Micro Focus ArcSight: Protecting Security Analytics with an Audit Quality SIEM Solution
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Securing the Data Overview</td>
<td>1</td>
</tr>
<tr>
<td>Common Event Format (CEF)</td>
<td>2</td>
</tr>
<tr>
<td>Categorization</td>
<td>3</td>
</tr>
<tr>
<td>CEF Standard and Technology Partners</td>
<td>3</td>
</tr>
<tr>
<td>ArcSight Audit Quality Assessments</td>
<td>4</td>
</tr>
<tr>
<td>Audit Quality—Raw vs. CEF</td>
<td>5</td>
</tr>
<tr>
<td>ArcSight Field Obfuscation</td>
<td>7</td>
</tr>
<tr>
<td>Audit Quality Log Data</td>
<td>7</td>
</tr>
<tr>
<td>Data Encryption with Micro Focus Voltage</td>
<td>11</td>
</tr>
<tr>
<td>Breach Protection and Risk Reduction for Threat Data</td>
<td>12</td>
</tr>
<tr>
<td>ArcSight with Polymorphic Linux</td>
<td>12</td>
</tr>
<tr>
<td>Summary</td>
<td>13</td>
</tr>
</tbody>
</table>
Securing the Data Overview

The goal of a security analytics SIEM platform is to provide visibility across the enterprise by collecting event logs and machine-generated data from a variety of systems and sensors. Once this data is collected it can be used to streamline compliance audits, enhance security posture, and adhere to service level agreements. Micro Focus® ArcSight establishes chain of custody by appending a timestamp from each ArcSight component that processes the event. Where many log aggregation and analytic products have a single time-stamp which is determined by the sender, ArcSight retains multiple time-stamps as it’s processed, providing full visibility into the event lifecycle. Because effective SIEM requires broad event collection, efficient storage, and straightforward analysis of large amounts of log data, ArcSight uniquely addresses those challenges along with simplicity in deployment and management, small to enterprise scale, and elimination of tradeoffs between performance and efficiency.

Having a single database for IT and Security Operations to leverage is powerful, but there are some considerations that should not be overlooked. Is the data “audit quality” that meets industry compliance requirements? Would the integrity of the data stand up in court if needed? Is the data protected and viewable by only those who need it? Remember, adversaries would love to get access to this highly valuable dataset and if companies adopt less secure data analytics platforms there are significant risks.

ArcSight not only provides litigation quality, trustworthy and dependable data, both in transit and at rest, but is also the only SIEM product that integrates with Voltage Format-Preserving Encryption to further encrypt and control what data is accessible, and to enable data privacy compliance across the threat analytics ecosystem. The following features make ArcSight arguably the most robust and secure SIEM platform in the industry:

- Platform Access Controls—Common Criteria Evaluation, FIPS-140 certification
- Built-in Multi-tenant controls for data and application change controls
- Database Access Controls—NIST 800-92 Log Management Standard
- Protected Data Streams—fully encrypted TLS communications between components
- Data Obfuscation—Field level obfuscation of sensitive data
- Format Preserving Encryption—provided by Voltage integration via the ArcSight SecureData Add-on for ADP
- Support for highly secure Polymorphic Linux

In this paper, we will discuss how ArcSight uniquely protects this data while providing litigation quality of data in the collection, storage, archiving, and correlation of log data.
Common Event Format (CEF)

ArcSight utilizes CEF, also known as Common Event Format. CEF addresses the NIST 800-92 requirement, which requires putting data into consistent formats. In addition to preparing data prior to correlation, CEF provides intelligent, accurate, real-time data processing to aid analysts and operators in deriving meaning from log data.

Normalization

This NIST 800-92 standard supports and advocates the process of normalization. Here is an excerpt from NIST 800-92:

“In normalization, log data values are converted to a standardized format and labeled consistently. One of the most common uses of normalization is storing dates and times in a single format. For example, the times when events occurred could be stored in twelve-hour (i.e., 2:34 P.M.) or twenty-four hour (14:34) format, with time zones indicated through different types of notation. In the original data, the event date and time could have had many different labels within individual logs, such as Event Time, Timestamp, and Date and Time. Converting data to consistent formats and labels makes analysis and reporting easier.”

