SIEM Integration Best Practices: Making the Most of Your Security Event Logs

What logs to collect and how to process them in your SIEM and beyond

Written by Bryan Patton, principal strategic systems consultant, Quest Software®

INTRODUCTION

Collecting security event logs is vital to detecting and analyzing security incidents. But logs are so voluminous that processing them through your SIEM can be prohibitively costly and overwhelm you with more alerts than you could ever hope to investigate, allowing potential security breaches to go unnoticed in all the noise. As a result, many organizations limit their log collection to their most critical machines — a decision that can leave them unable to properly detect, investigate and remediate security breaches or comply with many regulatory mandates.

But the true root of the problem isn't the mountains of log data; it's the poor model many organizations use to process it. Fortunately, there are better approaches that enable you to collect all your log data and store it cost-effectively for years for forensics and compliance audits, while feeding your SIEM only high-value security events to slash costs and empower security teams to spot true threats.

THE GREAT LOG COLLECTION DEBATE

Many organizations put their SIEM at the center of their log collection, archiving and analysis model. But when combined with a couple of hard facts and an inconvenient truth, this model often leads organizations to make the potentially catastrophic choice not to collect all their log data. Here's how the reasoning frequently goes.

Hard fact: Log collection is essential for security and compliance.

If you want to be able to discover active threats and thoroughly investigate incidents, then you simply must collect every log with security value.

This basic truth is easy to illustrate. Suppose you collect log data only from your highest value servers. One day, using that data, you spot an intrusion on one of your critical servers, which isn't connected to the internet. Where did the intrusion begin? It must have come from another server or workstation in the network — but without the log data from those "non-critical" machines, you can't determine how the infection started and how it spread.

If you want to be able to discover threats and investigate incidents, then you must collect every log with security value.

Moreover, although you can clean up the critical server, you can't determine what other machines were compromised and therefore you can't know if you've cleared out the cancer from your environment. In short, your organization is at continued risk.

Inconvenient reality: Logs are big — really big.

Of course, the basic fact that you need to collect all security logs that could have forensic value quickly bumps up against an inconvenient reality: Everything about logs is big, including:

- The number of endpoints, devices and other resources, both on premises and in the cloud, to collect data from
- The number of events per second and the bandwidth, CPU and throughput needed to process them
- The storage required to archive logs
- The number of different log formats
- The complexity involved in correlating multiple events
- The number of different use cases and interested parties

Moreover, it's important to note that we're not just talking about raw native logs here. You very well might also be collecting large volumes of data using other types of security technologies, such as your:

- Endpoint detection and response (EDR) tool
- Intrusion detection system (IDS)
- Anti-virus (AV) software
- Next-generation firewall (NGFW)
- Virtual private networks (VPNs)
- Data loss prevention (DLP) tools
- Privileged account management (PAM) or privileged session management (PSM) solutions
- User behavior analysis (UBA) technologies

Hard fact: Processing tons of log data through a SIEM is expensive and overwhelming.

That adds up to a huge volume of data being generated, day in and day out. Since many SIEM vendors charge by the volume of data ingested, sending all that data through your SIEM quickly becomes prohibitively expensive. In fact, the SIEM might simply be unable scale to the level required to process all that data without a costly upgrade, making this approach even more untenable for your budget.

Moreover, sending all log data through your SIEM is costly in another way as well: It usually results in a flood of false alarms that quickly overwhelm security teams. Since they simply cannot investigate every alert, they have to guess about what's most important and let everything else go. As a result, they waste time and effort chasing down innocuous events while truly important incidents get missed, sometimes with catastrophic consequences.

Mistaken conclusion: We shouldn't bother to collect all our log data.

Faced with these hard facts and inconvenient truths, all too often, organizations conclude that they simply shouldn't bother to collect all their log data, since they can't use it effectively anyway. This leaves them in a very tough spot: unable to quickly detect suspicious activity, unable to properly investigate incidents, unable to assess their remediation efforts and unable to comply with many regulatory mandates.

Fortunately, this conclusion follows only if you adhere to a log collection and processing model in which you send every log that you collect through your SIEM. But the truth is, just because you collect a log doesn't mean you have to process it through your SIEM. It doesn't even mean you have to monitor or review that log at all. In fact, if you're using a SIEM for event log collection, you are likely grossly overpaying for that capability and not getting the intended value out of the SIEM.

