Sponsored Content

GCN

Mining Gold from Machine Data

During the past several years, agencies have embraced technologies that enhance productivity, mission readiness and cost efficiency. These new technologies have changed the way data centers are operated and managed. Today's data centers consist of complex, layered groups of siloed and interconnected technologies working in an environment without boundaries. At the same time, virtualization and cloud infrastructures introduce additional complexity, especially in the event of outage or performance degradation. When problems arise, these complexities make it very difficult to quickly identify and resolve issues, causing user dissatisfaction and dangerous periods of downtime that can impact agencies' missions.

The best way to gain understanding and visibility into your data center, whether on-premises, in the cloud or in a hybrid environment, is by gathering and analyzing machine data generated by the infrastructure software, operating systems and devices in the environment. This machine data often goes unmined and unused, but can power valuable Operational Intelligence and insight. Machine data includes application logs, IT router information, GPS and sensor data, transaction records, security data and more. By unlocking the intelligence in this data and understanding how all data interrelates, agencies can resolve problems faster, reduce downtime and improve user satisfaction.

FULL VISIBILITY

Holistic visibility is more important than ever because most IT services—including email and web services—consist of a series of connected applications and infrastructure components. Monitoring each component separately provides limited insight, but the ability to view and query all of the data connected to a service together provides exponentially more value.

Machine data contains a definitive record of all the activity and behavior of your customers, users, transactions, applications, servers, networks and mobile devices. Machine data includes logs, wire data, configurations, data from APIs, message queues, change events, the output of diagnostic commands, call detail records, sensor data from industrial systems and more.

By mining machine data, agencies can gain a comprehensive view of all data in their environment. End-to-end insights enable agencies to proactively monitor the entire environment to ensure uptime; rapidly pinpoint and resolve problems; identify infrastructure service relationships; establish baselines; and create analytics to report on SLAs or track SLAs of service providers.

“At the highest level, you might be able to see that there is a problem with a web service, but you may not know exactly where in the application stack or infrastructure that problem actually lives. But with the right tools and access to the right data, you can quickly search through the individual components to get to the root of the problem,” explains Jon Rooney, Director at Splunk, which provides a software platform for Operational Intelligence that is used by the majority of government agencies. Splunk software and cloud services enable organizations to search, monitor, analyze and visualize machine data generated from virtually any source, format or location—turning silos of data into integrated, actionable information and operational insights.

WITH DATA-DRIVEN OPERATIONS, AGENCIES ARE ABLE TO DRIVE MEANINGFUL ACTIONS AND DELIVER OPERATIONAL INTELLIGENCE.

Traditionally, it has also been difficult to gain full visibility into cloud-based infrastructure; in some cases, IT departments don’t even have root access to the backend machines running their applications and services. It can also be difficult to analyze and troubleshoot mobile applications. Machine data harnessed from cloud infrastructures and mobile applications can enable greater visibility to help find the root cause of problems faster, optimize infrastructure usage, improve operational efficiencies and deliver business value. With data-driven operations, agencies are able to drive meaningful actions and deliver Operational Intelligence.

The Nevada Department of Transportation has looked to the insights found in machine data to significantly improve visibility into both IT operations and security issues. Prior to adopting this approach, the IT staff used manual processes for lengthy system log reviews that were often unreliable. After implementing Splunk software, the staff built two dashboards to graphically present and manage logs for diagnosis and troubleshooting. One of the dashboards collects data from servers, switches, routers and firewalls throughout the network to inform IT staff of events like errors, time-outs and crashes. With this information, the staff can now perform root-cause analysis and fix problems much more quickly. The visibility enabled by Splunk software has helped reduce system errors from as many as 35,000 per day to just 2,500 per day.

FASTER PROBLEM-SOLVING

A nonfunctioning website or slower-than-molasses service can bring an agency’s mission to its knees. In a worst-case scenario, the time spent resolving issues can stretch from days to weeks or even longer. Users don’t care what the problem is—they just want a functional service, and it’s the IT staff’s job to make these hiccups as short-lived as possible. The best way for IT staff to identify and solve problems like downtime or slow connections is by having a holistic view of the data.

Traditional tactics required manually looking through logs of all types and finding patterns in a sea of data or homegrown approaches for log-data extraction. But if the data is already correlated and ready for analysis, it’s just a matter of asking questions of that data. For example, if a service went down at 12:53 p.m., an analyst can look across all of the data sources to see what happened at exactly that time. Very quickly,

Machine Data: The Critical Big Data

By this time, most know what big data is—varied data from both structured and unstructured data sources that is growing exponentially every year. The sources of big data include applications, email, social media, sensors, video, audio and text.

The fastest growing, most complex and valuable segment of big data is machine data. Machine data holds critical information on user behavior, security risks, capacity consumption, service levels, fraudulent activity, customer experience and much more. IDC predicts that 40 percent of all data will be machine-generated by 2020, up from approximately 11 percent in 2005.

Analyzing machine data requires a new breed of monitoring and analysis tools. With these tools, organizations can more quickly understand the root cause of problems, leading to faster fixes. These tools can also help detect security threats, improve compliance and fine-tune network management.

that analyst can pinpoint the issue to a specific configuration change in a database or error on a web server.

Access to relevant information enables agencies to accurately measure and analyze how software, devices and services are being used, who the heaviest users are, and how service levels are delivering against agency objectives. With that information, leaders can make more informed decisions to determine charge-backs and support audits, compliance mandates and strategic initiatives, such as data center optimization or tools consolidations.

IMPROVE OPERATIONAL EFFICIENCIES AND REDUCE COSTS

By understanding how systems and data interact, IT departments can get smarter about optimizing their operations and building efficiencies. Today, there is a siloed view into how data centers operate. For example, when planning for capacity in virtual environments, critical information that may be relevant to the operating system is usually missing. Without looking at the data from both the virtual environment and operating system, resource consumption is often under-reported, resulting in inaccurate resource allocation.

Analyzing machine data generated by these environments can help identify bottlenecks, optimize the allocation of resources and find underutilized software or hardware that often takes up valuable resources. These optimizations reduce the costs of providing IT services and enable agencies to operate with greater agility and efficiency.

“With the right visibility into your IT infrastructure, you can see where you are getting the most value and where you have issues,” Rooney says. “Are you rolling out new applications in a way that minimizes the impact to production systems? Do you have too much or too little capacity to service your customers?”

That was the case in Clackamas County, Oregon, where the county’s IT department implemented Splunk software to gain more visibility across different organizations and datasets. In addition to greater visibility, Splunk software helped the entire operation become more efficient. Holistic insights enabled the IT team to uncover patterns that resulted in implementing practices that saved money and rooted out waste. In one case, the IT team used Splunk software to respond to a query from the sustainability department about ways to save paper. When examining the data, the systems administrator noticed a large number of keywords related to coupons. With that information, the IT operations team was able to trace back to an employee misusing government resources and rapidly correct the situation.

CONCLUSION

Agencies today face increasingly complex IT environments, standalone technology silos and pressures to increase efficiency and reduce costs. With data from all tiers of applications and hardware infrastructure in a single central location, agencies benefit from faster troubleshooting and analysis. Correlating all machine data enables agencies to unlock insights from that data and gain Operational Intelligence needed to simplify processes, reduce costs, improve user satisfaction and increase efficiency and productivity.