Normalization is the process of taking values contained in an event and mapping them into a standardized schema. Most enterprise environments consist of different types of devices ranging from routers, VPNs, firewalls, physical access systems, along with an even wider range of applications and servers. It is seldom that two manufacturers will use the same logging mechanism, log format, or value representation. The ArcSight CEF format consists of 450+ fields in its schema that log data can be mapped to.

For example, a Cisco PIX will not report an accepted packet in the same way as a Check Point Firewall or even the same as a Cisco Router. The fact that the formats are all different makes it virtually impossible to correlate the events or store the log data in a common database without normalization. The following two events are from different devices from the same company, but the formats are noticeably different:

**Cisco Router:**
Dec 21 12:10:27: %SEC-6-IPACCESSLOGP: list
102 permitted tcp 65.65.65.65(1355) ->
10.10.10.10(80), 1 packet

**Cisco PIX:**
Dec 21 2018 12:10:28: %PIX-6-302001: Built inbound TCP connection 125891 for
faddr
65.65.65.65/1355 gaddr 10.10.10.10/80 laddr
10.0.111.22/80
The previous two log entries are difficult to compare manually and virtually impossible to compare with an automated correlation engine. This problem becomes worse when comparing logs generated from a different brand of firewall:

**Check Point:**

```
"14" "21Dec2018" "12:10:29" "eth-slp4c0" "ip.of.firewall" "log" "accept"
"wwwhttp" "65.65.65.65" "10.10.10.10" "tcp" "4" "1355" "" " " " " " "
" " " " " " " " " " firewall* " len 68"
```

Since all of these formats and fields are significantly different, it would be practically impossible to correlate or report across devices without normalizing the data first. ArcSight uses a comprehensive normalization schema that makes it possible to store log data in a common location and in a consistent format.

**Categorization**

ArcSight orders data into categories and subcategories and has users interact with that structure before looking at the changing data values. Categorization provides a common taxonomy or language that allows users to quickly identify similar events from disparate sources. Categorization enables cross-device correlation or reporting without requiring administrators to know the exact event syntax for each vendor and platform.

An additional benefit of categorization is the ability to seamlessly integrate new data sources without having to rewrite reports or correlation rules. An example would be a company that now has two types of intrusion detection system (IDS) solutions from a recent acquisition. Events are now coming from an IBM ISS IDS as well as the original Sourcefire IDS. Having reports developed based on the categories instead of the vendor-specific event syntaxes eliminates the need to rewrite all of the company’s IDS reports.

**CEF Standard and Technology Partners**

Normalization is consistent with the ‘copy provision’ under the U.S. Federal Rules of Evidence and also provides full indexing, which is recommended as a forensics best practice to speed and aid investigations. Additionally, the NIST 800-92 standard supports the process of normalization. CEF was developed to provide a common taxonomy between the plethora of cryptic messages across a multitude of heterogeneous log sources.

To assist technology companies that want to adopt, test, and certify their compatibility with the CEF standard, ArcSight has formed a CEF certification program. The objective of this program is to provide an all-encompassing program that provides partners with documentation, access to a hosted ArcSight Express/ESM solution, for testing and web support as part of the CEF certification process.
Many major solution vendors have formally committed to supporting the CEF standard as an option for log and event output. You can view the list of CEF certified products on the ArcSight Marketplace.

ArcSight Audit Quality Assessments

ArcSight has undergone Common Criteria Evaluation, FIPS-140 certification, and has had a third-party assessment against the Federal Rules for Evidence by Kahn Consulting Inc.

ADP Delivers Threat Intelligence That Includes Regulated Data

ArcSight delivers a global-scale SIEM solution for ingesting and processing high volumes of event data, including personal data. The ArcSight Data Platform (ADP) Event Broker and ADP Smart Connectors ingest data from wide-ranging sources, with coverage of more than 400 source types. ADP collects data from mobile phones and devices, data centers, applications, cloud services, call logs, web data, laptops, servers, and more—and enriches it in real-time to give analysts organized information for interpretation and action. This massive data flow potentially includes personal data relevant today that is subject to data privacy regulations such as GDPR, the New York State Department of Financial Services (23 NYCRR 500), California Consumer Privacy Act (CCPA2018)—and the host of industry, state, and international data protection mandates focused on protection of consumer information.

