sumologic

Sumo Logic's Real-Time Forensics and Push Analytics™: Meeting the Challenge of Big Data Log Management

A Sumo Logic White Paper

Big Data *for* Real-Time IT

sumologic

605 Castro Street
Mountain View, CA 94041
P 1 650 625 0183 +

www.sumologic.com +

Executive Summary

The huge volume of log data generated by today's enterprises presents IT professionals with both a challenge and an opportunity. The challenge is managing large amounts of unstructured data. The opportunity is turning that data into actionable insights that can have a meaningful impact on the business or organization.

Sumo Logic's log management service employs several technologies that meet the challenge and capture the opportunity. Two of the most business-critical are Real-Time Forensics, which offers an unprecedented level of flexibility and responsiveness in pinpointing potential or emerging problems; and Push Analytics, which provides new insights into operational, security and compliance issues, including answers to questions that previously may not have even been asked.

"Sumo Logic ... is in a unique position to make observations that no person or in-house system could hope to make."

Real-Time Forensics: From "If only..." to "What if..."

Organizations that rely on systems and applications running 24/7 maintain a healthy paranoia about the potential for unplanned service interruptions, slow response times, security breaches and other problems that can affect revenue streams and/or customer satisfaction. The ability to fix problems quickly is essential. Avoiding problems in the first place is even more desirable.

Machine-generated logs—the diagnostic and runtime data produced by services, applications and –devices--contain information that can potentially hasten resolution of problems or prevent them altogether. But because the data sets produced by these logs are so large and unwieldy, many organizations don't even bother to try to understand what the data can offer. Those that do try expend considerable amounts of time, money and resources, and often don't see a very satisfactory return on these investments.

sumologic

605 Castro Street
Mountain View, CA 94041
P 1 650 625 0183 +

www.sumologic.com +

"Fixing problems quickly is essential. Avoiding problems in the first place is even more desirable." "Sumo Logic ... is in a unique position to make observations that no person or in-house system could hope to make."

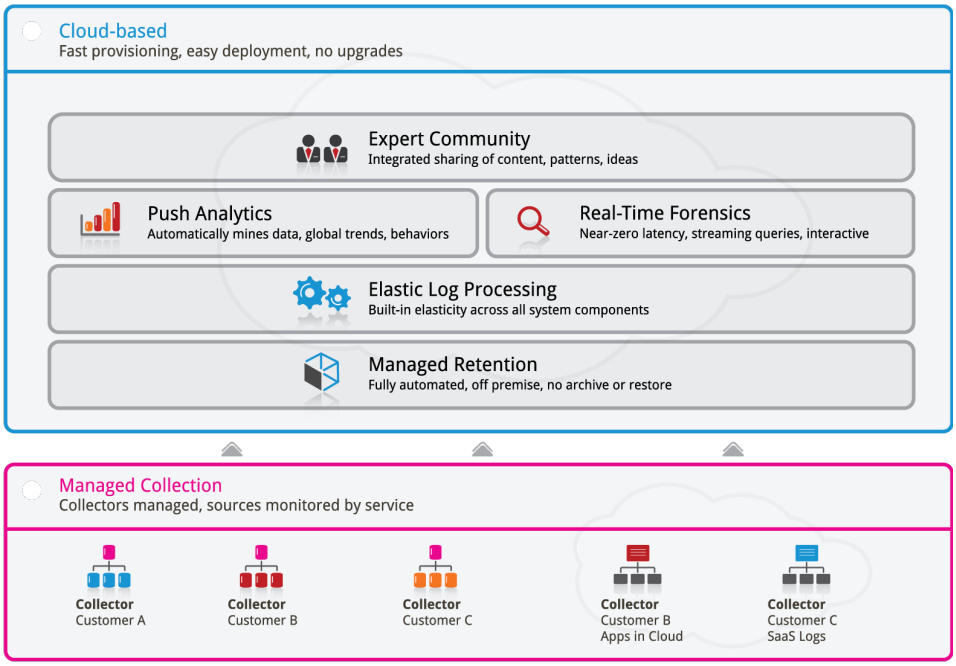
The result is that most of the time and effort spent on operational, security and compliance problem solving is backward-looking. It tries to understand where, when and how a problem occurred. Finding the answer, if in fact it is found, often takes hours or days, when an earlier diagnosis could have minimized its impact and prevented its recurrence.

Two of the biggest barriers to timely problem resolution are managing the volume of log data and knowing the right questions to ask. The Sumo Logic service resolves both these issues through a three-part approach:

+ Scalability. Because it is a cloud-based service, Sumo Logic does not impose limits on the amount of processing power a customer can apply to a problem. If a company suddenly faces a problem on a previously unmatched scale, it can add resources from Sumo Logic in as little as 15 minutes. If it were trying to tackle the same problem with an in-house approach, it would take days or weeks to add servers, software and personnel to increase processing capability.

