SIEM Migration
A Tactical Approach

True Zero Technologies LLC
5116 Kenwood Drive
Annandale, VA 22003

http://truezerotech.com

DUNS #: 026026373
TIN #: 83-3964542
CAGE #: 8CXQ5

Background

Security Information and Event Management (SIEM) solutions have undergone a transformation over the past ten years. Legacy SIEM solutions relied on strict data structures and relational databases to capture security events and generate detections. This legacy approach created significant performance issues and loss of data quality which led to poor search performance, missed detections, and high response time. Combine that with the increased adoption of cloud and micro services, containerization, and devops/devsecops has created a vastly different landscape from both an operational and security perspective. All of these advancements have led to increased speed and productivity but at the same time created significant challenges to how we monitor, detect, and respond to security and operational incidents.

Modern problems require modern solutions, as they say. Organizations in both the public and private sector have made the decision to migrate to next generation SIEM solutions that provide not only the speed and scalability to achieve near real-time monitoring, but also capabilities that allow for enrichment, integration of machine learning, and provide opportunity to automate actions where possible.

Objective

Leveraging our collective experience and past performance, True Zero has generated this white paper to define a tactical migration strategy that will provide customers with the must know information to make a successful migration. The goal of the migration is to achieve the following key objectives:

- Achieve a like-for-like operating capability
- Reduce or eliminate downtime required for migration
- Achieve cutover flexibility to ensure operational goals are achieved

This white paper assumes that the next generation SIEM (i.e., Splunk Enterprise & Enterprise Security) is already implemented in a production environment according to Splunk best practices and is operationally ready.

Migration Process

SIEM migrations have two primary challenges. First is ***how to on-board data to support both SIEM solutions during the migration process***, and second, ***how to convert use cases from a legacy SIEM to the new SIEM***. The remaining challenges are more operational in nature such as how to update security team operating procedures and response protocols using the new SIEM, but we will address those later.

Let's first start with data on-boarding and then address the challenges with use case migration.

Data On-Boarding

Legacy SIEM's rely on strict data standardization and formatting that forces native log formats to comply with. Examples of this is HP ArcSight's reliance on the Common Event Format (CEF) or McAfee Nitro's reliance on JavaScript Object Notation (JSON). In both of these products it requires strict parsing rules to convert all logs into these formats. This approach has some benefits from a data standardization perspective, but limits flexibility and requires significant work to modify parsers to include additional data, which the SIEM may or may not support.

Splunk Enterprise takes an opposite approach by not caring what the native format is, its goal is to get the data written to disk as quickly as possible and handle parsing at search time. This allows for superior customization, tailoring, and enrichment that can be updated anytime and in an on-demand fashion. This combined with the lack of relational database overhead and the utilization of a flat-file database schema provides customers with extreme performance benefits whether searching over a short or very long period of time.

With this understanding in place, we have a few options to address data on-boarding that ensures both SIEM environments get the data they need allowing side-by-side operation during the migration. We will first start by addressing agent-based collection and then provide information for syslog based collection.

Agent Based Collection

Option 1 – Dual Agent (Recommended)

The first option is to install both SIEM agents on the endpoints that need to feed data into their respective SIEM. This approach ensures each SIEM receives data as each expects it and reduces any dependencies between the two SIEM solutions. As an example, let's say our customer is currently running ArcSight and has ArcSight Collector agents deployed on their Windows and Linux servers. They would simply install the Splunk Universal Forwarder agents on the same servers and perform collection per Splunk best practices.

There are some drawbacks with this approach. Overcoming file permission issues and file lock scenarios can be cumbersome. File lock scenarios occur on Windows systems and if another process has the file opened for read/write can prevent Splunk from accessing the file. To resolve these issues customers should pay close attention to the following:

1. File Permissions:

   Depending on which user the Splunk agent runs as, ensure that user has the necessary permissions to open log files for ingestion.