The GDPR (Article 32) calls for increased use of encryption of personal data. To provide confidence that event data is not exposed without authorization, ArcSight supports Voltage SecureData with Format-Preserving Encryption (FPE) at the point of data ingestion. ArcSight SecureData Add-on for ADP increases enterprise data protection inflight and at rest, while enabling usability for analytics.

Voltage Format-Preserving Encryption (FPE) is based on next-generation data masking technologies, and is the technology used in the SecureData Add-on. FPE uses NIST-standard FF1 mode AES (Advanced Encryption Standard) to replace sensitive data elements with usable—yet de-identified—equivalents that retain their format, behavior, and meaning. Protection is applied at the data field level and access policy persists with the data itself. Policy controlled secure reversibility enables data to be selectively re-identified in trusted systems that need live data. FPE is the leading, standards-based, industry-vetted method of pseudonymization that de-identifies sensitive classes of event data in line with regulatory mandates, addressing audit requirements and simplifying compliance with a single, integrated approach. The ArcSight SecureData Add-on for ADP encrypts sensitive data on ingestion, and protects it for privacy compliance persistently: as it travels through the enterprise and in threat analytics—without the need for frequent decryption.

Federal Requirements

- NIAP EAL3 (Augmented) Certified against IDS Analyzer Protection Profile (as opposed to just Common Criteria Certified, which means vendor can pick and choose areas of conformance claims—we went against Protection Profile, which means we had to match and meet all requirements)
- FIPS 140-2 Certified Encryption for communication channels
Product has been STIG evaluated and accredited at various DOD and Intel agencies
- DOD CAC and HSPD-12 Card Integration

**Industry Standards**
- CVE compliant and OVAL compatible
- Originator of CEF standard
- Gartner Magic Quadrant Leader (multiple years)
- Various industry awards

**Regulatory Compliance Standards**
- FISMA Content Pack (NIST 800-53 standards)
- IT Governance Content Pack (ISO 27002 standards)
- PCI (Payment Card Industry)
- SOX (Sarbanes-Oxley)
- GDPR

**Audit Quality—Raw vs. CEF**

The ArcSight normalization process preserves 100% of the data from the original event and no data is changed during normalization. Additionally, ArcSight establishes chain of custody by appending a timestamp from each ArcSight component that processes the event.

There are several strategies for preserving raw event information in ArcSight for forensics purposes and therefore you can choose the technique that makes sense for your business.

![Diagram](image)

**Figure 1.** Audit quality logs
1. Do not preserve the raw event—All of the data from the original event is captured and preserved in the ArcSight schema, and fields are not changed in any way that would affect their meaning. This is a valid approach used by many customers and validated against the Federal Rules for Evidence by Kahn Consulting Inc.

2. Enable “Preserve Raw Event” in the ArcSight SmartConnectors—This option stores the entire text of the raw event within a field of the event in the ArcSight Schema. Note that some sources do not have a true raw event as the original message is stored in a binary structure (like Windows and Check Point). For these sources, a raw event that is a close human-readable representation of the original binary structure is stored.

3. Send raw events separately to ArcSight Logger. This technique has the added benefit of allowing longer retention on the raw event, which is typically smaller than a CEF event from a SmartConnector. In this approach, raw log data would be sent directly to Logger over UDP or TCP 514 (Syslog) and a “Storage Group” would be dedicated to storing the raw log data.

Preserving Raw Data
Raw log data can be preserved in two ways—direct log collection into Logger or as an additional data field in the normalized event. However, it has been documented that the normalized event data (CEF) maintains the litigation quality of raw logs.

Regardless of the litigation quality of the CEF format, ArcSight users can preserve the original raw log data by sending data directly to Logger. In this approach, the data would either be sent to Logger via Syslog or written to a log file on the local system and made available through a secure protocol such as SCP or SFTP.

In our example we will assume the latter, being a secured method of retrieving the raw log file. Below details the process.

