



*Sponsored by LogRhythm*

**Scaling Analytics to Meet Real-Time  
Threats in Large Enterprises:  
A Deep Dive into LogRhythm's  
Security Analytics Platform**

*September 2013*

**A SANS Analyst Program Review**

*Written by Dave Shackelford*

**Review Environment** *PAGE 2*

**Security Analytics Architecture and Review** *PAGE 5*

**Use Cases** *PAGE 14*

**Searching and Reporting** *PAGE 21*

# Introduction

Today's attackers are getting smarter, attacks are stealthier than ever and the time to catch attackers before they do damage is shrinking dramatically, according to Verizon's Data Breach Investigations Report.<sup>1</sup>

Log management, event monitoring, and security information and event management (SIEM) platforms have helped thwart attacks in the past, but those tools are struggling to meet the loads produced by modern data centers and the need for quick response to advanced and persistent attacks.

SIEM platforms and similar tools are meant to help aggregate security and other data, correlate events and monitor activity within their environments. The next generation of security event management and monitoring tools is focused on analytics and very large data sets. The idea is that, by incorporating more data into their analysis, security teams can perform more predictive analysis for baseline development, as well as correlate increasingly diverse information that could potentially help discover needles in the haystack.

What most publications today fail to report is the impact these analytics may have on performance. With so many sources and varieties of information to collect, aggregate and sort through, a SIEM system's scalability must be assessed from multiple perspectives.

Until now, many teams have assessed their SIEM tool's capability to scale with a single metric—maximum collection rates based on events collected per second. Instead, they need to measure the volume, velocity and variety of data processed for real-time analysis, alerts and queries.

In this review of LogRhythm's real-time analytics capabilities, we evaluated its capacity for analyzing large quantities of data at high speed, emulating real-world environments and simulating several scenarios and use cases. We found LogRhythm was able to support large log volumes through a variety of processes and techniques that parallelize collection and processing, perform holistic real-time analysis and minimize search times while maximizing log storage capacity, thereby facilitating analysts' efficient review of past events.

---

<sup>1</sup> [http://verizonenterprise.com/resources/reports/rp\\_data-breach-investigations-report-2013\\_en\\_xg.pdf](http://verizonenterprise.com/resources/reports/rp_data-breach-investigations-report-2013_en_xg.pdf)

# Review Environment

For our review, we explored a security analytics testbed in a virtual environment provided by LogRhythm that was designed to validate specific performance metrics under very heavy load.

## Components

LogRhythm's architecture includes components for log collection, processing and persistence, real-time analytics and management. LogRhythm also offers optional agents that can collect log data from applications, independently generate host forensic details and perform file-integrity monitoring and other host-monitoring functions. The testbed was designed to stress the scale of agent deployment and log collection, using physical appliances to simulate a scalable analytics deployment and design that would support a large environment with large quantities of data to collect and analyze. The product deployment consisted of the following:

- **7 Log Manager appliances**—These aggregate log data, providing distributed and redundant log collection and management. Log Managers are horizontally scalable and include failover features.
- **2 Advanced Intelligence (AI) Engine appliances**—These provide sophisticated real-time correlation and analysis of all enterprise log data, covering logs, metadata, stored forensic data and other data types.
- **5,000 System Monitor agents**—These optional agents supplied logs and system event information to the centralized data collection platform. LogRhythm does not require the use of agents for collection, but many enterprise customers choose to utilize additional agent-based capabilities such as encryption, compression and the ability to spool data in the event of a network outage. We deployed 5,000 agents in our testbed as a demonstration of scalability for deployments where such additional capabilities are desired. Our testbed used the agents for Linux systems; agents for AIX, HP-UX, Solaris and Windows are also available.
- **1 Event Manager appliance**—This provides centralized event and alarm management and administration.

Because several messages can be associated with a single event, LogRhythm avoids the traditional SIEM performance metric of events per second, preferring messages per second (mps). This distinguishes the raw, text-based messages received by the system from events or activities recognized by it.

## Review Environment (CONTINUED)

The environment was configured with the following parameters and capacity considerations:

- 100,000 mps aggregate collection load
- 8.6 billion messages per day passing through the LogRhythm system
- 120,000 lab-generated unique log sources
- 1 log source, generating 20,000 mps
- 5,000 system monitor agents (with a capacity of 35,000 agents)
- 100 percent of data processed within the distributed architecture
- 100 percent of data analyzed in real-time by the AI Engines; no data was queued
- 100 percent of data available, persistent and archived
- 100 percent of data searchable

Although our testbed could handle up to 35,000 log agents, using this many was time-prohibitive for this review, so we deployed 5,000. These sent their messages to a single Log Manager appliance, demonstrating the capacity and availability of the architecture overall. (The other six Log Managers were included to showcase LogRhythm's distributed processing capabilities, as discussed under "Processing.") A syslog generator was also in place, as was a data generator that logged to a distinct flat file.

Figure 1 illustrates how the components of the testbed interacted and exchanged data.

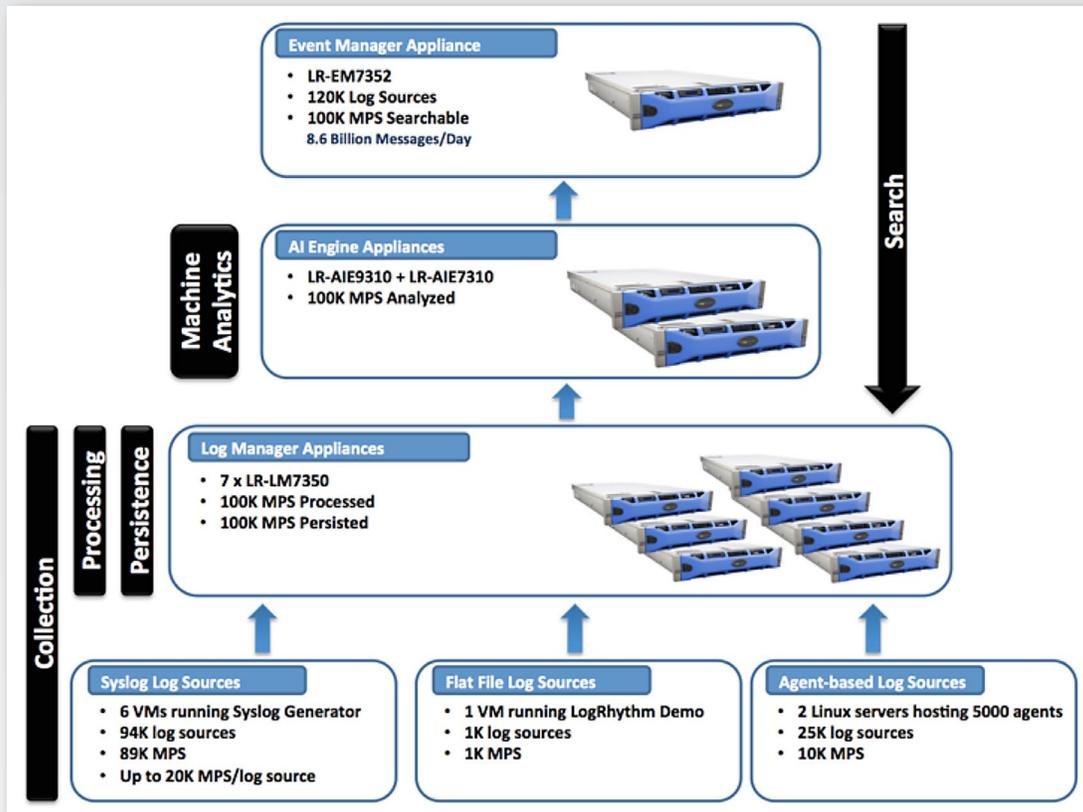


Figure 1. LogRhythm Review Environment

## Review Environment (CONTINUED)

When evaluating use under load, the evaluation environment included several “Knowledge Base Modules,” which are bundles of alerts, reports, rules and other content that provide expert interpretations of events and other data. Some areas these modules cover include:

- Automating PCI compliance
- Monitoring privileged users
- Detecting advanced persistent threats

In addition, we assessed a number of specific use cases, which are covered in more depth in the next section. These included machine analytics capabilities—which assessed unauthorized authentication activity and insider threat cases—security event search functions that delved into privileged user activity, permission provisioning activity and PCI compliance events, and, finally, archive recovery and data integrity for the data set overall.

# Security Analytics Architecture and Review

In a recent Network World blog post, Jon Oltsik described three key things that security analytics platforms should provide:<sup>2</sup>

- A complete picture of network behavior, which is critical for both understanding context and detecting anomalies
- Access to multiple sources of “security intelligence”; this can help add context to events and behavioral trends
- An understanding of “network state”—systems connected, configuration details, status changes and more

The LogRhythm architecture model addresses five distinct aspects that are essential to a “big data” evaluation of event and log information:

- Collection
- Processing
- Persistence
- Machine analytics
- Search

Our review examined the LogRhythm model’s approach to each of these, as well as several use cases in each category.

## Collection

The collection of event data is likely the most commonly measured aspect of event management platforms (and, by extension, security analytics systems). This is key to business and security teams alike, as poor performance of aggregation and central collection functions can degrade systems, applications and even network links.