2. File Lock:

   On windows systems, file locks can prevent Splunk from reading log files. To overcome this, utilize the "MonitorNoHandle" parameter in your inputs.conf.

Additional information:
https://docs.splunk.com/Documentation/SplunkCloud/latest/Data/Monitorfilesanddirectories

Option 2 – Splunk Feed

The second option is to collect all data first in Splunk and then convert/forward event feeds to the third party SIEM. This will be dependent on how the legacy SIEM expects data to be formatted and its ability to ingest various event feeds. One option is to utilize Splunk's app for CEF which can convert native logs into CEF format and then forward to a third-party system:

https://splunkbase.splunk.com/app/1847/

The Splunk app for CEF is ideal for ArcSight integrations, or any SIEM that standardizes on CEF, but for other SIEM's a syslog feed would meet the requirements as long as the legacy SIEM can ingest raw syslog. To accomplish this, configure Splunk to split feeds at the indexers using the following configuration:

https://docs.splunk.com/Documentation/SplunkCloud/latest/Forwarding/Forwarddatatothird-partysystemsd#Syslog_data

Option 3 – Legacy SIEM Feed (Least Recommended)

The final option is to forward all events from the legacy SIEM to Splunk. The biggest issue with this approach is the loss of native log formats and is also a step backwards in terms of the migration as it puts the legacy SIEM in front of the next gen SIEM. Splunk and its community of developers build bundles of configurations called Technology Add-Ons (TA's) to parse native log formats. By converting them to a standard syslog or CEF format would require significant re-work of these pre-built configuration packages and is not sustainable long term.

Syslog Based Collection

For syslog-based collection it's rather straight forward. It is preferable and recommended that customers have syslog aggregators running in the environment to centralize syslog data collection. This can be a Linux based system running either Syslog-NG or Rsyslog, which provides extensive options to collect, route, and store syslog data. In some cases, customers may be forwarding syslog directly to the legacy SIEM systems, which will require a different approach to address collection.

Syslog Aggregation

If your environment is already running Linux based syslog aggregation servers, then it simply requires installing the Splunk Universal Forwarder on the syslog aggregation servers to ingest and send data directly to Splunk. This can be done in similar fashion to the dual-agent recommendations in previous sections if installed alongside a legacy SIEM collector agent.

Direct Syslog

In the event syslog data is configured to send directly to the legacy SIEM application, it is highly recommended to establish syslog aggregation server(s) mentioned in previous sections. This is a best practice for syslog collection and will ensure a scalable and more manageable solution down the road. Once the aggregation servers are installed and configured, simply modify all applications/appliances to send a separate syslog feed to the new syslog aggregation servers and install a Splunk Universal Forwarder on the syslog servers to ingest into Splunk.

Another option that isn't highly recommended, but can work, is to configure the Legacy SIEM to forward all syslog events it receives directly to Splunk. Depending on the legacy SIEM software this configuration/setup can vary, so it is recommended to follow the configuration guides provided by the vendor. (i.e. Configure ArcSight filters and forwarding definitions). The main drawback to this approach is the potential for the SIEM to modify native log formats which will require large amounts of configuration adjustments in Splunk from a parsing and field extraction perspective.

Content Migration

At its core, content migration can simply be viewed as converting all preexisting detections from the legacy SIEM to the new SIEM, accomplishing a like for like monitoring capability. However, this stage becomes complicated as next generation SIEM applications like Splunk have different or new approaches to solving old and dated problems. It is recommended that customers take time to review current content and re-baseline it against current and future security priorities. This generally leads to the removal of dated content that provides little to no value, while opening the door to new approaches that gain greater value for your security operation team.