ALTERNATIVE DATA COLLECTION AND PROCESSING MODELS

The way out of this conundrum is not to give up on collecting the log data you need for both security and compliance purposes. Rather, it is to replace your old log data collection and processing model with a better model, one that will enable you to:

- Collect more (and sometimes better) data
- Archive that data cost-effectively for forensics and compliance purposes
- Feed your SIEM with lower volumes of higher quality data, thereby reducing both costs and false alerts
- Expand the possibilities of what you can do with log data beyond feeding it to your SIEM
- Be more agile in adopting new data analysis technologies

Create a single, stable logging pipeline

This alternative model is based on building a single, stable logging pipeline in which you collect logs and other security data once and then selectively send that data to various consumers. One of those consumers might be a SIEM, but there could be others, either instead of a SIEM or in parallel with it, as illustrated in Figure 1.

This strategy offers multiple benefits. First, the various log data consumers can come and go, but your logging pipeline remains the same. As a result, you can be far more flexible and agile in taking advantage of new technologies and retiring those that no longer deliver value. In addition, you can collect the logs just once and be selective about which data to send to each consumer, saving processing time and network bandwidth.

Plus, you eliminate the lag time involved in processing log data through your SIEM before it is sent to other consumers. If tools like your security analytics solution don't get log data until your SIEM gets done with it, their work will be delayed, so you won't be able to respond to incidents as quickly. And if the SIEM goes down, all the tools downstream will be effectively down as well.

Alternative and parallel consumers of logs

Let's briefly review some of the recommended log consumers in a bit more detail:

- **Archival —** At a minimum, be sure to archive all the log data you collect in a centralized, cost-effective storage where it is not in danger of being modified or improperly accessed.
- **Search —** You also want to have basic search capabilities so you can find the data you need for investigations, audits and other purposes.
- **UBA and other advanced analytics —** Ideally, you also want to have advanced

> Sending tons of log data through a SIEM is prohibitively expensive and can drown your security team in a sea of false alerts.

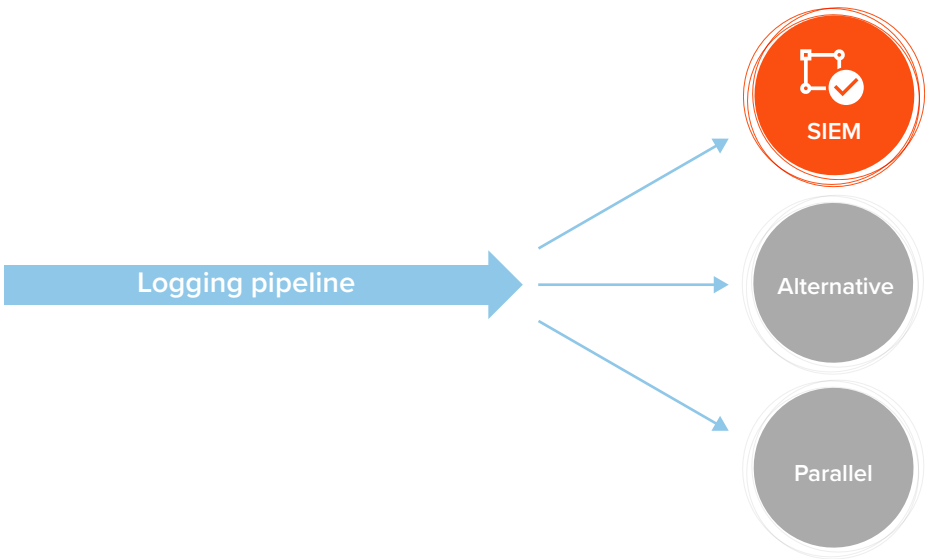

Figure 1. Build a single, stable logging pipeline and send the data to multiple consumers.

Quest

Just because you collect a log doesn't mean you have to process it through your SIEM.

analytics. Most SIEM solutions can be considered real-time analytics, but you might also want non-real-time analytics. For example, you might want to analyze PowerShell events from all your workstations in order to baseline what is normal for your environment to help you spot suspicious activity. Similarly, you might want to build a picture of your normal network traffic, or determine which applications are rarely or never used in your environment. Those analyses are quite valuable but they do not require a SIEM.