+ Access to all data. Sumo Logic gathers and looks at all of a customer's log data. On-premise or in-house solutions have to limit the data set they are analyzing, because they have finite resources at their disposal. Sumo Logic can look at all a customer's data, increasing the chance that a diagnosis will be made quickly. In addition, Sumo Logic does not limit log data analysis to a particular set of equipment vendors. If a data set is not familiar, the Sumo Logic service will automatically mine the structure out of the logs and start building parsers without having to involve the customer. Sumo Logic then remembers the structure and applies it if another customer has logs from the same source, saving both time and resources and leading to faster, more accurate solutions.

+ Improved query options. The Sumo Logic combination of elastic processing power and a complete set of data means that more questions can be posed more quickly. Instead of following hunches and making educated guesses, IT teams can quickly scour massive amounts of data in search of the anomalies, error reports or patterns that will pinpoint the source of a problem.



Sumo Logic elastic log processing enables customer to scale their storage and data processing requirements on demand while providing applications, user and network trend through push analytics.

Sumo Logic's Real-Time Forensics capability changes and increases the business value of the IT department by giving IT managers immediate insights into key operational issues that demand immediate attention. Instead of always looking backwards, IT staff can ask more complex, "what if..." questions that will produce insights leading to operational improvements.

Consider the following examples where real-time operational insights can be critical:

An application service provider that provides service-level agreements (SLAs) with its customers suffers financially and in reputation if there is any degradation in the service it provides. If a customer reports a problem or the IT staff sees a drop in performance, it's critical that the source be identified quickly. But with 50 or 100 nodes involved, each generating streams of operational log data, an on-premise or in-house solution is reduced to essentially playing a guess game, making its best guess at the problem and then looking for supporting evidence in the log data. With Sumo Logic's service, all of the log data from all the servers is collected in one place, where forensics can be run very quickly. Starting with what a user sees at the front end, the problem can be traced back through different nodes in the cluster such as

databases, applications and/or servers. The result is a complete picture of that particular incident, what happened and why. And once that is known, the organization can take action to rectify the problem and prevent it from happening again.

A security example might involve IT managers noticing or suspecting an attack from an unauthorized outside party. With traditional responses, the process of identifying the specific target of the attack might involve a lengthy iterative process, depending on the number and location of servers. But with the Sumo Logic service, all logs from all servers are collected centrally, making it easy to reduce hundreds or thousands of logs to quickly determine exactly who tried to log in, when, and from what IP address. If, for example, a password cracker were being used, that would generate a log detailing a large number of failures from a particular host. Sumo Logic's Real-Time Forensics would detect the anomalous activity instantly and report it, allowing the IT staff to pinpoint the source and respond appropriately. Queries against that log could then be turned into alerts, where the staff would be notified in real time should the same, or similar attack occur.

"When looking at log management alternatives, including on-premise and managed solutions, we came to the conclusion that Sumo Logic is the right fit for us."

– Imperva

Another security example: managers suspect that a recently departed employee may have made unauthorized downloads during his/her last hours of employment. IT managers may be asked to identify all of that employee's computer activity for the last 12 to 24 hours of employment. But because the request comes after the employee left, the IT staff now has to go back through logs of past activity. Sumo Logic's service greatly simplifies the process by centralizing log data from all possible relevant sources –server logs, operating system logs, proxy logs, etc.—and filtering it by a single user name. The result: a full report on the employee's activity within an hour.