The following attributes are important to consider when evaluating event collection:

- **Volume of data.** LogRhythm provides a horizontally scaled collection layer that enables the distribution of data collection and supports tens of thousands of data sources remotely or with agents. To validate this during the performance review, we observed the system collecting an aggregate of 100,000 mps, and that the system could scale and handle at least 5,000 concurrent agent connections on one Log Manager.
- **Velocity of data.** A single Log Manager appliance can sustain an average rate of 15,000–20,000 mps during collection. We confirmed that each Log Manager was receiving on average 15,000 mps, and our testbed included a high-volume log source producing 20,000 mps.

---

<sup>2</sup> [www.networkworld.com/community/blog/big-data-security-analytics-trifecta](http://www.networkworld.com/community/blog/big-data-security-analytics-trifecta)

- **Variety of sources and data types.** LogRhythm can collect any message-based data from any source, through a variety of remote or agent-based collection techniques. We saw that 120,000 unique log sources were generating events, with approximately 50 unique forms of log data represented.
- **Distributed architecture.** LogRhythm's distributed collection architecture can collect data close to the source, ensuring efficient and secure forwarding to a central location. (This function was not specifically examined during the review, although we collected data from a variety of devices in our test environment.)

## Volume and Velocity

To validate the collection capabilities of the LogRhythm analytics platform, we first reviewed the Log Rate Analysis report for the current review period. The "Logs per Second" metric, indicated in Figure 2 with a red arrow, shows a rate of more than 100,000 mps.

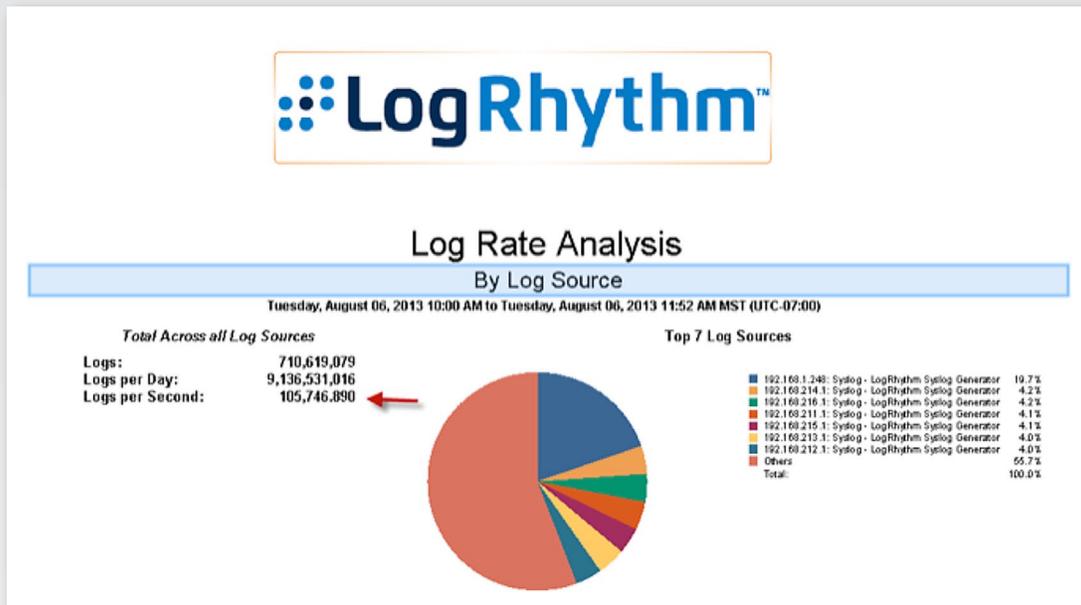


Figure 2. Aggregate Collection of More Than 100,000 mps

We then drilled down into the individual log sources to verify this rate and the generation of 20,000 mps by a single log source, as shown in Figure 3.

Summary				
<b>Total Log Sources:</b>		85,517		
<b>Total Log Messages:</b>		710,591,655		
<b>Avg Total Log Messages per Day :</b>		9,136,178,421	<b>Avg per Day per Log Source :</b> 106,835	
<b>Avg Total Log Messages per Sec:</b>		105,742.810	<b>Avg per Sec per Log Source :</b> 1.240	
Total Logs	Logs/d	Logs/s	Log Entity	Log Host
21,366,003	274,705,753	3,179.460	Sites	192.168.1.228
21,571,114	277,342,894	3,209.990	Sites/Site5	192.168.215.128
21,613,648	277,889,760	3,216.320	Sites/Site3	192.168.213.128
21,619,436	277,964,177	3,217.180	Sites/Site2	192.168.1.225
21,707,537	279,096,904	3,230.290	Sites/Site1	192.168.1.242
21,839,693	280,796,053	3,249.950	Sites	192.168.1.246
21,840,678	280,808,717	3,250.100	Sites	192.168.1.234
21,851,492	280,947,754	3,251.710	Sites/Site2	192.168.1.243
21,863,427	281,101,204	3,253.490	Sites/Site1	192.168.1.222
21,875,655	281,258,421	3,255.310	Sites	192.168.1.236
21,895,837	281,517,904	3,258.310	Sites/Site1	192.168.211.128
21,898,301	281,549,584	3,258.680	Sites	192.168.1.244
21,901,718	281,593,517	3,259.180	Sites/Site6	192.168.216.128
21,921,141	281,843,241	3,262.070	Sites	192.168.1.247
22,001,730	282,879,386	3,274.070	Sites	192.168.1.231
22,010,622	282,993,711	3,275.390	Sites/Site4	192.168.214.128
22,046,931	283,460,541	3,280.790	Sites	192.168.1.245
28,448,712	365,769,154	4,233.440	Corporate HQ	192.168.212.1
28,775,755	369,973,993	4,282.110	Corporate HQ	192.168.213.1
29,030,190	373,245,300	4,319.970	Corporate HQ	192.168.215.1
29,366,121	377,564,413	4,369.960	Corporate HQ	192.168.211.1
29,510,690	379,423,157	4,391.470	Corporate HQ	192.168.216.1
29,705,993	381,934,196	4,420.530	Corporate HQ	192.168.214.1
139,881,500	1,798,476,429	20,815.700	Firewall	192.168.1.248

*Figure 3. Details of Log Collection, Including 20,000 mps from One Source*

To verify the concurrent number of System Monitor agents, we checked the System Monitor tab of the LogRhythm console, which showed 5,000 agents reporting simultaneously, indicated in Figure 4 by the arrow.

Last Log Manager : USB01DEMOLM07 (5000 items)			
Action	Host Entity	Host	Last Heartbeat
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1	8/6/2013 11:33:13.843 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_10	8/6/2013 11:33:18.397 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_100	8/6/2013 11:34:05.073 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1000	8/6/2013 11:41:35.600 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1001	8/6/2013 11:41:36.177 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1002	8/6/2013 11:41:36.880 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1003	8/6/2013 11:41:37.490 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1004	8/6/2013 11:41:37.787 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1005	8/6/2013 11:41:38.207 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1006	8/6/2013 11:41:38.613 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1007	8/6/2013 11:41:38.987 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1008	8/6/2013 11:41:39.283 AM
<input type="checkbox"/>	Primary Site	A_MultiAgentTest_1009	8/6/2013 11:41:39.579 AM

*Figure 4. System Monitor Tab Showing 5,000 Agents in Use*

This can also be verified by reviewing the performance statistics shown in Figure 5.



Figure 5. Performance Statistics Displaying Agent Connections

To verify the collection capability on a single Log Manager, we checked the Performance Monitor. Note in Figure 6 that this Log Manager is collecting more than 20,600 mps.



Figure 6. Performance Monitor Showing Message Collection Rate

## Variety

To evaluate the variety of event data and sources for collection, we went to the Log Sources tab, which displayed the total number of sources represented, as shown in Figure 7.

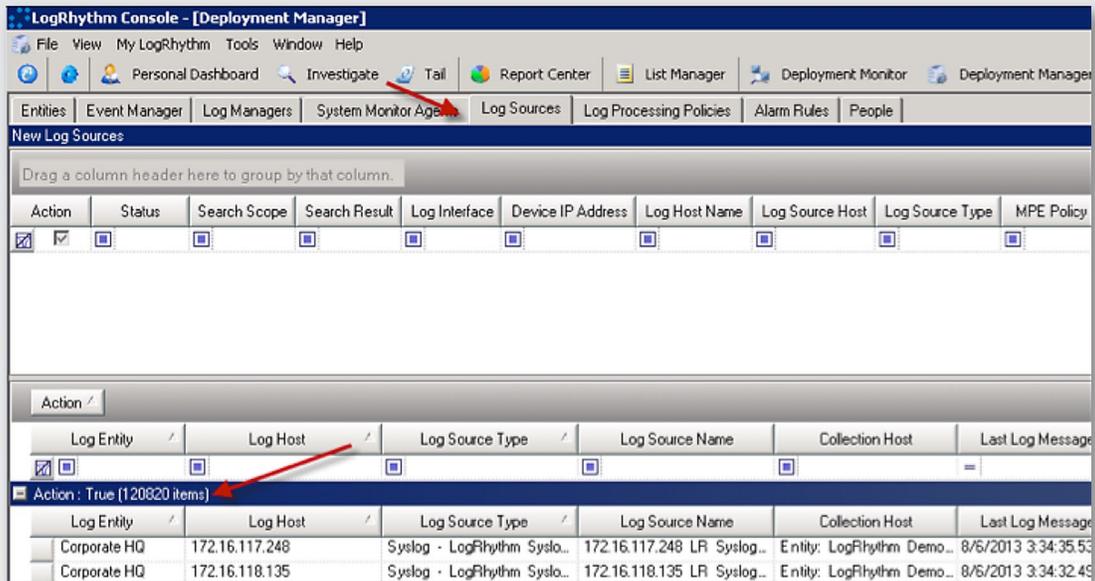


Figure 7. Log Sources Tab Displaying More Than 120,000 Lab-Generated Sources

## Processing

How well a security analytics platform can handle the data it receives is another important factor. When security teams are performing forensic analysis or real-time incident management, they need the most up-to-date data as soon as possible. The aggregation and sifting of log and event fields into a consistent format for analysis can be an extremely resource-intensive process that must execute as quickly as possible.