Legacy SIEM Inventory

First, start by taking an inventory of the current enabled detections from the legacy SIEM. Dedicate time to categorizing and prioritizing this content to help get an understanding of where the current detections are focused as well as determining which detections provide the most value to the security team. An example inventory may look like the following:

Detection	Category	Priority	Notes
Brute Force Access Attempt	Authentication	Medium	Generates considerable false positives, however, when a true positive is detected it leads to immediate action.
Threat IP Access Attempt	Network	Critical	Each alert requires immediate attention as it detects an active attack on the network.
Lateral Movement Detected	Endpoint	High	Mostly true positive, requires immediate action.
Scanning activity detected	Network	Low	Many false positives, not much value

Additional information can be gathered from the legacy SIEM such as reports on false positives by detection, or statistics around the overall number of distinct alerts per detection. Information like this can help show customers valuable information that will influence the priority and integrity ratings of alerts they are currently receiving from the legacy SIEM. For instance, some alerts may fire very frequently, and majority end up being false positives or some may fire very infrequently but lead to actual investigations and remediation.

Inventory Review & Migration Mapping

Once this inventory is complete it is recommended to make broad strokes first and mark detections that are no longer needed based on lack of value or ones that no longer meet operational security goals. This should reduce the size of the list considerably. Once the list is paired down customers should begin the process of identifying new Splunk searches to replace legacy detections, either by utilizing a multitude of available resources or building custom correlation searches from scratch.

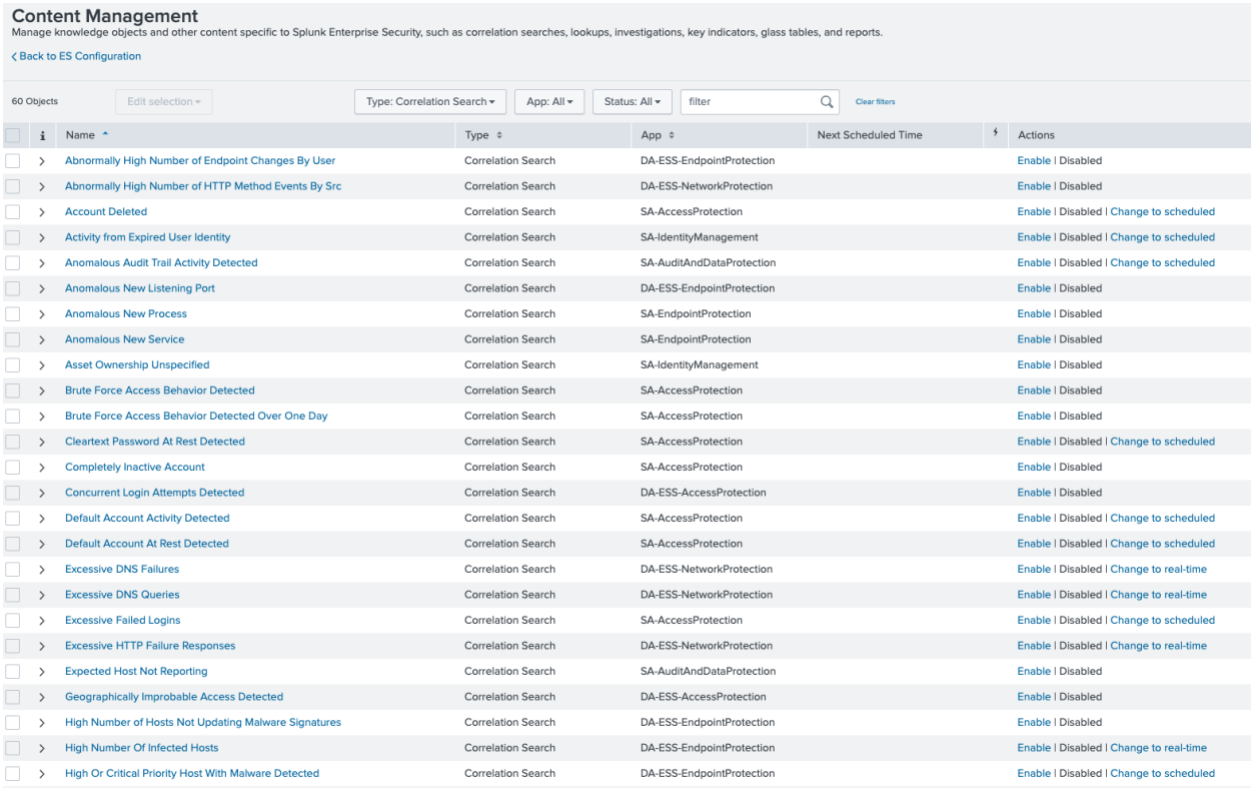
The following sections provide resources that have a lot of pre-built content to meet typical use case requirements.

