305 Main Street
Redwood City, CA 94063

Toll-Free 1 855-LOG-SUMO +
Int'l 1 650 810 8700 +
www.sumologic.com +

sumologic

The Sumo Logic Solution:
IT Operations Management

Introduction

IT professionals face daily challenges in maintaining computing and communications infrastructure to support customer-facing and revenue-generating applications and services. Today's enterprise-scale IT infrastructures — a hybrid of physical, virtual and cloud environments — have become so complex that the number of possible points of failure has mushroomed.

A critical element in preventing and resolving issues within the IT infrastructure is the ability to quickly uncover root causes of problems precisely enough so that steps to resolution become immediately clear. Such information exists in operational log data, but the challenge is unearthing that information quickly from terabytes of logs.

"Agility and speed really matter when you're running production systems. With Sumo Logic, the most recent log data is available instantly so that you can investigate and analyze network, system and user behavior immediately."

As IT infrastructures have become more complex and dispersed, the volume of operational log data has become immense, making it very difficult to perform troubleshooting and root cause analysis when problems arise. Home-grown tools or aging on-premise log management systems were not architected to address complexities and scale of modern IT environments, and simply can't keep pace.

Sumo Logic's Log Management and Analytics Service

Sumo Logic's next-generation log management and analytics service changes this equation, delivering actionable insights into infrastructure operations while dramatically reducing complexity and cost.

As a massively scalable, multi-tenant service, Sumo Logic performs log data collection, processing, storage and analysis within a centralized and highly secure cloud-based platform. It effortlessly handles all of your log data, regardless of volume, type or location enabling IT teams to uncover operational insights buried under terabytes of logs in seconds. As a result, log data is emerging as one of the most strategic business assets within the enterprise.

Sumo Logic is designed from the ground up to handle Big Data-scale environments. Among Sumo Logic's breakthroughs is its near-zero latency Real-Time Forensics engine that delivers real-time search results from petabytes of log data. Real-Time Forensics makes critical new events occurring within IT infrastructure instantly available for analysis. Anomalous conditions can be spotted as they occur, enabling operations teams to respond immediately to prevent network outages, eliminate system downtime, resolve application issues and improve SLAs. In short, Sumo Logic reduces mean-time-to-investigation and mean-time-to-resolution dramatically.

Sumo Logic also scales to support orders of magnitude more data than legacy premise-based log management systems. Its patented Elastic Log Processing™ engine scales each component of the service independently to meet every customer's compute, storage and data processing requirements on demand.

Sumo Logic also takes a unique approach to log data collection. Data is securely and reliably collected through either local collection (via Sumo Logic Collectors) or through hosted collection (via https or directly from Amazon S3). All data is collected in raw, or unstructured format with no need to parse or understand the data upfront; all data processing and parsing is handled in the cloud. By separating collection from processing and parsing, which occur entirely in the Sumo Logic service, there is no need to update complex parsing logic. Consequently performance is significantly improved and management overhead significantly reduced.

To help enterprises manage exploding volumes of log data, Sumo Logic is built around a globally distributed data retention architecture featuring built-in data redundancy. Sumo Logic eliminates the need for costly SAN and NAS infrastructures and removes the complexity of data archiving, backups and restores.

Another major breakthrough is Sumo Logic's patent-pending LogReduce™ technology that eliminates the time-consuming and tedious tasks of manually reading log records, writing scripts and handcrafting queries. LogReduce reduces millions of log lines into a handful of human digestible patterns that enable IT teams to get to insights without having to manually writing queries to slice and dice the data. This enables our customers to quickly find important and emerging system and device behavior patterns that would otherwise require days of analysis. Sumo Logic's patent-pending Push Analytics™ leverages LogReduce to automatically uncover insights and it then pushes those insights proactively to IT teams in order to facilitate immediate investigations.

The Challenge of Identifying Operational Problems

In the earliest days of enterprise computing, the toughest problem to diagnose might have been figuring out which vacuum tube was faulty or which wire was improperly connected. But in today's highly complex IT infrastructures, there is almost an incalculable number of possible root causes of operational problems.

