have resulted in the proliferation of narrowly focused tools, no single tool covers an entire infrastructure. This causes visibility gaps in heterogeneous environments where customers run IBM WebSphere application servers with other technologies such as Apache or Internet Information Servers (IIS) web servers, IBM WebSphere MQ Series message buses along with Oracle and MySQL databases, Cisco or Juniper firewalls and F5 load-balancers, VMware or Xen hypervisors. Packaged and homegrown monitoring tools that focus only on a portion of the environment often result in false green lights.
- **Limited flexibility and forensic capability:** Many homegrown and packaged monitoring tools are based on sampling, filtering and forwarding a subset of information provided by the monitored technology. For certain well-understood, well-foreseen situations, they perform reasonably well – but for a majority of difficult-to-predict scenarios, they fall short on troubleshooting and root-cause analysis. Since these tools usually don’t have the information needed for a deep drill down, they force the operations staff to manually scrape logs from each individual server or device to gain a deeper understanding of the problem. Plus their rigidly defined, schema-based approach necessitates months and years of customization that still leave them too inflexible to use on the fly for any type of reporting.
- **Expensive to maintain:** Not only do traditional approaches require multiple point products; many of these tools are expensive to purchase, deploy and maintain. When the environment changes, re-tooling is often as painful as it is costly. Further, the cost of these tools prohibits deploying them in development environments, leaving developers relying on non-standard tools for debugging. This reinforces a silo mentality, making it harder for developers to troubleshoot production problems.
- **Access restrictions:** Often developers can only pinpoint root cause or fix issues with distributed applications when given access to the production environment. Not only is this prohibited from a security standpoint but it also violates the separation of duties necessary for regulatory compliance. This forces IT operations to manually share information with development teams, which further increases MTTI/MTTR and the overall impact of the outage, especially in situations where every minute an application is unavailable costs the business ‘top line’ revenue.

Embrace the Complexity of Distributed Applications With a Novel Approach

Splunk is unique in that it can index, search, alert and report on all of the IT data across an entire application infrastructure – custom and packaged application logs, stack traces, message queues, database audit trails, even the logs, status and metrics for your OS, hypervisor and network.

Customers using Splunk to troubleshoot and manage packaged and custom applications routinely report cutting problem identification and analysis time by 60% to 80%. They spot and diagnose problems before they impact the business, eliminating up to 90% of escalations and avoiding millions of dollars in lost revenue and productivity.

Monitoring their distributed environment "the Splunk way", the Macy's support team detected abnormalities across their entire application environment as soon as they occurred. And they pinpointed the specific instance so they could resolve the issue before it resulted in an outage. Splunk helped Macy's eliminate downtime on both macys.com and bloomingdales.com through a record-breaking holiday season.

What makes Splunk unique for complex, distributed applications?

Massively scalable cross-tier visibility across the entire IT infrastructure

With Splunk, you can manage the full range of applications that make up the most complex IT environments. Splunk is universally compatible and can index the data generated by all the hardware and software components of SOA platforms such as IBM WebSphere, enterprise applications, web services, databases, Microsoft applications, J2EE Servers, middleware and operating systems without the need for version-specific connectors or pre-built integrations.

By importing data from your existing monitoring infrastructure, Splunk effectively augments it with rapid drill down and analysis capabilities, and the parallelism built into its search engine makes it particularly effective at scaling to handle terabytes and petabytes of data per day.

(For more on performance and scalability read the *Splunk and MapReduce* technical paper: <http://www.splunk.com/view/SP-CAAACGF>)

A leading, global financial services firm has deployed Splunk across their entire 1000+ application infrastructure.

"We just didn't have a way of aggregating message flows and examining them. Now we're monitoring message flow with Splunk as part of our trading application. We're monitoring for problems post-mortem and protecting against recurrences. In the future we'll probably use Splunk to monitor for performance assurance." — Team Leader, Research and Development

Deep visibility and flexibility

With Splunk, you can troubleshoot and pinpoint root cause in minutes or seconds instead of hours or days. Splunk's time-series indexing and powerful search language help you collate and correlate events across multiple infrastructure components. You can drill-down into the unfiltered data from any element and find root cause quickly. You can trace transactions across an application infrastructure in minutes or correlate hardware failure notifications with application performance or availability issues, massively slashing MTTR/MTTI.