Sample models

Figure 2 illustrates just two of the possible ways to orchestrate your overall log flow in a way that enables you to collect all security logs that could have forensic value but not send all of that data through your SIEM. The main difference between

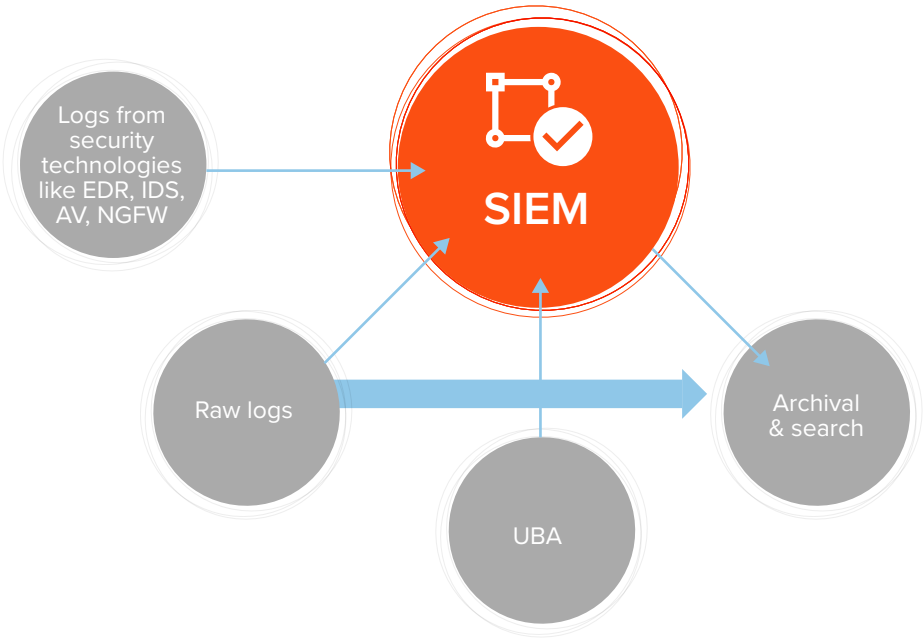

Figure 2, Model A: Sending security and UBA data to a SIEM and then to an archive

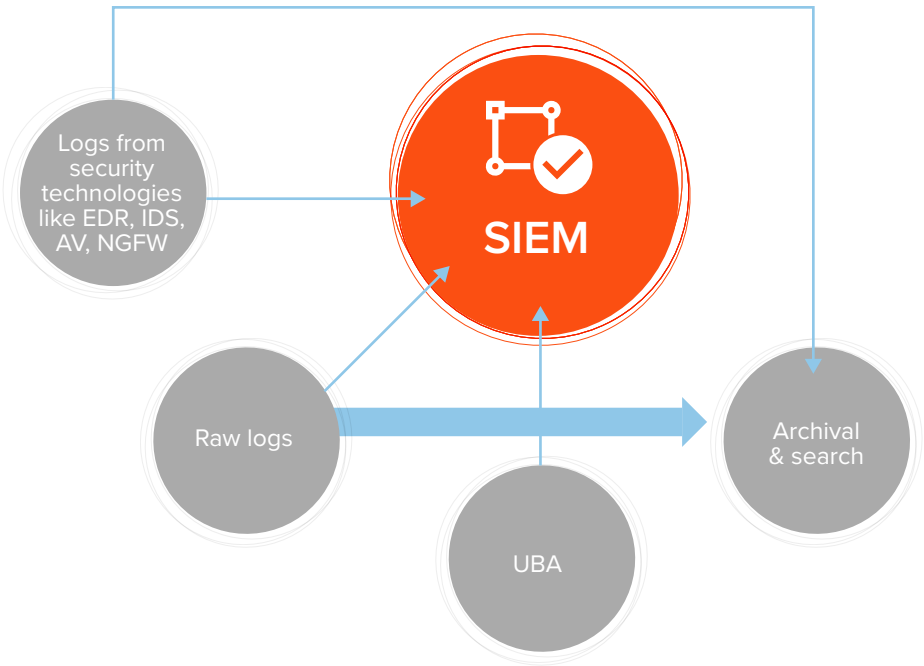

Figure 2, Model B: Sending security and UBA data to a SIEM and an archive in parallel

these two models is whether you send the data from your security technologies and UBA tools only to your SIEM, or to both your SIEM and your archive in parallel. But in both models, you send all the log data to a cost-effective long-term repository and process only the most important logs through your SIEM.