What do we mean by "operational problem?" Quite simply, it's anything that impedes or prevents business systems from performing their expected functions. Examples can range from the almost imperceptible (i.e., a sub-second increase in response time) to the catastrophic (i.e., a complete shutdown of a mission-critical servers). Most often, the nature and scale of a problem are somewhere in between. But the nature of complex and fully connected IT environments is such that problems seldom remain contained for very long. Rather, issues in one part of IT infrastructure almost instantly ripple through the rest. The result is that a minor glitch in a single router can wind up causing an enterprise-wide impact and business slowdown.

Small problems are always going to arise. What keeps them from impacting business operations is largely the speed with which they can be detected, analyzed and resolved. And the single biggest barrier to rapid detection is the complexity of IT infrastructure. When a problem occurs, the device will usually cry “help!” – delivered as an operational log line – but that cry is often drowned out by the noise generated by routine operations, expressed in terabytes of operational log entries.

Consider four areas of IT infrastructure where operational problems can arise:

- Networks
- Server
- Virtualization infrastructure
- Cloud computing infrastructure

Networks

A modern IT operation is made up of hundreds, even thousands, of hardware components involved in computing, storage, communications, and security. All of these components are interconnected by networking infrastructure such as routers, switches, firewalls, proxies, etc. A single router could be connecting dozens of servers and other devices running hundreds of applications. When a router experiences an issue, the ripple effect is felt across all servers and applications that are connected to that router.

Typically the first thing that happens after such a fault is that all devices and applications depending on that router start writing huge numbers of errors into their logs. This is followed by users calling IT or support complaining that their applications don’t work. One failure creates a firestorm of noise, making it harder to find the actual problem whose impact is obvious but the root cause is deeply hidden. Being able to extremely quickly isolate the root cause within a mountain of related error logs to the one that indicates that a particular router was misconfigured during latest maintenance is critical.

Servers

Sometimes it is the server’s fault - physical, virtual, or cloud-based. Servers run many applications that can clash with each other, require constant operating system upgrades to ensure stability and security, and interact with sometimes unpredictable users. Further complicating isolating server issues is the fact

> "Sumo Logic's intelligent parsing of structured and unstructured log data, coupled with the ability to run real-time queries, provide us with new insights into our platform's key performance trends and behaviors."
>
> Limelight NETWORKS

that servers are getting bigger and they are tasked to run more and more software on an ever expanding number of CPU cores, disks, and memory. An issue that arises within a server running critical business software like ERPs, Databases, HR systems can cause critical business processes to come to a screeching halt. Finding issues on servers is difficult, especially if those servers are not functioning well or if customers don't have the up to date server logs collected and managed in a different, always available location.

Even if you have the logs elsewhere, sifting through dozens of different server log types such as operating system, kernel, 3rd party libraries, applications, security software, is time consuming and frequently results in chasing false-positives.

Virtualized environments

The trend toward virtualization has been of great benefit to IT organizations, providing a high degree of flexibility and the opportunity to increase efficiency. At the same time, it introduces greater complexity to IT operations. Because virtualization introduces another infrastructure layer and makes it easy to move servers and applications between devices, it increases the chance of unforeseen faults within applications and operating systems. For example, the ease of provisioning virtual machines can result in too many virtual servers running on a particular piece of hardware, competing for what is now an inadequate amount of CPU, memory or network capacity. The result could be a failure or decrease in response times of critical applications that share the same virtualized hardware. In this example, looking for root causes within the given application or virtual machine will not yield results - the root cause is actually over provisioning and evidence is hidden within the logs of virtualization infrastructure itself.

Cloud Computing Infrastructure

Like virtualization, the increasing use of cloud computing brings many benefits. But when problems arise, the ability of an IT staff to identify the source of the problem can be compromised when its visibility into operations stops at the point of demarcation with the cloud. Cloud providers jealously guard the internal logs so all you have to analyze is logs of the nodes (servers) that are provisioned to you. Furthermore, ephemeral nature of nodes makes it more difficult to troubleshoot. If you are not collecting and aggregating log data in real-time from your nodes if those nodes fail due to your application or operating system failure, those logs are gone forever. There is no chance to learn and solve the issue to prevent future failures.

> "We now have the ability to quickly aggregate and summarize key findings from our log data. This informs us where performance improvements can be made in our web services infrastructure and provides early warning signs of potential issues that can be addressed proactively – before they become problems."