Out-of-the-box Content

Splunk Enterprise Security comes with a large repository of pre-built content and detections customers can use to monitor active threats on their networks. You can view these available detections within your Enterprise Security deployment by navigating to:

Configure -> Content -> Content Management

and filtering by selecting the “Type” dropdown and selecting “Correlation Search”.



Content Management

Manage knowledge objects and other content specific to Splunk Enterprise Security, such as correlation searches, lookups, investigations, key indicators, glass tables, and reports.

‹ Back to ES Configuration

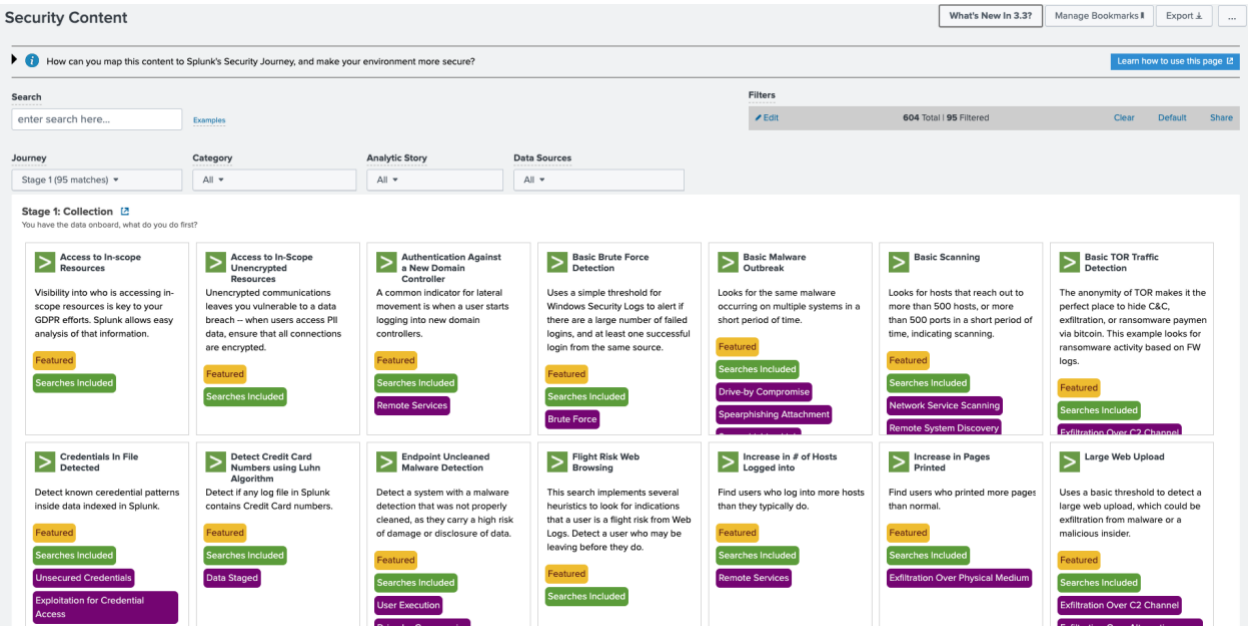
60 Objects | Edit selection | Type: Correlation Search | App: All | Status: All | filter | Clear filters