1. The application or database would use third-party tools or scripts to extract the audit data from the application layer and write it to a log file. The log file will be rotated on a scheduled basis and placed into a directory that has read permissions for Logger. Logger will, in turn, connect to the system(s) on a regular interval via a secure protocol (such as SCP or SFTP) and collect the log files and write them to a storage group locally on Logger (named “RAW” in our example).

2. For correlation, reporting and alerting, ArcSight requires data to be parsed into the normalized, CEF schema. This will require the raw event data to then be forwarded from the Logger appliance to a SmartConnector that will normalize the data into CEF format.

3. Once the event data has been parsed and normalized, it is sent back to Logger, which will, in turn, write the normalized data to a separate storage group (named “CEF” in our example).

4. Finally, once the raw data has been stored and a copy has been normalized, the normalized data can be used for reporting, alerting, and forwarding on to ArcSight Express/ESM for multidimensional correlation, advanced analytics, and workflow integration.
ArcSight Field Obfuscation

ArcSight natively has the ability to obfuscate sensitive data on a per destination basis. For example, a SmartConnector can send normalized events with particular sensitive fields values hashed to ESM, while sending them un-hashed to a Logger appliance. Then through access controls, only special users can search and retrieve this data through a Logger Integration command.

Audit Quality Log Data

The use of logs in compliance audits and litigation requires organizations in to demonstrate that the log data is trustworthy and dependable, both in transit and at rest. Even a few missing or compromised log events can lead to inaccurate audit results or may create doubt around the validity of audit reports.

Integrity checks are enforced in accordance with the NIST 800-92 Log Management standard and ArcSight supports raw data collection across all devices. Granular role-based access controls protect both system objects and event data. All log data can be collected in its native/raw form in addition to the analysis optimized format.

Compliance and litigation chain-of-custody dictates the preservation of the Confidentiality, Integrity, and Availability (C.I.A.) of all event data, whether in transit or at rest. If any of these key criteria cannot be substantiated during the retrieval, processing, or management of log data then the associated inference obtained by the log events themselves becomes inadmissible in a court of law.

Confidentiality—Security event information should only be shared amongst authorized individuals that satisfy the distinction of duty. Security event data should be classified as ‘confidential’ through the organization’s security classification and supported by security policies and processes.
Examples of breaches to the confidentiality of security event data are:

- The systems storing the log data does not have necessary access control and directory permissions to mitigate unauthorized users from tampering with the data
- The data is collected by users with manual intervention and delivered to the parties responsible for storing the data
- The delivery mechanism for the security event data sends the data across public or unsecured networks

**Integrity**—Security event information should be authentic and complete to ensure that this data can be relied upon to be sufficiently accurate for the purpose of security monitoring, alerting, and reporting. The integrity of data is not only whether the data is accurate, but whether it is trustworthy and dependable.

Examples of breaches to the integrity of security event data are:

- Manually pasting information into a spreadsheet and deleting or modifying any portion of the original data to infer meaning with bias
- Collecting security event data with technology or tools that remove any part of the original message or alters the original message in any way

**Availability**—The systems or technology used for managing security event data is always available to collect, store, and process the information. This means the architecture must be fault tolerant so that no data is lost and is always available to the appropriate users.

Since ArcSight was founded in assisting large military customers in protecting their infrastructures against the possibility of Cyber-Terrorism, ArcSight employs the concept of C.I.A. in all of its components to ensure litigation quality of data. Here are descriptions of the ArcSight key technologies and the mapping to C.I.A.:

**ArcSight SmartConnectors**

**Confidentiality**—ArcSight SmartConnectors use a 128-bit encrypted SSL connection to communicate event data between other components such as ArcSight Logger and ArcSight Express/ESM. The SmartConnector can be installed directly on the end device or within a protected DMZ to further protect the confidentiality of the log data.

**Integrity**—SmartConnectors normalize security event data in accordance with the NIST 800-92 standards and 100% of the data from the original event is preserved and no data is changed during normalization. Additionally, ArcSight establishes chain of custody by appending a timestamp from each ArcSight component that processes the event. This is a key differentiator for the ArcSight platform.
Among the event time-stamps retained in each event are:

- **Device Receipt Time/End Time**—This time represents the time the event occurred (or ended), as reported by the device. For sensor devices that may aggregate a series of events or attacks, an additional Start Time may be logged in addition to End Time.

