Security Information/Event Management Security Development Life Cycle Version 5

If your enterprise is like most, you are collecting logs from most every device with security relevance. The flood of events is probably more than any human can keep up with let alone correlate. This is the role of the Security Information/Event Management (SIEM) system. The SIEM collects log data, normalizes it into a consistent format and allows for cross checking of events from multiple systems. They allow for detailed reporting and the sending notification with a high degree of confidence. SIEM products are rapidly becoming an important part of regulatory compliance monitoring as well.

All this functionality does (of course) come at a price. SIEM solutions are typically complex to engineer and deploy. They can be expensive to purchase and maintain as well. The following Security Development Life Cycle will help guide you through some of the pre and post deployment considerations for a SIEM installation. It is hoped that some of our "lessons learned" contained herein will help you avoid some of the pitfalls in your own SIEM endeavors.

1. **Project planning:**
    a. Determine the need for a Security Information/Event Management (SIEM) solution.
        i. What problem are we solving with this solution (log retention, regulatory compliance, security management, tying together alerts from disparate security systems, consolidation of manpower, etc.)?

        ii. What products are you going to be taking log data from?

            1. Ensure the desired logs can be brought into the SIEM system. Most SIEM solutions offer agent based and agentless data collection capabilities. Ensure capabilities exist to gather and bring in logs in a manner that is consistent with your security architecture.
            2. Plan for some degree of excess capacity, both in hardware and also in software licenses. As you demonstrate the value of the solution you can expect to receive additional tasking to take logs from additional products not on today's roadmap. You may also need to allow for corrections to your initial message volume estimates.

3. Look for the ability to create your own log parsing capabilities if the vendor does not have the capability of reading the logs with a pre-built agent. This can be faster and cheaper than having the vendor create an agent for you if you have to take logs from a custom application.
4. Learn the data to determine how it can be used to provide extra value. Leverage a SME for each data source
5. Involve data-owners early on, and provide access, to improve data-owner cooperation.

iii. What areas of the business are you going to take data from?
1. Plan for some excess capacity or the ability to add capacity easily, you can expect to gain additional customers you have not planned for in the initial design.

iv. What areas of the business are you going to offer services (and which services) to?
1. Plan for some excess capacity or the ability to add capacity easily, you can expect to gain additional customers you have not planned for in the initial design.
2. Ensure that the product allows for easy, granular, and secure Access Controls so that access can be provided to only the appropriate level of data.

b. Determine the financial viability of a SIEM solution
i. Software or appliance costs
1. Central SIEM server or appliance
2. SIEM Agents or collectors (if licensed per-agent)
3. Database Server software (if required)
4. Software modifications needed to support log gathering
5. Software needed to manage the systems (if required)
ii. Hardware costs
1. Servers to run the SIEM (unless deploying an appliance solution)
2. Storage (Local Disk, SAN, or NAS) (if required)
3. BCP hardware or hardware for a high-availability configuration.
4. System management hardware (Tape Backup, Monitoring, hardware management, etc.)
5. Any local costs relating to Data Center space
iii. Bandwidth costs – These will be dependent on the log volumes, you need to ensure you have adequate bandwidth available.
1. Log source to log collector data transfer (will vary with log levels)
2. Log collector to central server data transfer (will vary with log levels)

3. Data transfer within the SIEM infrastructure
4. Client to server traffic
5. Database replication traffic
6. Server and Database Backup traffic.

iv. Customization costs
1. Discuss with the SIEM vendors the feasibility and any costs associated with changes you may require to be made to the product
2. Determine the cost and lead times for any custom agent creation you may require.

v. Maintenance costs
1. Hardware annual maintenance
2. Software annual maintenance costs

vi. Staffing costs
1. Sysadmin(s) salary
   a. Who will do the sysadmins work for your SIEM systems?
2. Content Developer(s) salary
   a. Who will be developing your SIEM content, such as reports, correlation rules, alerts, etc.?
3. Database Administrator
   a. Do you require a DBA (on a full or part time basis)?
   b. Can you split the salary cost between internal groups/projects?
4. SAN Administrator
   a. Do you require the services of a Storage Area Network admin (on a full or part time basis)?
   b. Can you split the salary cost between internal groups/projects?