LogRhythm handles some of these process demands through load balancing. To evaluate the total processing capability of the review architecture and its capability to balance load, we checked the Deployment Monitor interface to see the activity supported by our seven Log Manager appliances, as shown in Figure 8.

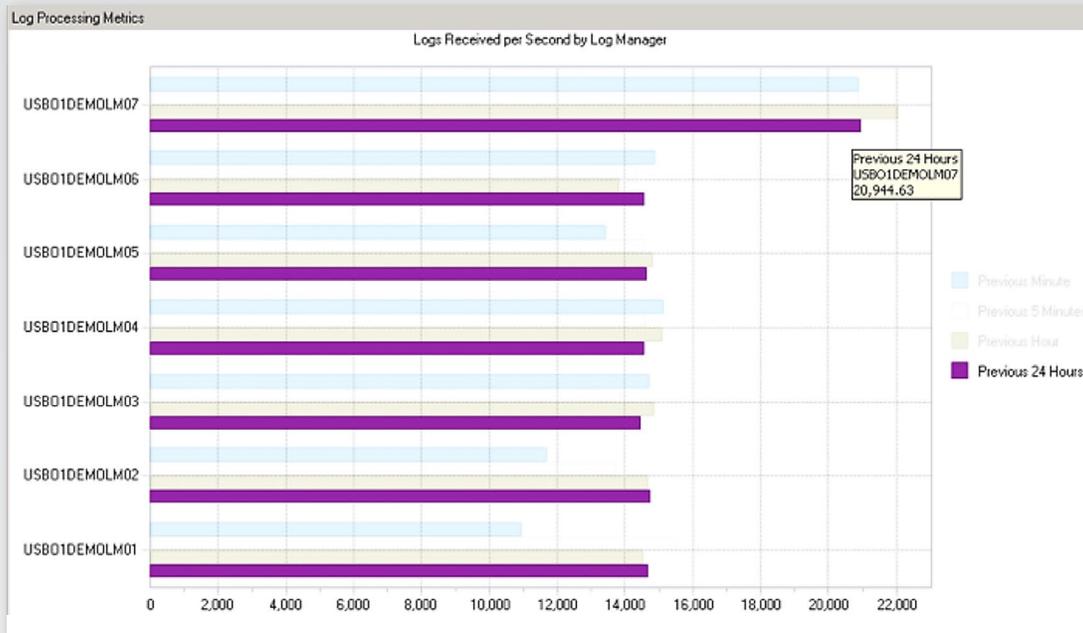
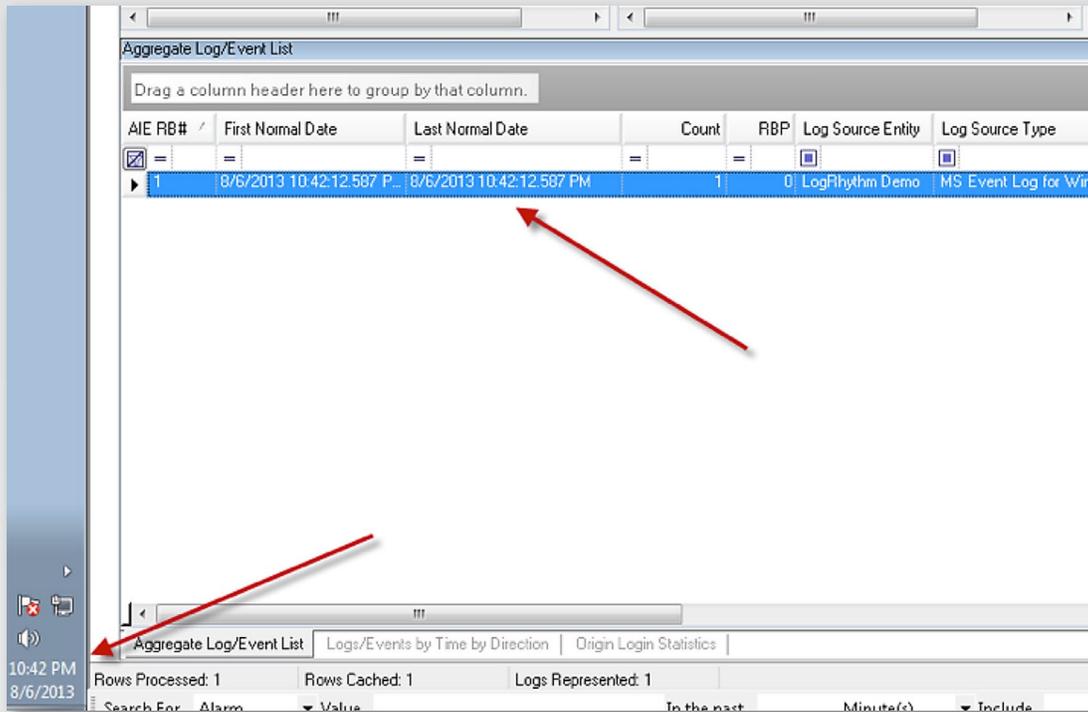


Figure 8. Deployment Monitor Displaying Load Across Log Managers

It was obvious which systems were being used in tandem, as well as the total processing metrics for each platform. Figure 8 shows one appliance processing more than 20,000 mps, and the other six with loads distributed equally.

To see how rapidly LogRhythm could process data and make it available for searching and analysis, we reviewed activity monitors to see when specific events were first seen by the system and then verified when they were processed. An example of this appears in Figure 9, with the system collecting an event at 10:42 p.m. and the event being processed and made available within that same minute.



The screenshot displays the 'Aggregate Log/Event List' window. The table below shows a single event entry where the 'First Normal Date' and 'Last Normal Date' are identical, both being 10:42:12.587 PM on 8/6/2013. A red arrow points from the system clock in the bottom-left corner to the 'Last Normal Date' column, and another red arrow points from the system clock to the 'First Normal Date' column, highlighting the minimal time difference between collection and processing.

A/E RB#	First Normal Date	Last Normal Date	Count	RBP	Log Source Entity	Log Source Type
1	8/6/2013 10:42:12.587 P...	8/6/2013 10:42:12.587 PM	1	0	LogRhythm Demo	MS Event Log for Win

System Clock: 10:42 PM 8/6/2013

Rows Processed: 1 | Rows Cached: 1 | Logs Represented: 1

Figure 9. Comparison of Activity Timestamp with System Time

## Persistence

In large, busy environments, the ability to continually collect, store and process event data for future use is also important. In addition to volume, velocity and variety of collection types, the Log Manager (LM) appliance is designed to sustain concurrent read/write operations, so heavy write activity doesn't degrade the ability to search through large volumes of variety of data types. We observed rapid access to data via search functions, even under a sustained load of 15,000 mps per Log Manager.

To determine the volume and velocity of event data persisting on the system, we reviewed the various Log Manager metrics, including log processing, archiving logs to disk and logs being "flushed" through the system. The metrics for all seven LM appliances appear in Figure 10.

Appliance	Metric	Value
LM211	LogRhythm AI Engine Data Provider Rate Logs Flushed / Sec	14,452.107
	LogRhythm Mediator:Processing Logs Archived to Disk / Sec	14,392.568
	Rate Logs Processed / Sec	14,397.776
LM212	LogRhythm AI Engine Data Provider Rate Logs Flushed / Sec	14,248.196
	LogRhythm Mediator:Processing Logs Archived to Disk / Sec	14,345.885
	Rate Logs Processed / Sec	14,336.271
LM213	LogRhythm AI Engine Data Provider Rate Logs Flushed / Sec	14,334.375
	LogRhythm Mediator:Processing Logs Archived to Disk / Sec	14,327.897
	Rate Logs Processed / Sec	14,351.854
LM214	LogRhythm AI Engine Data Provider Rate Logs Flushed / Sec	14,421.584
	LogRhythm Mediator:Processing Logs Archived to Disk / Sec	14,363.512
	Rate Logs Processed / Sec	14,377.612
LM215	LogRhythm AI Engine Data Provider Rate Logs Flushed / Sec	14,385.118
	LogRhythm Mediator:Processing Logs Archived to Disk / Sec	14,420.033
	Rate Logs Processed / Sec	14,403.059
LM216	LogRhythm AI Engine Data Provider Rate Logs Flushed / Sec	13,972.898
	LogRhythm Mediator:Processing Logs Archived to Disk / Sec	13,937.121
	Rate Logs Processed / Sec	13,941.275
LM217	LogRhythm AI Engine Data Provider Rate Logs Flushed / Sec	20,808.110
	LogRhythm Mediator:Processing Logs Archived to Disk / Sec	20,722.869
	Rate Logs Processed / Sec	20,718.917

Figure 10. Processing of Data Across Log Managers

## Protecting Stored Data

To protect the integrity of this data being collected, LogRhythm cryptographically hashes data when written to disk and verifies the hash during read operations. We ran searches against data with integrity checking in place, and there was no discernable degradation of system performance. To facilitate longer-term storage, LogRhythm archives older data in a highly compressed format.