- **Agent Receipt Time**—This time represents when the event was collected by the ArcSight Connector (Agent). If the event is relayed to other forwarding connectors, Original Agent Receipt Time may also be utilized as an additional time-stamp.

- **Manager Receipt Time**—This time represents when the event made its way to the central manager (ESM). Note that one factor in complete normalization is converting timestamps to a common format. Since the devices may all use different time zones, ESM normalization can convert the timestamps to UTC (GMT).

For systems that don’t track these different time-stamps, it is very difficult for an analyst to know for certain when an event occurred. It’s a common occurrence for servers and appliances to have misconfigured date/time. In such cases, where time can be off by hours, days or even years, the correlation rules and the analysts doing historical searches could easily miss critical event data. ArcSight even has the ability to easily determine which systems have misconfigured time through simple built-in Timestamp Functions where the actual time an event received is compared with the time the source server or appliance claims the event occurred. Knowing the accuracy of event-times is critical in establishing context and when providing evidence of malicious activities.

![Figure 3. Example Event Time Stamps as seen in ArcSight Command Center](image)

**Availability**—ArcSight SmartConnectors provide local caching at remote sites, which mitigates the impact of a connectivity loss between remote offices and central log aggregation points that would otherwise lead to a loss of critical event data that may be the missing link in an audit or investigation. SmartConnectors support automated failover to a secondary centralized ArcSight destination (ArcSight Logger or ArcSight Express/ESM) in the event that the primary destination is unavailable. Logs are transmitted and stored reliably—to ensure that critical events (such as logs that indicate compliance violations) are not dropped or lost due to saturated transmissions links, lack of buffers at the source, or unreliable transport protocols.
The ArcSight Express/ESM Manager is already secured and hardened at the application layer. The customer would be responsible for implementing standard OS best practices to secure the Event and user data.

ArcSight Logger

Confidentiality—Logger is a hardened appliance form-factor that only permits connections via SSL. Access to Logger requires authentication and utilizes group permissions to dictate access granularity. There is no access to the underlying file structure and data that is archived off of Logger is digitally signed, compressed, and ‘chunked’ into randomly sequenced digests.

Integrity—Logger creates a message digest of all the messages it receives before committing to storage to maintain data integrity and then digitally signs the data with a secure hash before writing it to a compressed flat file. Logger can also receive data directly from event sources in a ‘raw’ format if it is required for the data to be forwarded to other systems or for litigation purposes.

Availability—Logger can be deployed in a number of high-availability scenarios that offer active-active or active-standby availability. Logger is also provided with connectivity to enterprise Storage Area Network (HBA) environments whereby the data can be stored on the SAN.

ArcSight ESM/ESM Express

Confidentiality—The ArcSight Express/ESM Manager is already secured and hardened at the application layer. The customer would be responsible for implementing standard OS best practices to secure the Event and user data. ArcSight Express/ESM only permits encrypted connections to the application via SSL. ArcSight uses its own, built-in authentication by default but also supports third-party authentication mechanisms, such as RADIUS Authentication, Microsoft Active Directory, two-factor authentication or a custom JAAS plug-in configuration. Access to Express/ESM requires authentication and utilizes group permissions to dictate access granularity. Since the software is installed on customer managed servers, the customer must provide the necessary access control and permissions of the underlying file system.

Integrity—Since the event data is read and correlated in memory in Express/ESM while the data is in transit, there is no effect on correlated events from tampering with the base events in the database.

Availability—ArcSight Express/ESM is architected for high availability through the use of discrete components, automatic component restart, and cached event queues. For example, if the Manager is restarted for some reason, ArcSight SmartConnectors simply cache events to send when the Manager is running again. Likewise, the Manager will automatically suspend and resume operations in the event of Database failure.
Data Encryption with Micro Focus Voltage

Safe, Open Threat Analytics

Enterprises are concerned about exposure of sensitive and regulated data. But most data protection techniques run counter to the needs of the business, which require faster, more comprehensive analytics at scale. Traditional encryption methods, such as CBC-mode AES (Cipher Block Chaining mode Advanced Encryption Standard) have an enormous impact on data structures, schema, and applications. Many data masking transformations create duplicates and destroy referential integrity, causing join operations on database tables to map improperly and reducing the ability to perform analytics on the data. But options are not limited to either accepting the risk of exposure or locking down access to threat analytics.