Name	Type	App	Next Scheduled Time	Actions
Abnormally High Number of Endpoint Changes By User	Correlation Search	DA-ESS-EndpointProtection		Enable | Disabled
Abnormally High Number of HTTP Method Events By Src	Correlation Search	DA-ESS-NetworkProtection		Enable | Disabled
Account Deleted	Correlation Search	SA-AccessProtection		Enable | Disabled | Change to scheduled
Activity from Expired User Identity	Correlation Search	SA-IdentityManagement		Enable | Disabled | Change to scheduled
Anomalous Audit Trail Activity Detected	Correlation Search	SA-AuditAndDataProtection		Enable | Disabled | Change to scheduled
Anomalous New Listening Port	Correlation Search	DA-ESS-EndpointProtection		Enable | Disabled
Anomalous New Process	Correlation Search	SA-EndpointProtection		Enable | Disabled
Anomalous New Service	Correlation Search	SA-EndpointProtection		Enable | Disabled
Asset Ownership Unspecified	Correlation Search	SA-IdentityManagement		Enable | Disabled
Brute Force Access Behavior Detected	Correlation Search	SA-AccessProtection		Enable | Disabled
Brute Force Access Behavior Detected Over One Day	Correlation Search	SA-AccessProtection		Enable | Disabled
Cleartext Password At Rest Detected	Correlation Search	SA-AccessProtection		Enable | Disabled | Change to scheduled
Completely Inactive Account	Correlation Search	SA-AccessProtection		Enable | Disabled
Concurrent Login Attempts Detected	Correlation Search	DA-ESS-AccessProtection		Enable | Disabled
Default Account Activity Detected	Correlation Search	SA-AccessProtection		Enable | Disabled | Change to scheduled
Default Account At Rest Detected	Correlation Search	SA-AccessProtection		Enable | Disabled | Change to scheduled
Excessive DNS Failures	Correlation Search	DA-ESS-NetworkProtection		Enable | Disabled | Change to real-time
Excessive DNS Queries	Correlation Search	DA-ESS-NetworkProtection		Enable | Disabled | Change to real-time
Excessive Failed Logins	Correlation Search	SA-AccessProtection		Enable | Disabled | Change to scheduled
Excessive HTTP Failure Responses	Correlation Search	DA-ESS-NetworkProtection		Enable | Disabled | Change to real-time
Expected Host Not Reporting	Correlation Search	SA-AuditAndDataProtection		Enable | Disabled
Geographically Improbable Access Detected	Correlation Search	DA-ESS-AccessProtection		Enable | Disabled
High Number of Hosts Not Updating Malware Signatures	Correlation Search	DA-ESS-EndpointProtection		Enable | Disabled
High Number Of Infected Hosts	Correlation Search	DA-ESS-EndpointProtection		Enable | Disabled | Change to real-time
High Or Critical Priority Host With Malware Detected	Correlation Search	DA-ESS-EndpointProtection		Enable | Disabled | Change to scheduled

Security Essentials App

https://splunkbase.splunk.com/app/3435/

The Splunk Security Essentials App has a large repository of pre-built content that is categorized and cataloged to provide for easy viewing and filtering. Additionally, it provides a wealth of knowledge in terms of how to respond to a detection, known false positives, and guidance around how to implement. Lastly, it will help you detect if required data is available in your Splunk system prior to implementing. Utilizing the bookmark feature it makes it easy to track selected detections to be used to replace legacy detection content.



Splunk Enterprise Security Content Update

https://splunkbase.splunk.com/app/3449/

Splunk provides an app tailored directly to Enterprise Security that contains a large list of content mapped to the MITRE ATT&CK framework. The ES Content Update app has direct integration with ES that allows for quick deployment of content directly from the app along with many other investigative searches and information.

Custom Detections

The last option is to build custom correlation searches using Splunk's extensible Splunk Processing Language (SPL) and leveraging all of the frameworks provided by Enterprise Security such as:

- Asset & Identity Framework
- Threat Intelligence Framework
- Risk Based Alerting

There are many guides and resources available to guide content creators in the creation process, but a good starting point is to reference Splunk's documentation on creating new correlation search definitions:

- Splunk Docs – Creating Correlation Searches

Additionally, some customers seek to implement predictive analytics to help understand what is considered normal in their environment and base alerting on deviations from that norm. Leveraging the Machine Learning Toolkit will provide the necessary components to begin leverage machine learning tactics from a security monitoring perspective.