We reviewed the integrity check tools and capabilities of LogRhythm's SecondLook Wizard, which allows security analysts to rapidly restore archived logs and data into the platform for forensic review. Figure 11 illustrates of the selections necessary to receive a warning upon detection of possibly compromised data.

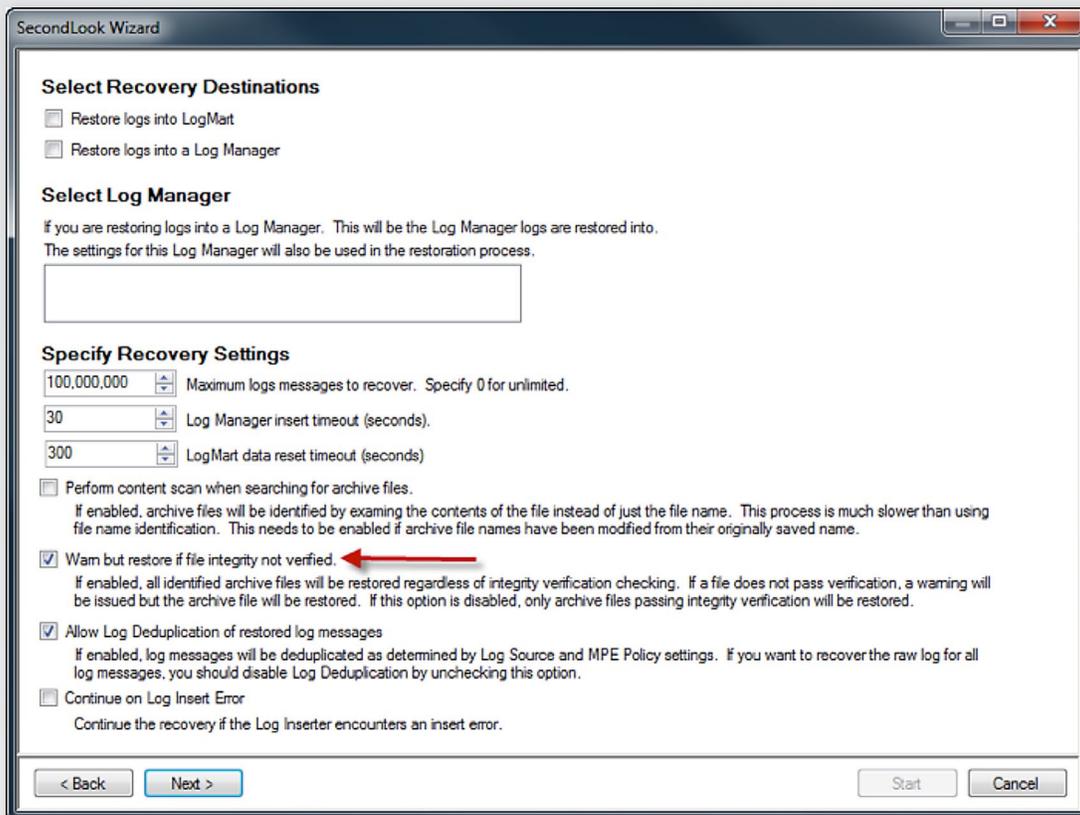


Figure 11. Integrity Check on Archive Retrieval

### Machine Analytics

The heart of any analytics platform is the complex correlation and analysis of data that has previously been collected and processed. Any event management platform that performs complex analytics should exhibit or possess the following capabilities:

- Each AI Engine can analyze data at very high velocity, and analysis can be distributed across multiple AI Engines in support of horizontal scaling or distributed analysis. We observed that 100 percent of the collected data was processed by the two AI Engine servers in the review deployment.
- To meet the need for real-time throughput, the AI Engine leverages stream-based, in-memory analysis to recognize activity within seconds of collection. We verified that alarms raised by the AI Engine were visible within seconds of the associated data being collected.

Monitoring for sophisticated threats and modeling complex behaviors also requires a holistic view of the environment. The AI Engine's analytics rules can target the spectrum of collected data across all device types.

We noted that each AI Engine server seems to be very flexible in parsing and interpreting data, enabling broad analysis capabilities. Its data engine supports a variety of analytic techniques from correlation to multidimensional behavioral analysis. Several real-world use cases that demonstrate the analytics capabilities follow.

# Use Cases

Our review saw several advanced analytics use cases in action, measuring the system’s capability not only to consume data, but also to perform real-world threat detection and response. All use cases were examined while the system was under continuous load (as described earlier, in “Review Environment”) to demonstrate usability. We found the system was able to accomplish its tasks and provide a responsive experience for the user while handling billions of messages each day. The first use case demonstrates the use of an AI Engine Rule that monitors authentication success activities from accounts that are not approved PCI users.

## Detecting and Blocking Suspicious Users

For example, the “log observed” rule block uses LogRhythm’s log classification for any authentication success activity against any user not listed among either the software’s administrative users or the PCI users (a list created for this testbed). This demonstrates the system’s capability to learn and then apply that in an adaptive way to dynamic environments through system-generated watch lists. Figure 12 shows the rule as it appears in the AI Engine Rule Wizard, which enables analysts to construct advanced correlation rules.

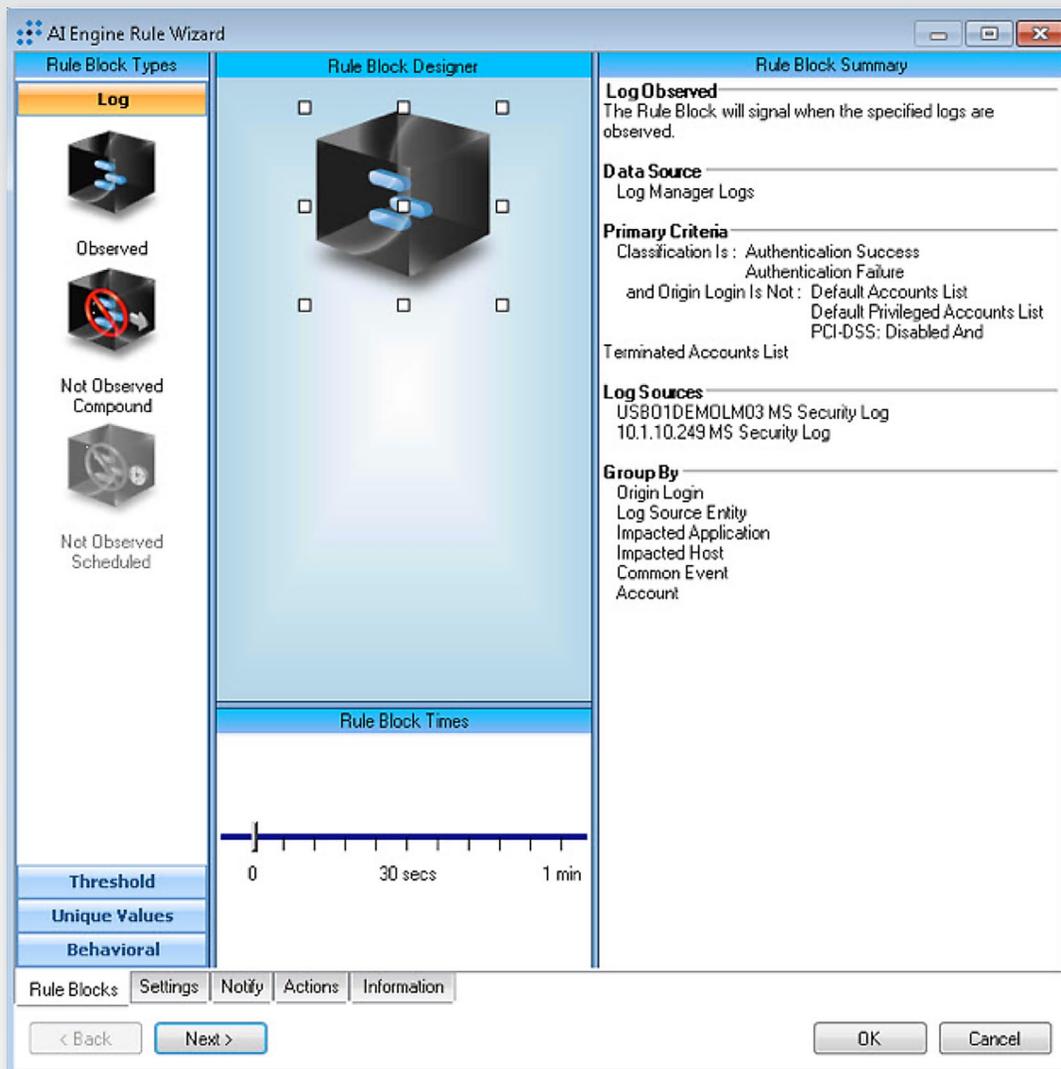


Figure 12. AI Rule for Log Classification

## Use Cases (CONTINUED)

Next, Figure 13 shows the specifics of the rule, where we have set up a SmartResponse action to automatically add any user recognized in the preceding step to a “Suspicious Users” list.