By turning searches into pre-defined alerts, Splunk helps you proactively identify issues before you get a call from support. You can easily use Splunk to monitor your production environment for changes, letting you spot problems caused by configuration updates. The flexibility of Splunk lets you throw any problem at it; Splunk adapts to help resolve issues in minutes or even seconds. A customer responsible for application support at a Fortune 100 consumer goods giant uses Splunk to view their Apache web servers, shared Oracle WebLogic app servers, LDAP infrastructure and IBM WebSphere MQ software that support about 50 applications. Across this infrastructure, Splunk not only lets the applications team trace transactions across tiers, but also easily determine root causes when problems arise.

"Splunk is a completely different way of looking at things. It is extremely flexible and very handy when unexpected issues come up. I can build searches on the fly, search across my entire infrastructure in minutes and customize my views to show exactly what problem happened and how it manifested itself." — Mark Frost, Senior IT Specialist, Pepsi Beverages Company

Easy to maintain

The main advantage of Splunk lies in its ability to quickly search and report on any event, without any need for complex customization. Splunk indexes any and all machine-generated IT data, so changes or additions to SOA or hardware infrastructure components require little to no integration effort. Other tools require weeks or months of configuration, while Splunk usually delivers value in a matter of days.

Development and security friendly

Splunk manages all phases of multi-tier deployments, but it's particularly effective in SOA environments where developers may be involved in troubleshooting many different tiers deployed on various hardware infrastructures. Typically, when a problem is identified, developers get called in to troubleshoot. For security and compliance reasons, they usually don't have direct access to these servers so they ask someone in operations to zip and FTP the relevant log and trace files. With Splunk, developers have their own secure views of log data from live production servers. With Splunk deployed, developers who troubleshoot issues in production systems and SOA deployments now have direct, role-based access to the data they need to quickly find and resolve business-critical issues.

Splunk provides developers with much-needed visibility into application behavior without complex and laborious instrumentation. Splunk helps developers trace critical transactions across the stack, identifying slow, failed or resource-expensive transactions. Splunk also helps them monitor response times as well as error and status codes to understand the impact of changes on application performance. Splunk alerts developers when critical application metrics do not meet thresholds and application behavior deviates from the norm.

The Director of Software Architecture at Comcast believes, "Splunk is a required best practice. The ability to find and correct issues before we go into production more than pays for Splunk."

Customer Showcase

The following are examples of customers using Splunk to manage large-scale, distributed applications:



Summary: A Simple, More Flexible Approach Is Needed

Most web-based application environments, such as the ones based on IBM WebSphere, are intricate, multi-layered and constantly changing. The traditional way of managing performance and availability don't offer a holistic solution that deals with unexpected issues. These critical infrastructures need a new approach – one that's massively scalable, flexible and embraces both variety and complexity. Splunk is refreshingly innovative at resolving knotty issues around end-to-end monitoring and troubleshooting in these dynamic environments.

About Splunk

Splunk is software that provides unique visibility across your entire IT infrastructure from one place in real time. Only Splunk enables you to search, report, monitor and analyze all your real-time streaming and historical IT data. Now you can troubleshoot application problems and investigate security incidents in minutes instead of hours or days, monitor to avoid service degradation or outages, deliver compliance at lower cost and gain new business insights from your IT data. Over 1,850 enterprises, service providers and government organizations in 68 countries use Splunk to realize new levels of service quality, reduce IT operations costs and mitigate security risks in record time. The world's leading technology providers and over 50 system integrator, value-added reseller and managed service provider partners are driving new business and fueling their offerings with Splunk. Download your own free copy of Splunk today at www.splunk.com/download

For guidelines and examples on using Splunk in an application management environment, please download the product and review our documentation: <http://www.splunk.com/base/Documentation/AppManagement>