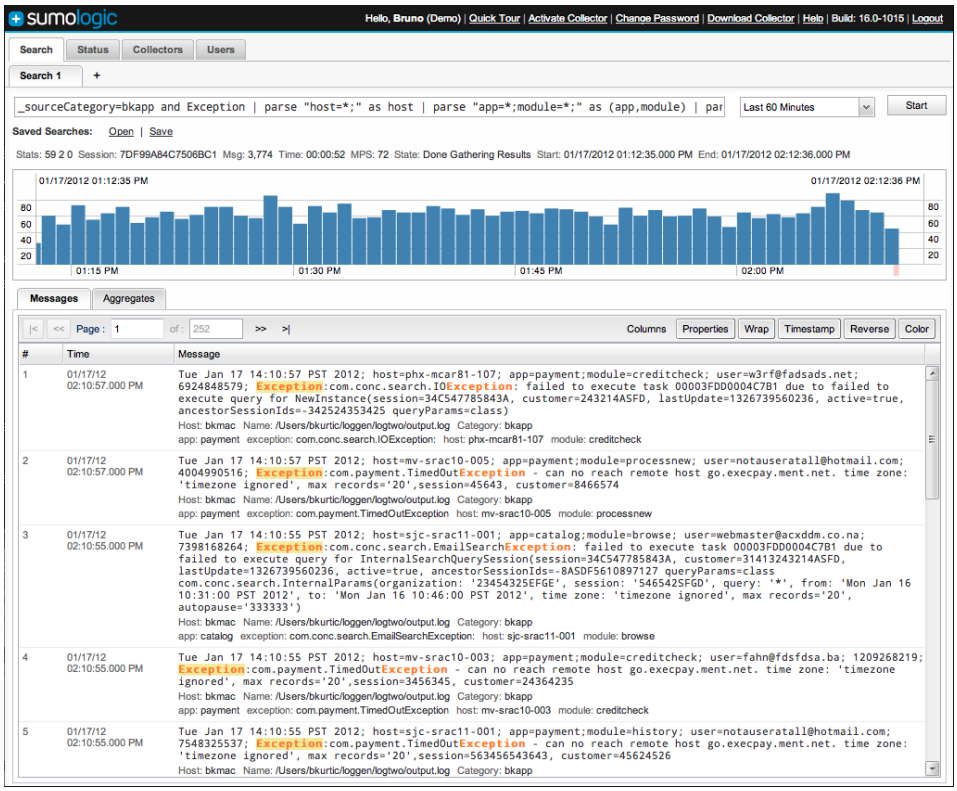
With Sumo Logic's Real-Time Forensics, the combination of processing power, completeness of data and flexibility of queries reduces the mean-time-to-investigation and mean-time-to-resolution.

Push Analytics™: Answering the Questions. Asked and Unasked

While Sumo Logic's unique forensic capability enables customers to process the huge volume of log data to detect and diagnose problems, it's equally important to be able to wade through that same volume of data and emerge with valuable operational insights. Sumo Logic's Push Analytics can extract interesting and important events from mountains of log data and proactively alert IT staff.

Push Analytics works in two dimensions. First, it analyzes all the log data collected, looking for anomalies and unfamiliar patterns--events that IT

managers might never think to make inquiries about. Then, out of all that data it will make a summary list of the most compelling and business-critical events and present those findings to an IT manager, who can select those items of greatest interest and drill down to investigate further.



Sumo Logic's near-zero latency Real-Time Forensics engine delivers real-time search results from terabytes of logs making critical new events occurring within IT infrastructure instantly available for analysis.

A second dimension to Push Analytics is that it learns from not only the experience of a single customer but learns from the experience of all its customers. That aggregated knowledge is applied in helping IT staff make sense of anomalies and previously unseen patterns.

An example of Push Analytics at work would be a business offering storage and processing services to large numbers of users. As a server runs out of disk space, it will generate a log entry somewhere that says "no space left on this device." Typically, there may be hundreds of such alerts from that server in a day. But they may not come to the attention of IT staff until a threshold is crossed. As more and more devices run out of disk space, performance will gradually degrade, and at some point customers will notice and complain. If,

"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

on the other hand, the IT staff had a report of the full-disk message on the first day, they would know to address the issue days in advance of when they might otherwise have noticed.

On the security front, consider the example of a server that normally gets a steady amount of traffic. At some point, a change is made in a firewall rule, and suddenly that particular port is exposed to a new service. An outside party who scans the network figures this out and starts hitting the service. This triggers a new type of alert that starts showing up in the logs. Sumo Logic's Push Analytics would alert the IT staff to this anomalous condition long before they might otherwise notice.

As important as it is to diagnose critical events as they occur, it can be equally important to know when something is not occurring. For example, a particular server may routinely generate an "I'm running" message every hour. But if that message hasn't been seen in the last six hours, that's important to know—and it could easily get buried amid all the operational log data generated every minute. An IT staff can write hundreds of monitoring rules to detect cases such as this, but still might be blind to dozens or hundreds of others that they don't know they could write. Sumo Logic's Push Analytics finds the answers to these unasked questions, both for what is happening that shouldn't be and what isn't happening that should be

Summary

Combining our highly scalable processing power, access to all log data, and highly flexible query capability—the Sumo Logic service provides IT managers invaluable assistance in maintaining uptime, meeting SLAs, preventing security breaches, and improving operating efficiency. Real-Time Forensics quickly gets to the root of problems, sometimes before they even become evident, while Push Analytics proactively notifies IT staff of trends and anomalies that otherwise could be invisible. Together they deliver unique capabilities and insights that improve customers' IT operations and bottom line.