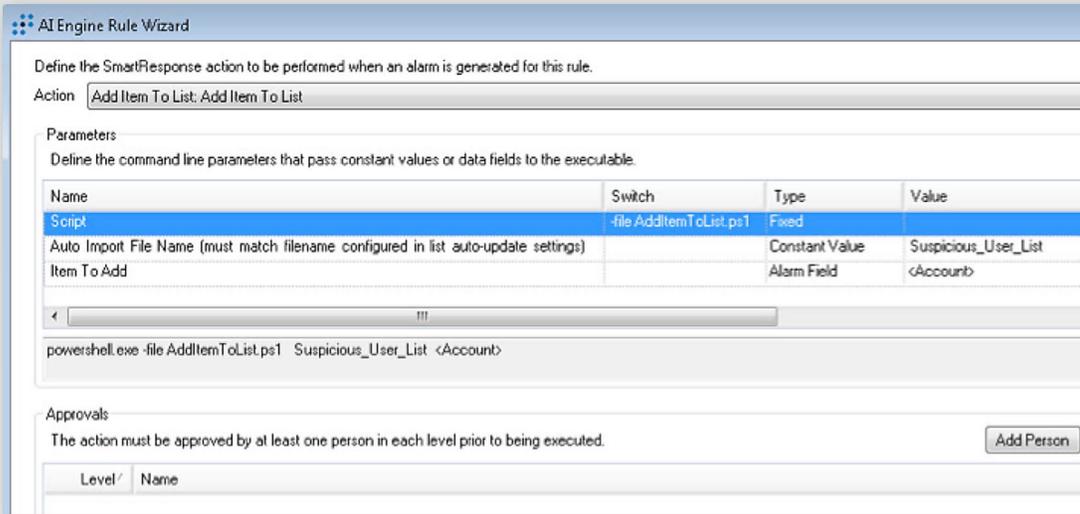


Figure 13. SmartResponse Actions in AI Rules

In Figure 14, the rule triggers and executes the SmartResponse, adding the user to the “Suspicious Users” list.

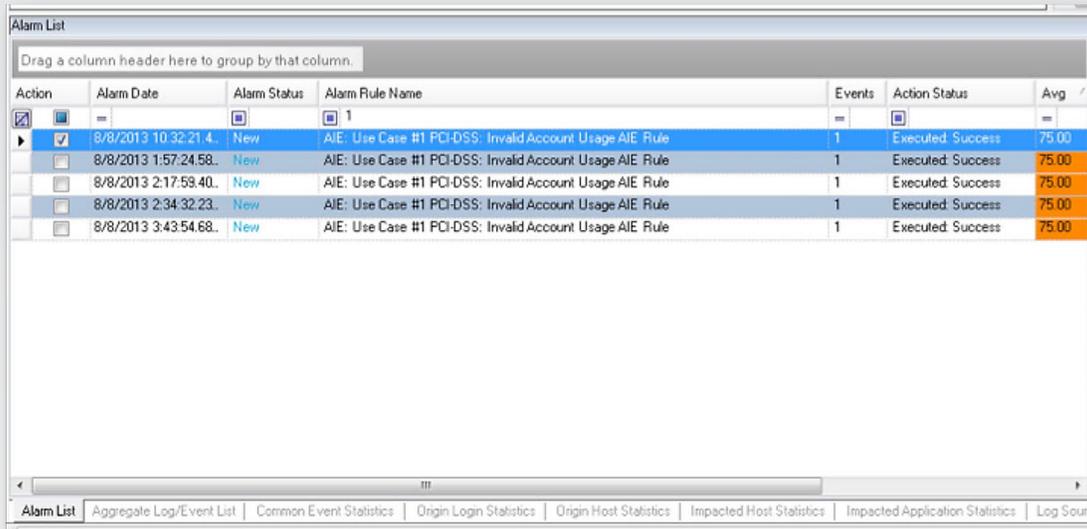


Figure 14. SmartResponse Rule Triggering

## Use Cases (CONTINUED)

Finally, in Figure 15 we see the users that have been added to the “Suspicious Users” List as a result of the rule triggering.

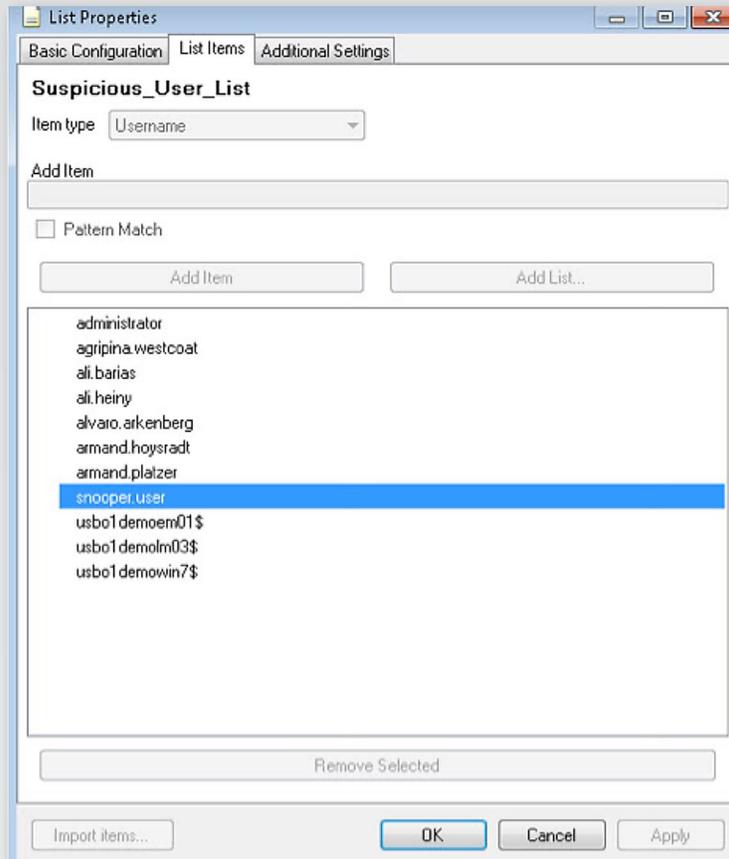


Figure 15. “Suspicious Users”

## Behavioral Whitelisting

The next use case demonstrates LogRhythm’s behavioral whitelisting, which enables the system to build a whitelist of objects while in a “learning mode,” and its capability to learn the users authenticating to a system and then identify a new user whose activity is outside our defined baseline. Any new users who fit this profile will be automatically logged out.

Figure 16 shows the AI Engine’s Behavioral Whitelist Rule Block, which uses LogRhythm’s classification to recognize any authentication success activity targeting a specific system.

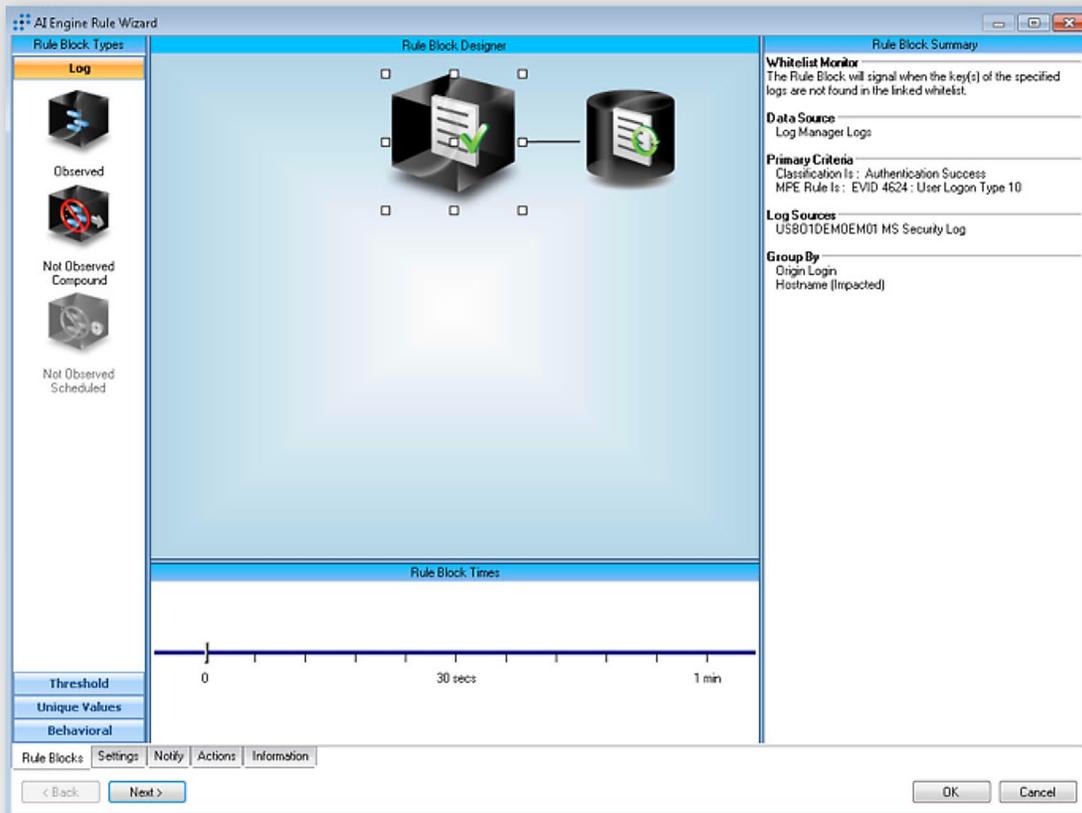


Figure 16. AI Behavioral Whitelist Rule Block

## Use Cases (CONTINUED)

Figure 17 shows the SmartResponse, a forcible logoff of any user outside the baseline; as a safeguard, this action is also marked for approval to listed staff, who must approve each forced logoff.