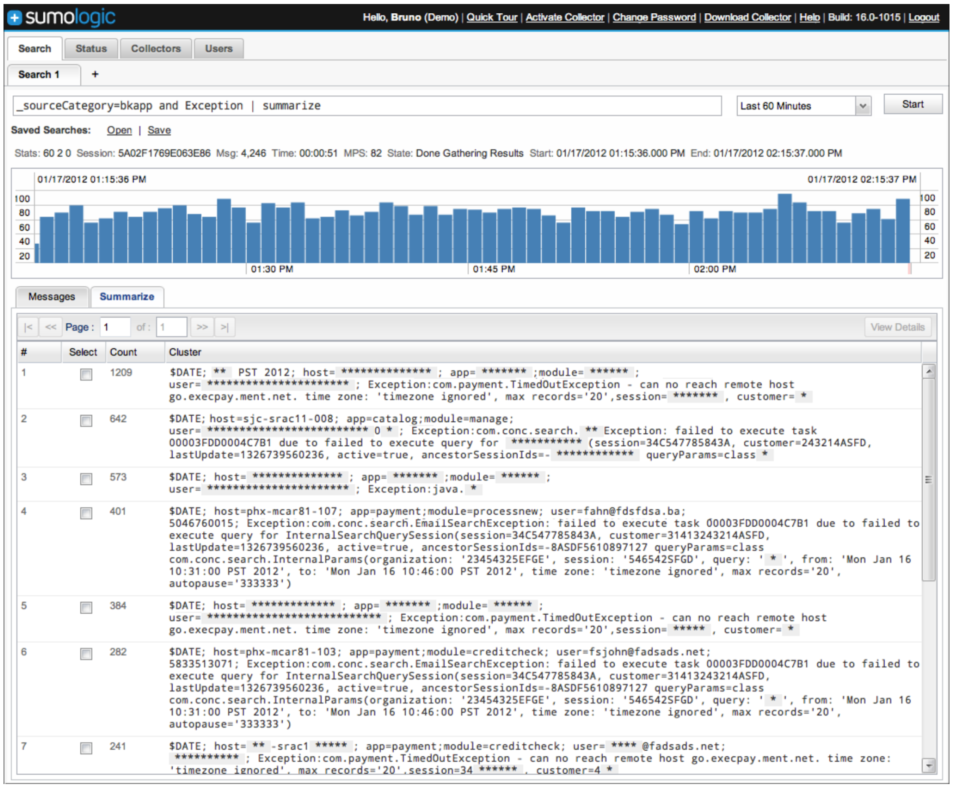
Limelight NETWORKS

Sumo Logic at Work

In the context of IT Operations, Sumo Logic plays a key role in ensuring smooth operations of modern IT infrastructure by dramatically reducing the time it takes to perform troubleshooting and root cause analysis. In addition, Sumo Logic delivers analytics capabilities that can help IT teams better understand utilization of their infrastructure, its performance characteristics, and deliver insights into which areas need further capital investments.

Sumo Logic's approach to log collection enables seamless collection in a non-intrusive way across your physical, virtual or cloud infrastructure. Our small footprint collectors can be deployed locally or remotely from the devices they are collecting from, or they can be baked right into your IaaS images such as Amazon Machine Images (AMIs). Data can also be collected with zero footprint through utilizing hosted collection (via https or directly from Amazon S3).

When something goes wrong within the IT infrastructure across data centers and/or in the cloud, conducting root cause analysis and troubleshooting in seconds is critical. The powerful combination of Sumo Logic Real-time Forensics and LogReduce™ enables IT Operations Teams, to with near-zero latency, isolate dozens of log lines reporting a router failure from



```
_sourceCategory=bkapp and Exception | summarize
```

Last 60 Minutes

Stats: 60 2 0 Session: 5A02F1769E063E86 Msg: 4,246 Time: 00:00:51 MPS: 82 State: Done Gathering Results Start: 01/17/2012 01:15:36.000 PM End: 01/17/2012 02:15:37.000 PM