ArcSight SecureData Add-on for ADP enables FPE at the point of data ingestion with ADP SmartConnectors, automatically applying data protection to the pre-configured fields in the normalized security events created from the raw data. FPE preserves data value and referential integrity across distributed data sets. This means applications, analytic processes, and databases use the protected data without alteration, even across distributed systems, platforms, and tools. Encryption keys are derived on-the-fly, eliminating the need to maintain a key database, and enabling high-volume, real-time threat data ingestion with ADP.

![Figure 4. FPE from the ArcSight SecureData](image-url)
Breach Protection and Risk Reduction for Threat Data

Outdated security models that do not maintain usability, relying instead upon controlling access to authorized parties, are no longer sufficient. If threat data containing personal information is exfiltrated through cyber-attack, but is protected by FPE, it will be unusable by the thieves. Moreover, breach of data protected by FPE can be considered a ‘safe breach’. GDPR waives the requirement for breach notification if the stolen data is encrypted (Article 34). By protecting threat data itself, ArcSight SecureData Add-on for ADP neutralizes the impacts of a data breach, including requirements for breach disclosure, and avoids related brand damage and loss of customer trust.

ArcSight SecureData Add-on to ADP for Simplified Deployment

Micro Focus offers the ArcSight SecureData Add-on for ADP to enable comprehensive data protection at the point of data collection, extending to systems where normalized security events data are shared. It uses Voltage FPE technology to provide end-to-end encryption for enterprise security data collected and enriched by ADP, limiting exposure of sensitive information and thus lowering risk in environments vulnerable to data breach. The SecureData Add-on updates ADP Smart Connectors to add Voltage protection during the data ingestion and normalization flow. The SecureData Virtual Appliance provided with the solution offers transparent, scalable, automated key management for authorized users and applications. Enterprises may deploy as many SecureData appliances as desired to achieve high-availability deployment and high-volume scale.

ArcSight with Polymorphic Linux

Micro Focus has partnered with Polyverse to bring the first Moving Target Defense (MTD)-protected SIEM to market with ArcSight ESM integrated with Polyverse’s code scrambling and polymorphic MTD technologies. This level of protection can be used by those organizations with the highest security requirements, such as those within the intelligence community and the Department of Defense. Polyverse’s Moving Target Defense prevents zero-day memory exploits by leveraging the Polyverse polymorphic version of Linux, which constantly scrambles unique location instructions. For example, their Polymorphic version of Linux randomizes and hardens open source Linux distributions creating a constantly changing attack surface that is extraordinarily difficult for attackers to penetrate. Using their polymorphic compiler, they can dynamically change register usage, function locations, import tables, and so on to produce individually unique binaries that are semantically equivalent.
Summary

Micro Focus ArcSight leads the industry in providing a highly secure and auditable log analytics security platform through its:

- Adherence to NIST 800-92 guidelines
- Native field level data obfuscation
- High integrity data store and complete visibility into log event life cycle
- Granular access controls with multi-tenant protections
- Native integration with Voltage Encryption technology

The ArcSight platform uniquely provides customers with assurance in the audit quality of our technology that meet the most stringent standards. Any enterprise data lake or data repository will become the target of those seeking to exfiltrate valuable data. For this reason, Micro Focus has invested heavily in protecting its security analytics platform. Furthermore, it is the opinion by many governing bodies that ArcSight, with its CEF log format, provides capabilities that support its use as a platform for the management of computer security log files as evidence. Regardless of whether customers would like to retain raw log data for compliance purposes or they prefer to use the litigation quality CEF format, ArcSight provides flexibility for collecting either raw log data, data in CEF format, or both raw and CEF at the same time.

Learn more about ArcSight at:
www.microfocus.com/arcsight

Learn more about the ADP SecureData Add-on at:
www.microfocus.com/media/flyer/arcsightsecuredata_add_on_for_adp_enabling_privacy_compliance_flyer.pdf

Learn more about ArcSight with Polymorphic Linux at:
www.microfocus.com/media/value-brief/on_point_to_secure_vulnerabilities_vb.pdf