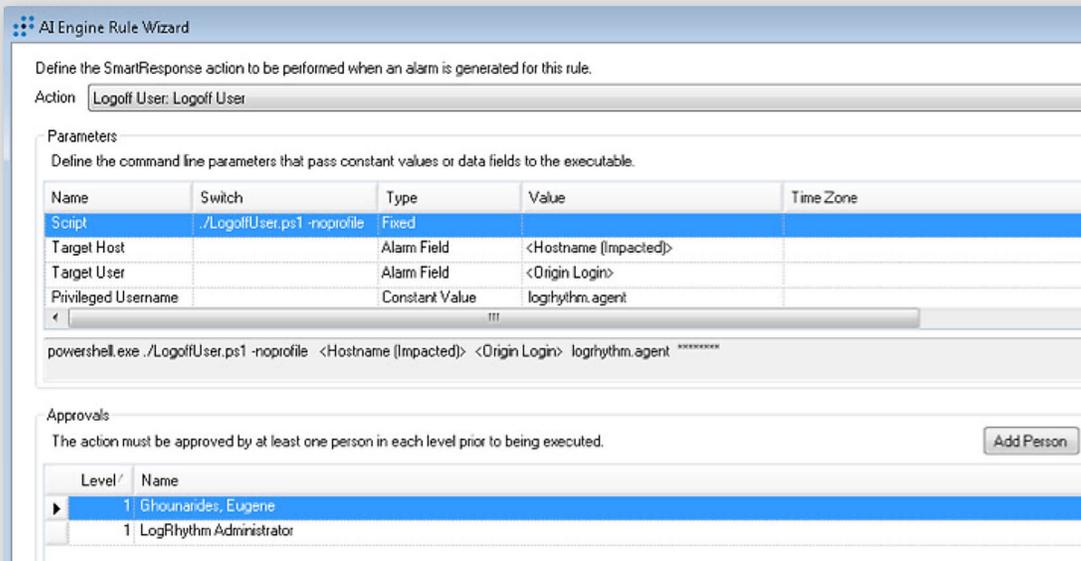


Figure 17. SmartResponse Action for Logging Off Suspicious Users

Finally, Figure 18 shows the Alarm List identifying a new user; with a mouse click, the analyst can approve or deny the SmartResponse action of automatically logging off the user.

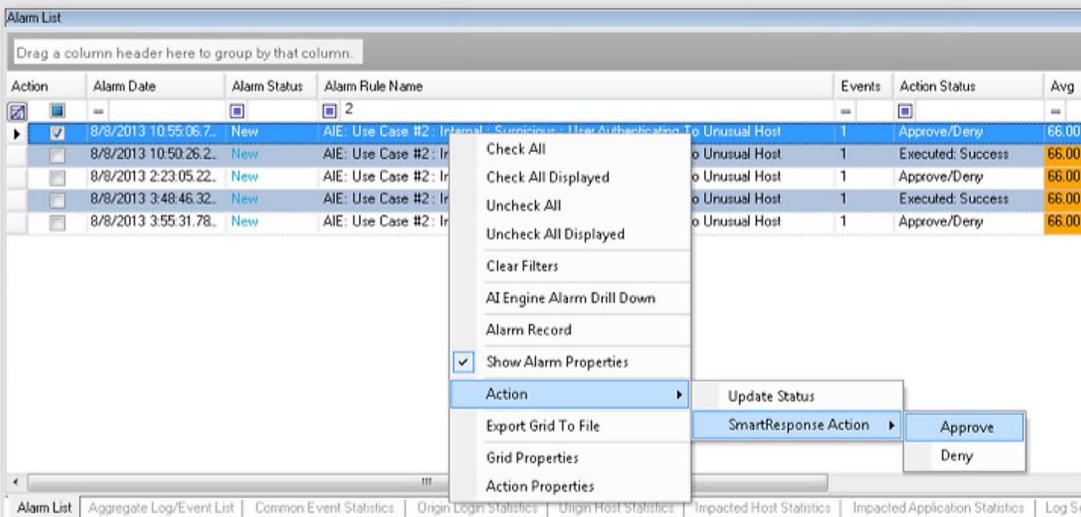


Figure 18. Approving/Denying the SmartResponse Action

## Detecting and Searching for Authentication Violations

The final use case in our exploration of LogRhythm’s advanced analytics shows an AI Engine rule that monitors concurrent authentications from different locations followed by any suspicious activity from the same account. Figure 19 displays the first rule block, observing log activity across the entire deployment and monitoring for authentications of identical users that originate from different city locations within a short amount of time.

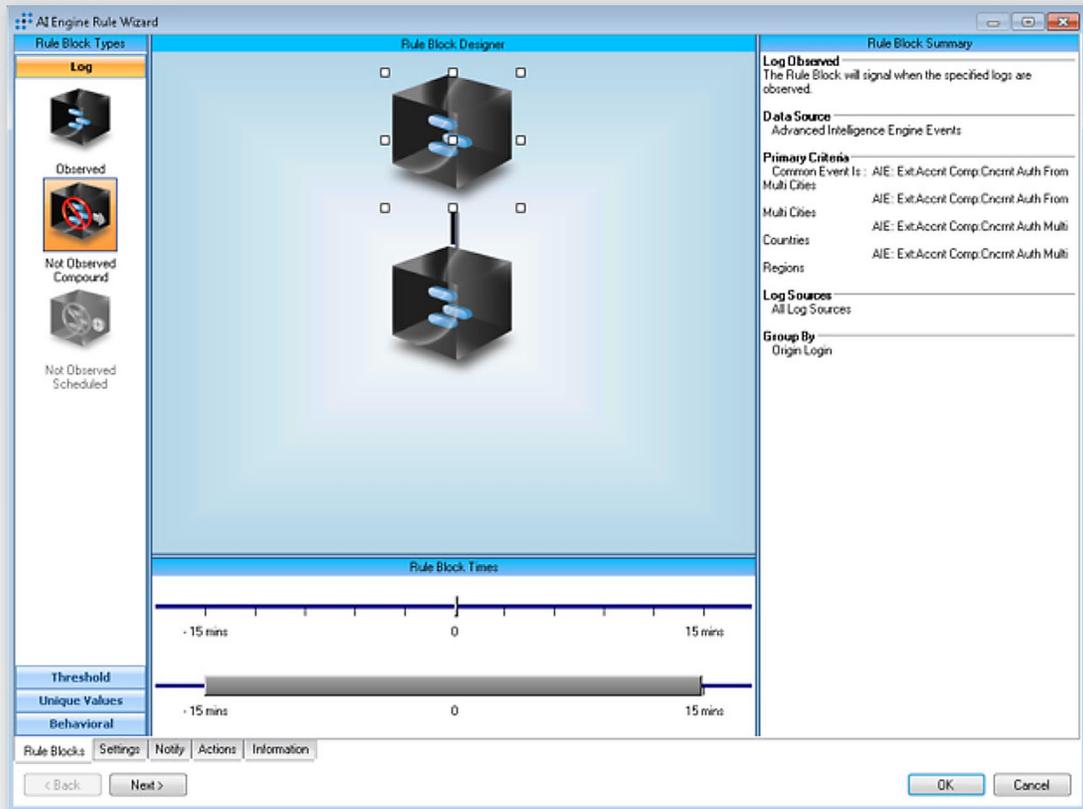


Figure 19. AI Engine Rule for Monitoring Authentication

## Use Cases (CONTINUED)

We set up another AI Engine rule to monitor suspicious activity across a number of security-related classifications. Using LogRhythm’s classifications allows any suspicious activity, regardless of the type of product or vendor, to trigger the second rule block.

Once this was in place, we tested the rule with unusual patterns of authentication, which triggered the alarm shown in Figure 20.

Action	Alarm Date	Alarm Status	Alarm Rule Name	Events	Avg RB	Action Status
	8/9/2013 2:06:18.25...	New	AIE: Use Case #3 : APT: Concurrent Auth From Disp Locations Follow By Susp	1	83.00	
	8/9/2013 2:06:28.33...	New	AIE: PCI-DSS: Attack Alert Rule	1	97.00	
	8/9/2013 2:06:38.74...	New	AIE: PCI-DSS: Malware Alert Rule	1	100.00	
	8/9/2013 2:06:39.85...	New	AIE: PCI-DSS: Compromise Alert Rule	1	100.00	
	8/9/2013 2:06:51.02...	New	AIE: PCI-DSS: Suspicious Activity Alert Rule	1	91.00	
	8/9/2013 2:06:52.04...	New	AIE: Use Case #3 : APT: Concurrent Auth From Disp Locations Follow By Susp	1	83.00	
	8/9/2013 2:07:12.32...	New	AIE: PCI-DSS: Malware Alert Rule	1	100.00	
	8/9/2013 2:07:35.64...	New	AIE: PCI-DSS: Suspicious Activity Alert Rule	1	91.00	
	8/9/2013 2:07:35.64...	New	AIE: PCI-DSS: Malware Alert Rule	1	100.00	
	8/9/2013 2:08:43.09...	New	AIE: PCI-DSS: Suspicious Activity Alert Rule	1	91.00	
	8/9/2013 2:08:43.20...	New	AIE: PCI-DSS: Malware Alert Rule	1	100.00	
	8/9/2013 2:09:03.48...	New	AIE: PCI-DSS: Priv Acct Auth Failure Alert Rule	1	75.00	
	8/9/2013 2:09:23.76...	New	AIE: Internal : Host Compromised : Botnet Zombie	1	63.00	
	8/9/2013 2:09:23.76...	New	AIE: PCI-DSS: Suspicious Activity Alert Rule	1	91.00	
	8/9/2013 2:09:47.23...	New	AIE: PCI-DSS: Malware Alert Rule	1	100.00	
	8/9/2013 2:09:47.40...	New	AIE: PCI-DSS: Malware Alert Rule	1	100.00	

Figure 20. Authentication Rule Alarm

Finally, Figure 21 shows the AI Engine drill-down view that displays all of the activities used to trigger the two AI Engine rule blocks—a combination of the authentication logs and the activities captured by the classifications.