c. If an enterprise asset inventory system does not exist already, start the effort to build that infrastructure 6 months prior to starting the SIEM implementation. You will want accurate asset information to maximize the value of your SIEM system.
d. Determine if an enterprise Identity management solution exists and if you can leverage this for mapping user identities, both for mapping ID's to users during investigation and also for user access to the SIEM itself.
e. Obtain the most accurate mapping of your network possible prior to the start of the deployment of the SIEM system. Network location is also critical to obtain the most accurate information from your SIEM product.
f. Star the effort to standardize systems on a single time zone, if the business exists in multiple locations consider standardizing on UTC. Standardized log times are very important to correlation. If this is not possible (or practical) then you will need use time correction at the SIEM agent or build your correlation based on the timestamp assigned to events when they are received by the Manager.

g. NTP – if you are not using it, start! Log times have direct impact on correlation of events. The more drift the less likely the events can be correlated with events from other devices.
h. Ensure the data you want to collect is actually being logged by the devices. Nothing is more disconcerting than being ready to start taking in the logs and finding out that they are not being collected.
i. Evaluate SIEM products, try hard to talk with actual users of the products you are considering, preferably without the vendor present.
    i. Determine the main usage scenario for a SIEM solution.
    ii. Determine Hardware requirements
    iii. Determine types of log data support needed and available agent support from SIEM vendors.
    iv. Communication between Log source and SIEM Servers must be authenticated to avoid supplantation and fake data injection. Some log data types, such as UDP syslog do not allow for authentication so other measures (such as tunneling) will need to be taken if authentication is required.
    v. If required, data encryption at the SIEM to guarantee confidentiality (e.g. login and password information, personal or customer information, etc.)
    vi. If required ensure the traffic between SIEM components is encrypted. This is desirable because you may get new requirements after the initial implementation, even if you do not need it on day 1
    vii. Immutability of the logs stored by the SIEM system to ensure they can not be tampered with. (e.g. using digital signatures)
    viii. Raw unmodified log data storage is commonly required for archive storage, evaluate cost-effective raw data storage
    ix. Availability of 24-hour support
    x. Ease of custom content creation
    xi. Complexity of adding completely custom data sources
    xii. Flexibility in modifying the user interface to meet usage requirements
    xiii. Responsiveness of vendor in addressing problems reported
    xiv. Correlation capability to unify disparate data sources
    xv. Ablility to initiate automatic action
    xvi. Compression of transmitted data

2. Systems analysis:

a. Determine the volume of log data (from all sources) you need to be able to accommodate.
    i. Decide what to log at what level of granularity
    ii. Determine which log messages from each log source will be collected by a SIEM solution
    iii. Base numbers on maximum projected log volumes

b. Determine the storage requirements (how long do you need to be able to store logs for)?
    i. Does there exist for your organization, any regulatory mandate to store or destroy data ?
    ii. How long do you need to have online and
    iii. how long in offline (restorable) log data.
    iv. Do you need to replicate database data
    v. Do you need to store raw unmodified log data? If so for how long?

c. Determine how many users the system will need to support
    i. How many users will need to author content
    ii. How many users will be consumers of content only
d. Determine BCP/DR requirements for the SIEM system
e. Do you require external user authentication (via Active Directory, SSO, or Token) ?
    i. Ensure the SIEM supports this mechanism
    ii. Ensure your external authentication system is prepared to support the SIEM application.

3. Systems design:

a. Hardware Requirements (unless an appliance solution is being deployed)
    i. Determine the hardware requirements for SIEM manager and agent servers

b. Design the SIEM architecture
    i. Determine the number of SIEM servers required to support the volume of logs
    ii. Determine the number of SIEM servers required to support any organizational separation of functions (Line of Business or geographical region)
    iii. Determine the number of servers required to support SIEM agents (some SIEM servers have limitations on how many agents they can support)
    iv. Try to architect you SIEM environment in tiers to enhance overall scalability.
    v. Validate SIEM hardware and architecture design with the vendor to avoid any problems later relating to scalability or performance. Ask the vendor to provide a capacity plan that you can use as a scalability roadmap.
    vi. Attempt to design log aggregation points into the architecture. If you need to take in syslog from 25 servers, it is more efficient to have all 25 servers syslog to a log server and run the log collection agent on that log server than it is to run 25 separate log collection agents.

vii. Allow for a Development Manager/DB in your architecture. It is possible to crash/lag a system in the process of creating SIEM content (rules, reports, etc.). Having a non-production system to build and test content on will pay big dividends the first time something being written fails and forces a manager restart.