Better yet, you can send just the highest value events, rather than whole logs, to your SIEM. For instance, as both models illustrate, you can archive your entire Windows security log (the fat line emanating from the "raw logs" box), but send only certain high-value events from it to your SIEM (the skinnier line exiting that box). Similarly, an EDR solution might record an event for every program executed on every endpoint; instead of sending all that data to your SIEM, you can pass on just the alerts it creates about security-related activity.

WHICH DATA TO SEND TO YOUR SIEM

Exactly what data should go to your SIEM and what shouldn't? In general, you want to send it everything that deserves real-time analysis that isn't better analyzed elsewhere, along with everything that has already been digested using higher level analysis. This guideline yields more of a continuum of items than a simple yes/no list. Here are my recommendations in rough order of preference.

High priority for SIEM

- Authentication and change events from Active Directory (AD) and any other identity stores, such as cloud access security brokers (CASBs)
- High-level data from security technologies such as VPN, EDR, AV, NGFW, UBA, DLP and PAM/PSM
- High-value and compliance-relevant data from applications, database servers, Exchange, SharePoint and the cloud, such as software-as-a-service (SaaS) applications
- System security events from domain controllers (DCs), high-value servers and devices, and virtualization infrastructure

Valuable but lesser priority for SIEM

- Data from workstations, including security and system data, information on PowerShell activity, and Sysmon data
- DNS and DHCP data
- Flow data from sensors, routers, firewalls and other devices
- Additional data from applications, database servers, Exchange, SharePoint and the cloud, such as software-as-a-service (SaaS) applications
- Web server logs
- File and resource access logs
- Systems management data

Send all your logs to a cost-effective long-term repository and process only the most important data through your SIEM.

With Quest solutions, you can feed your SIEM higher quality data and store your logs cost-effectively for years.

HOW QUEST CAN HELP

The next step is to flesh out this model with quality technologies. Quest offers best-in-class solutions to help in multiple ways, as illustrated in Figure 3.

Data collection

First, consider the data you're collecting. Native logs are simultaneously noisy and incomplete. In particular, they have significant gaps in critical areas like Active Directory and Group Policy changes, as well as most file activity. In contrast, Quest® Change Auditor offers complete coverage without all the noise. By feeding this enriched, lower volume audit data into your SIEM, you reduce your reliance on native logs and eliminate blind spots that could let threats slip past. Moreover, Change Auditor even enables you to protect critical objects from being changed in the first place.

InTrust® can gather log data from across your enterprise, including Windows event logs, UNIX/Linux, IIS and web application logs, PowerShell audit trails, endpoint protection system data, logs from proxies and firewalls, data from virtualization platforms, network device data, custom text logs, and Change Auditor events. InTrust delivers easy and reliable integration with Splunk, QRadar, ArcSight and any other SIEM tools supporting common Syslog formats (RFC 5424, JSON, Snare). You can store long-term event log data with InTrust, and then use its pre-built filters to forward only high-value security data to your SIEM to reduce costs, minimize event noise and improve threat-hunting efficiency and effectiveness. Plus, InTrust can alert you to unauthorized or suspicious user activity and even respond automatically to specific events, for example, by blocking the activity, disabling the offending user, reversing the change or enabling emergency auditing.

Archival

In addition to streamlining log collection across your environment, InTrust is the solution of choice for cost-effective long-term storage of that log data. InTrust's unique storage technology allows for tens of years of data, indexed and always available. And with InTrust's predictable per-user license model, you can collect and store as much data as you need in a highly-compressed repository (20:1 with indexing and 40:1 without), all for a flat fee. In the end, you can save on storage costs by up to 60 percent while satisfying data retention policies and ensuring continuous compliance with HIPAA, SOX, PCI, FISMA and other regulations.

With 10,000 agents writing event logs simultaneously, a single InTrust server can process up to 60,000 events per second, which adds substantial hardware