AIE RB#	First Normal Date	Last Normal Date	Count	RBP	Log Source Entity	Log Source Type	Classification	Common Event
1	8/9/2013 2:05:39...	8/9/2013 2:05:39...	1	97	Indeterminate	LogRhythm AI Engine	Compromise	AIE: Ext:Accrnt Comp.Cncl.
2	8/9/2013 2:28:33...	8/9/2013 2:28:33...	1	0	New York Office	LogRhythm Demo File - Co...	Failed Misuse	Failed Unauthorized Activity
2	8/9/2013 1:15:00...	8/9/2013 1:15:00...	1	0	Corporate HQ	LogRhythm Demo File - Da...	Authentication S...	User Logon
2	8/9/2013 1:23:41...	8/9/2013 1:23:41...	1	3	Corporate HQ	LogRhythm Demo File - Da...	Authentication F...	User Logon Failure
2	8/9/2013 2:26:47...	8/9/2013 2:26:47...	1	0	New York Office	LogRhythm Demo File - Co...	Failed Misuse	Failed Streaming Media
2	8/9/2013 2:42:54...	8/9/2013 2:42:54...	1	0	UK Headquarters	LogRhythm Demo File - Co...	Failed Misuse	Failed Unauthorized Activity
2	8/9/2013 1:54:27...	8/9/2013 1:54:27...	1	19	PCI Network Dev...	LogRhythm Demo File - Co...	Misuse	Unauthorized Activity
2	8/9/2013 1:09:23...	8/9/2013 1:09:23...	1	0	UK Headquarters	LogRhythm Demo File - Co...	Failed Misuse	Failed Unauthorized Activity
2	8/9/2013 2:30:42...	8/9/2013 2:30:42...	2	32	Corporate HQ	LogRhythm Demo File - FT...	Access Failure	Command Execution Failure
2	8/9/2013 2:21:06...	8/9/2013 2:21:06...	2	32	UK Headquarters	LogRhythm Demo File - FT...	Access Failure	Command Execution Failure

Figure 21. AI Engine Analysis of Logs and Events

# Searching and Reporting

The ability to direct a holistic query against all data is a critical requirement for investigators, auditors and first responders. In our evaluation of LogRhythm, we confirmed the following:

- The seven Log Managers were distributing the message load across the environment, and all seven performed searches during our queries.
- Simple and complex searches under a sustained load of 100,000 mps returned results within in seconds.
- Within seconds of its collection, results data becomes available for searches. This includes its use in large data set analysis, which affects how quickly users can arrive at solutions to complex problems. LogRhythm supports customized large data set analysis with quick filter-and-pivot to move through data efficiently, arriving at decisions quickly.

## Supports Large Data Sets

We reviewed use cases with large data sets and optimized analysis layouts, aiding users to quickly size up the current environment state and make informed decisions. Using the Investigator Wizard, we performed searches across all seven Log Manager appliances in the testbed, as indicated by the red arrow pointing to the Log Manager instances in Figure 22, a view of the wizard's repository selection pane.

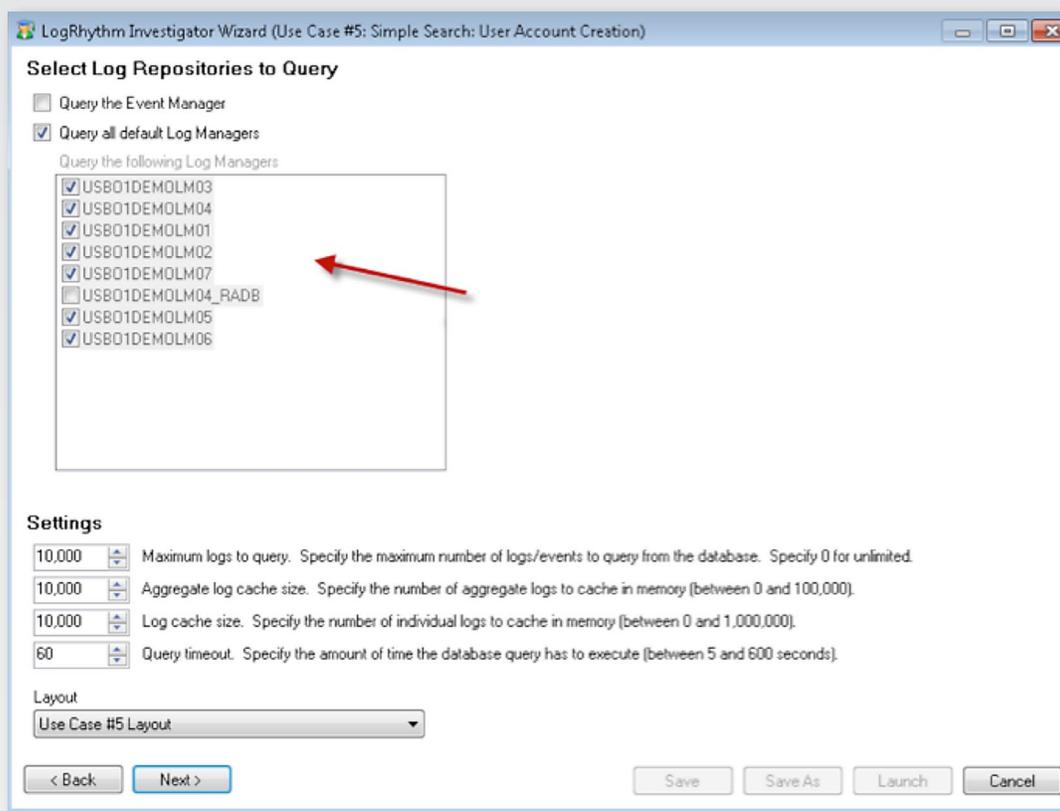


Figure 22. Investigation Spread Across Multiple Log Managers

# Searching and Reporting (CONTINUED)

## Rapid Response

We also validated rapid response times during searches by performing a search of all activity related to logins from the Administrator account over the last 90 days, as shown in Figure 23.

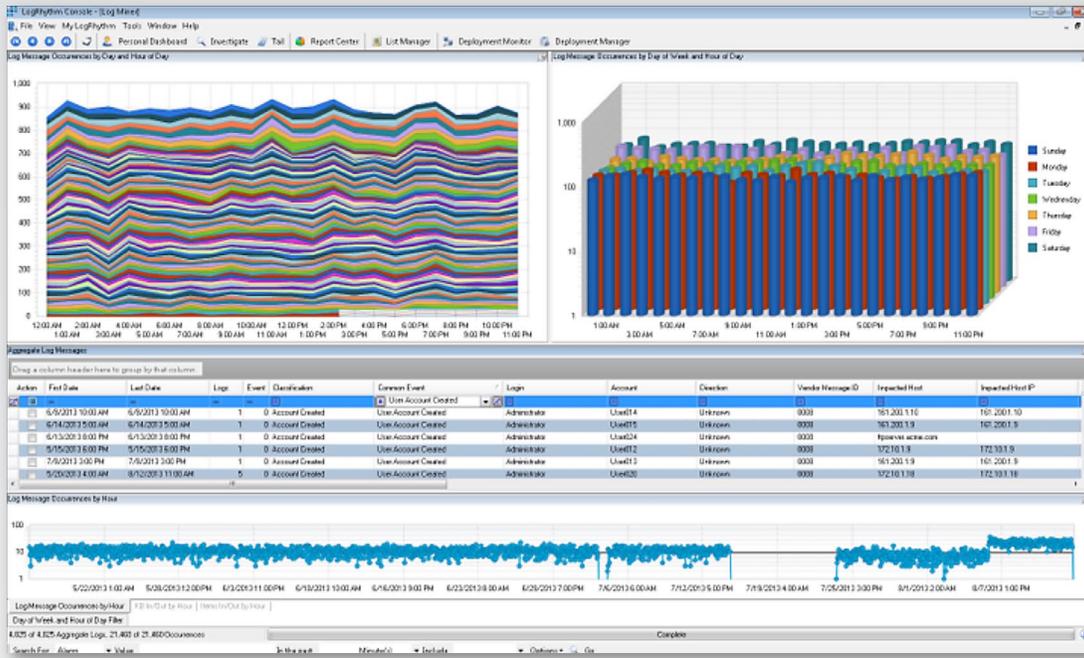


Figure 23. Large, Long-Term Search Results

Daily views of log occurrences appear in the graphs at the top of the screen; in the middle is our filtered data, and at the bottom, another view of occurrences by the hour. We performed the preceding search across all data from the variety of log sources reporting to the system, and it took only seconds to perform.