#	Select	Count	Cluster
1	□	1209	$DATE; ** PST 2012; host= ************** ; app= ******* ;module= ****** ; user= ********************** ; Exception:com.payment.TimedOutException - can no reach remote host go.execpay.ment.net. time zone: 'timezone ignored', max records='20',session= ******* , customer= *
2	□	642	$DATE; host=sjc-srac11-008; app=catalog;module=manage; user= ************************* 0 * ; Exception:com.conc.search. ** Exception: failed to execute task 00003FDD0004C7B1 due to failed to execute query for *********** (session=34C547785843A, customer=243214ASFD, lastUpdate=1326739560236, active=true, ancestorSessionIds=- ************ queryParams=class *
3	□	573	$DATE; host= ************** ; app= ******* ;module= ****** ; user= ********************** ; Exception:java. *
4	□	401	$DATE; host=phx-mcar81-107; app=payment;module=processnew; user=fahn@fdsfdsa.ba; 5046760015; Exception:com.conc.search.EmailSearchException: failed to execute task 00003FDD0004C7B1 due to failed to execute query for InternalSearchQuerySession(session=34C547785843A, customer=31413243214ASFD, lastUpdate=1326739560236, active=true, ancestorSessionIds=-8ASDF5610897127 queryParams=class com.conc.search.InternalParams(organization: '23454325EFGE', session: '546542SFGD', query: ' * ', from: 'Mon Jan 16 10:31:00 PST 2012', to: 'Mon Jan 16 10:46:00 PST 2012', time zone: 'timezone ignored', max records='20', autopause='333333')
5	□	384	$DATE; host= ************* ; app= ******* ;module= ****** ; user= *************************** ; Exception:com.payment.TimedOutException - can no reach remote host go.execpay.ment.net. time zone: 'timezone ignored', max records='20',session= ***** , customer= *
6	□	282	$DATE; host=phx-mcar81-103; app=payment;module=creditcheck; user=fsjohn@fadsads.net; 5833513071; Exception:com.conc.search.EmailSearchException: failed to execute task 00003FDD0004C7B1 due to failed to execute query for InternalSearchQuerySession(session=34C547785843A, customer=31413243214ASFD, lastUpdate=1326739560236, active=true, ancestorSessionIds=-8ASDF5610897127 queryParams=class com.conc.search.InternalParams(organization: '23454325EFGE', session: '546542SFGD', query: ' * ', from: 'Mon Jan 16 10:31:00 PST 2012', to: 'Mon Jan 16 10:46:00 PST 2012', time zone: 'timezone ignored', max records='20', autopause='333333')
7	□	241	$DATE; host= ** -srac1 ***** ; app=payment;module=creditcheck; user= **** @fadsads.net; ********** ; Exception:com.payment.TimedOutException - can no reach remote host go.execpay.ment.net. time zone: 'timezone ignored', max records='20',session=34 ****** , customer=4 *

hundreds of thousands other log lines representing normal infrastructure operations. Doing this at near-zero latency requires the ability to collect, index and analyze the most up-to-date log data from all infrastructure components without compromise.

LogReduce applied machine learning algorithms look for new patterns and changes in patters to identify and bring forward those indicators of failure. This is done without the user needing to write queries, look for specific terms, or point the algorithms into a very narrow data set, such as a specific router. LogReduce is designed to put your own data to work and help you find things you've never seen before. Furthermore, Sumo Logic LogReduce does not only help find root causes of past failures, it also helps observe subtle changes in infrastructure behaviors over time, compare them to baselines, and predict and prevent future failures. These capabilities are critical now more than ever give the rapid adoption of new computing paradigms and growing complexity of modern IT infrastructure.

Sumo Logic Push Analytics™ runs analytics in the background and can notify you when important issues occur or behaviors detected in your log data change, facilitating proactive operational monitoring and reducing downtime and business impact.

Summary

Sumo Logic's log management and analytics service provides an answer to the challenge of quickly identifying the source of operational problems. The Sumo Logic solution is rooted in three key advantages versus other approaches:

- Sumo Logic collects all of the log data from IT infrastructure in real-time and stores it in a secure central location.
- Sumo Logic's advanced analytic tools quickly get to the source of a problem by reducing huge amounts of raw data into meaningful and digestible insights about anomalies and behaviors.
- Sumo Logic proactively alerts IT staff to the existence of operational issues that may indicate a problem in its early stages.