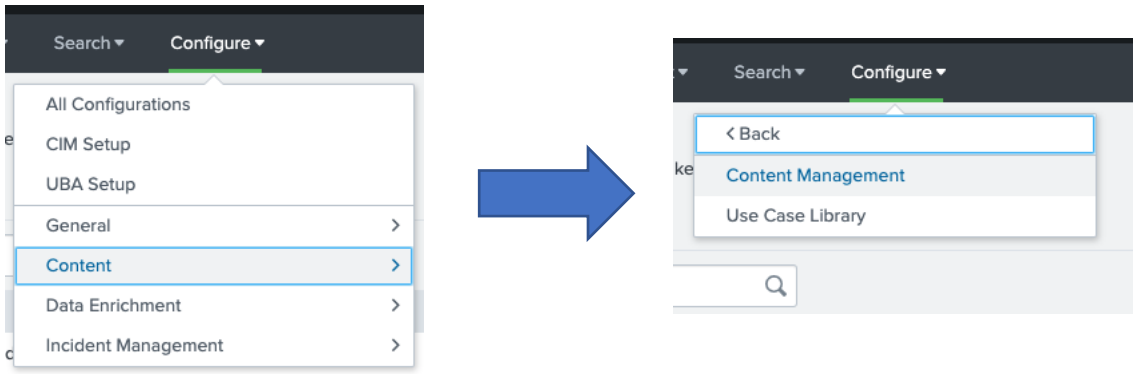
- https://splunkbase.splunk.com/app/2890/

Testing

At this stage you should have a paired down inventory of "must have" legacy detections that are now mapped to Splunk ES correlation searches, either selected from available curated content mentioned in previous sections or custom searches already prototyped in Splunk.

The next stage is to establish a timeframe to conduct unit testing of each of the newly selected Splunk ES correlation searches. This process should take at least 30 days, but depending on the number of correlation searches, could take multiple months. The key objective is to provide adequate time for correlation searches to run and alert so that proper evaluation can occur, and tuning can commence. Additional thought should go into how standard operating procedures need to change and incorporate Splunk Enterprise Security.

First begin by enabling the new correlation searches in ES using their default settings. This can be completed under the Configure -> Content -> Content Management:



As Splunk executes these searches and begins firing notable events (aka Alerts), understand that it will likely lead to a lot of noise and excessive alerts. This is expected! Analysts should begin reviewing the detections and leverage the ES Incident Management framework to track false positives and enter notes that include the reason for the false positive. Establish a weekly schedule to review all fired alerts and identify false positives and use this information to tune searches to improve their fidelity. Establishing a consistent feedback loop is critical to the on-going success of the new SIEM deployment.

Conclusion

SIEM migrations can seem daunting at first, but it becomes straightforward when broken down to its primary challenges and understanding where the real work needs to happen. In following this guide, we identified the two key challenges in doing a SIEM migration, ***on-boarding data to support both SIEM solutions during the migration process*** and ***converting use cases from a legacy SIEM to the new SIEM.*** Although customer requirements can vary, these steps should be fairly consistent environment to environment. The True Zero team is here to support customers large and small in their SIEM migration efforts and our past performance working with both federal and commercial Security Operation Teams affords us the unique opportunity to bring together collective knowledge and lessons learned to new and existing customers. Although this white paper is very targeted to specific aspects of a SIEM migration, the True Zero team provides end-to-end SIEM migration services that cover these topics and more.

The True Zero team hopes this white paper was informative and helpful and we are here to support if the need arises!

Jonathan Cooper

Jonathan Cooper

Vice President, Professional Services
True Zero Technologies, LLC
jcooper@truezerotech.com