## Searching and Reporting (CONTINUED)

Then, we tested a more specific search to discover new accounts on any system being monitored. Figure 24 displays the results from searching for all log activity where the event is “Account Added to Group.”

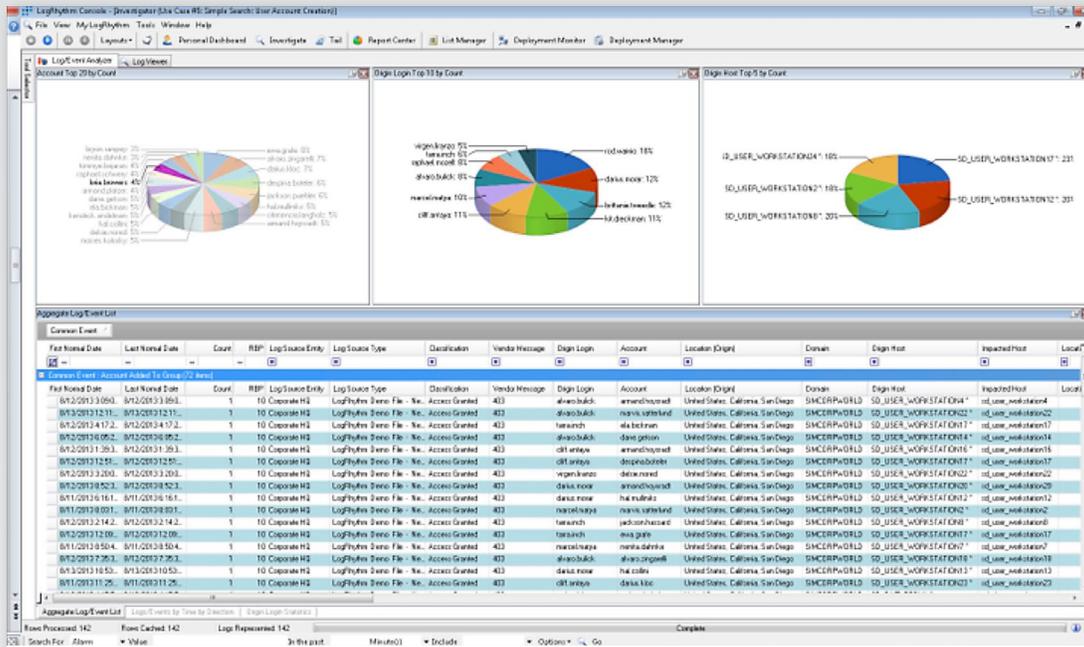


Figure 24. Specific Search for “Account Added to Group”

This search was also very fast, with almost no delay at all. At the top of the screen, we see pie charts illustrating the most active account creators and the most affected hosts, and detailed results appear below.

### Large-Scale Investigation

LogRhythm assigns a Common Event classification to all activity, allowing a search to key on the Common Event assignment, without regard to how different systems or applications may specify this activity in their log syntax. LogRhythm then breaks out the account (in this case, the user added to a group) and the origin login (the user who performed the administrative function) into two different metadata fields. This allows the security analyst to see quickly which account was affected and who was responsible for creating the change.

We configured the investigation window to display the most active accounts, origin logins and the hosts where an account was added to a group. Additional drill-down, filtering and correlation can be performed from the investigation results, and the search is performed across all data at hand, from the variety of log sources reporting to the system.

# Searching and Reporting (CONTINUED)

Finally, we analyzed the tool against large data sets by searching a 339,000-record collection of events for those classified as malware, with the results shown in Figure 25.

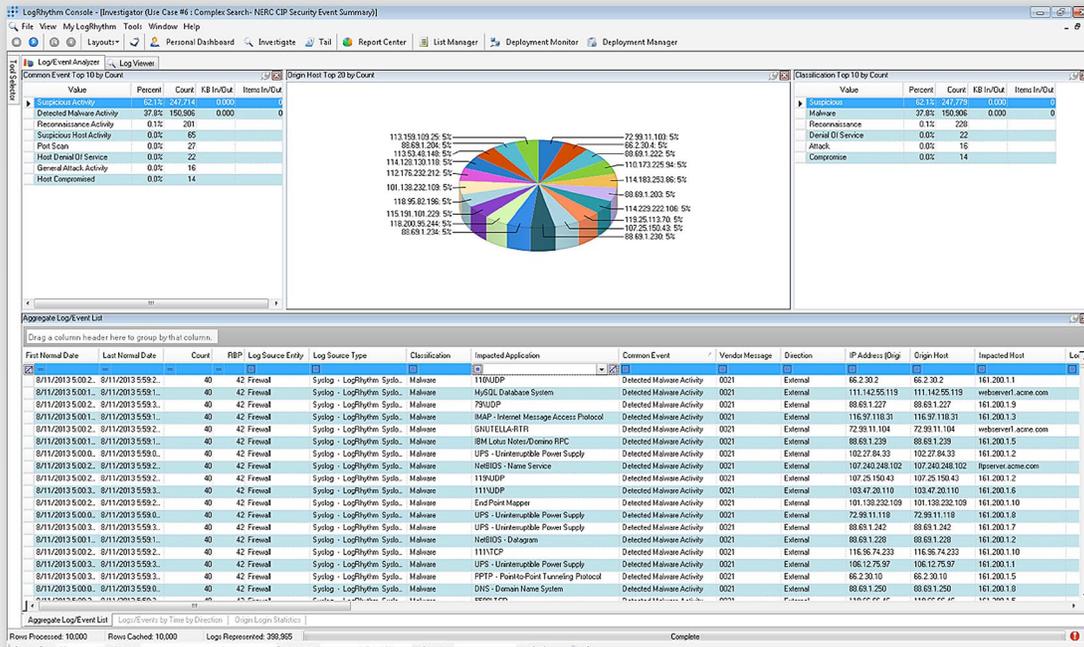


Figure 25. Analysis of a Large Data Set

In this case, more than one-third of a million records were analyzed in less than 30 seconds. This search leverages user-defined classifications as well as LogRhythm’s built-in ones to collect all activity—regardless of the log source type or log syntax—into a category indicating it is security-related and was performed across all data from the variety of log sources reporting to the system.

The layout of the search results is also configurable and supports drill-down, filtering and correlation; such searches can be used as a report and scheduled to run at configurable intervals. Figure 26 displays the reports that correspond to the North American Electric Reliability Corporation Critical Infrastructure Protection (NERC CIP) compliance standard and which reports map to specific controls.

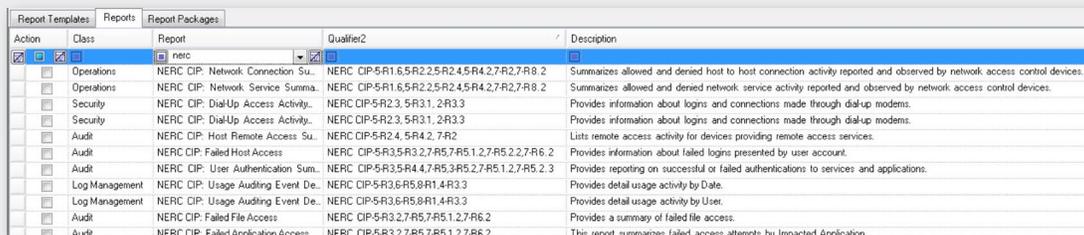


Figure 26. NERC CIP Reports

Ease of use has always been a strength of LogRhythm’s products, and that continues to be the case, based on our experience in this review. Functions were easy to find and define, advanced searches were easy to create and monitor, and the performance metrics in which we were interested were easy to locate within the platform’s console.

## Conclusion

As log management and SIEM platforms have evolved to better address the advanced nature of today's threats, it's become clear that collecting and correlating increasing amounts of data are essential to better monitoring and detection.

However, this "big data" analysis can come at the price of big performance hits. When trying to perform advanced, longer-term queries against high volumes of data from many sources, traditionally architected SIEM platforms will experience significant performance impacts. Poor performance and slow response are unacceptable during incidents where minutes (and sometimes, seconds) count. Therefore, it's important to examine the performance, as well as the features, of such tools.

In this review, we found that LogRhythm's security analytics platform performs as advertised and scales to meet the demands of extremely large and complex enterprise environments. We examined numerous use cases for both performance and capabilities. Our testbed demonstrated LogRhythm's holistic approach to improving the performance and scalability of security analytics platforms. This approach incorporates the traditional areas of consideration, such as collection and processing speeds, while adding persistence, real-time analytics and search categories that reflect the increasingly sophisticated nature of today's IT environments.

## About the Author

**Dave Shackelford** is the founder and principal consultant with Voodoo Security, a SANS analyst, instructor and course author, and a GIAC technical director. He has consulted with hundreds of organizations in the areas of security, regulatory compliance, and network architecture and engineering. He is a VMware vExpert and has extensive experience designing and configuring secure virtualized infrastructures. He has previously worked as CSO for Configuresoft and CTO for the Center for Internet Security. Dave is the author of the Sybex book *Virtualization Security*. Recently, Dave co-authored the first published course on virtualization security for the SANS Institute. Dave currently serves on the board of directors at the SANS Technology Institute and helps lead the Atlanta chapter of the Cloud Security Alliance.

**SANS would like to thank its sponsor:**