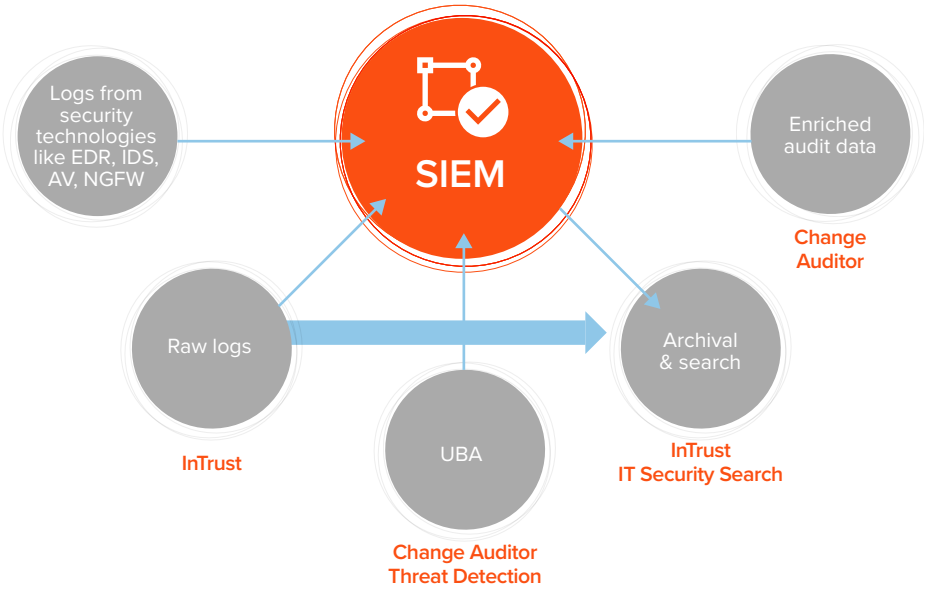

Figure 3. Quest offers best-in-class solutions that can help you build a more effective log collection and processing strategy.

Quest

cost savings. And if you need more volume, you can simply add another InTrust server and divide the workload — scalability is virtually limitless.

Moreover, you can protect your critical data from tampering or destruction, as required for both security and regulatory compliance. InTrust can create a cached location on each remote server where logs are duplicated as they are created; that way, even if somebody cleared the logs, you still have a copy of the original data.

Search

IT Security Search is a Google-like IT search engine that is available as part of several Quest solutions, including both Change Auditor and InTrust. It correlates disparate IT data from many Quest security and compliance solutions into a single console, enabling faster security incident response and forensic analysis across your on-premises or hybrid environment.

You can easily analyze user entitlements and activity, event trends, suspicious patterns, and more, all with rich visualizations and event timelines. For example, the integrated solution enables you to conduct full-text search on long-term event log data and other server data for compliance and security purposes, and search real-time information about changes to critical objects and sensitive data, whether on premises or in Office 365 and Azure Active Directory.

Plus, the solution features role-based access, so you can enable auditors, helpdesk staff, IT managers and other stakeholders to get exactly the reports they need and nothing more.

UBA

Change Auditor Threat Detection offers a unique approach to user threat detection. It models individual user behavior patterns using proprietary unsupervised machine learning and sophisticated scoring algorithms. Then it uses those models to detect truly anomalous activity that could indicate suspicious users or compromised accounts, ranking the highest risk users in your organization. Instead of drowning in false positive alerts, you'll be able to quickly zero in on true threats, such as data exfiltration attempts, malware infections, brute-force attacks and privilege elevation.

Even better, you can leverage your existing Change Auditor infrastructure and audit data to model user behavior, so there's no need to deploy additional agents and servers. A single virtual appliance is the only additional infrastructure required to enable advanced user threat analytics.

CONCLUSION

Don't let your SIEM-centric log data collection and processing model keep you from collecting the log data you need to ensure security and regulatory compliance. Instead, build a stable, efficient logging pipeline that collects native logs, data from your other security technologies and enriched data from solutions like Change Auditor, and then selectively sends that data on to your various consumers. At a minimum, you should send all of the data to a secure, cost-effective long-term archive with advanced search capabilities. Beyond that, you'll want to send selected logs, or selected events within those logs, to your SIEM, UBA and other advanced analytics solutions.

Quest solutions also help you speed threat detection, forensic analysis and incident response.

ABOUT QUEST

Quest provides software solutions for the rapidly-changing world of enterprise IT. We help simplify the challenges caused by data explosion, cloud expansion, hybrid datacenters, security threats and regulatory requirements. We're a global provider to 130,000 companies across 100 countries, including 95% of the Fortune 500 and 90% of the Global 1000. Since 1987, we've built a portfolio of solutions which now includes database management, data protection, identity and access management, Microsoft platform management and unified endpoint management. With Quest, organizations spend less time on IT administration and more time on business innovation. For more information, visit www.quest.com.

If you have any questions regarding your potential use of this material, contact:

Quest Software Inc.
Attn: LEGAL Dept
4 Polaris Way
Aliso Viejo, CA 92656

Refer to our website (www.quest.com) for regional and international office information.