c. Design SIEM network connectivity
d. Design the SIEM database
    i. Determine the disk space requirements for your SIEM database(es)
        1. Include online and offline storage
        2. determine disk space, speed, and expansion capabilities
        3. Is SAN storage a requirement?
        4. Determine requirements of the SIEM vendor, many have specific requirements for the DB disk space (raid type, raw disk versus file system, number of spindles, partitioning, etc.)
    ii. Allocate space for any database backup or replication requirements
    iii. Allocate space for restoring and re-importing of archived data
e. Train the Implementation team to deploy the SIEM product

4. **Implementation:**

a. Order, Receive, and Rack server hardware
b. Install selected Operating System
    i. Configure to local standards.
    ii. Patch OS to current levels
c. Connect and configure network
    i. Assign IP addresses
    ii. Connect network
    iii. Test connectivity
    iv. Test any network related High availability features
    v. Configure any SAN connectivity (if required)
d. Install SIEM software or deploy an appliance
    i. Load DB (unless installed by SIEM setup)
        1. Load any DB High Availability solution you are going to run now as well. (DataGuard, RAC, etc.)
        2. Test any Database high availability features
    ii. Load SIEM manager software
    iii. Basic manager configuration
e. Install SIEM agents
    i. Load agent software
f. Install System management software
    i. Load backup software
    ii. Load any locally required system management software and agents

g. Design and Implement access controls on user groups to restrict the visibility of events where appropriate.
    i. Many groups only need to see their log data and do not need to be able to see all events in the system
    ii. Build ACL's based on group membership
    iii. Log the logs and audit the auditors: ensure all SIEM access and audit logs are kept so there is a record of SIEM usage
h. Build initial SIEM content
    i. Build content for multiple levels of technical knowledge
    ii. Managers are typically looking for high level abstracted data
    iii. Engineers are looking for content with very detailed information
    iv. Determine requirements for reporting, schedule recurring reports to run automatically

5. Integration and testing:

a. Configure SIEM agent software
    i. Configure it to transmit events to the manager
b. Validate events are being received at the manager from the agents
    i. Check to see all expected events are being received
    ii. Validate the events are being parsed and classified properly
c. Validate the manager is processing events properly
d. Validate data normalization
e. Validate correlation function
f. Validate database archiving capability
g. Validate database restore functionality
h. Test notification functionality
i. Test reporting functionality
    i. Test report dissemination
j. Test any High Availability configuration & current maximum capacity. A stress test prior to brining in any production data is highly advisable.

6. Acceptance, Deployment:

a. Provide access to test user community
    i. Have these users validate content suitability for their assigned roles

b. Build production accounts for user community
    i. Build accounts with appropriate rights
    ii. Disseminate user accounts and software

c. Training of End Users and SOC Personnel in SIEM operation
d. Migrate business processes to new SIEM environment
e. Integrate into the CSIRT/Incident Handling process
f. Educate internal groups on capabilities and limitations of the SIEM product, this can include Audit, Management, and especially Legal.

g. If possible (based on regulatory and legal requirements) develop white list based filtering to prevent your database from getting filled with useless events. While you would like to have every log at your fingertips, the cost is storage and bandwidth can be exorbitant. Determine some local tradeoffs and filter at the log collection point to reduce overhead.

7. Maintenance:

a. Design a process for patching the OS on the SIEM servers
b. Design a process for patching the SIEM application
c. Design a process for patching the SIEM database software
d. Design a process for developing SIEM content to meet business requirements
e. Design a process for emergency content development to satisfy the needs of CSIRT conditions
f. Design a process to archive SIEM data for offline storage
g. Design a process to restore archived SIEM data for later analysis or legal requirements.
h. Design a process for managing user accounts in the SIEM system
    i. Creation
    ii. Deletion
    iii. Password reset/unlock
i. Ongoing training efforts for Administrative and Operations staff
j. Develop a formal process for maintaining SIEM content
k. Create a “lessons learned” feedback loop to allow processes involving the SIEM to be improved based on the incident handling process.
l. Anticipate the need for new agents or upgrades to current ones as new log sources are added or existing applications are upgraded. Ensure your SIEM vendor has a plan to update agents in a timely manner as new application versions are released. Try to get in front of new applications being deployed so you have time to get agents built if they do not exist at the time.

This document has been improved by the contributions of Tom Chmielarski, Anton Chuvakin, Augusto Pasos de Barros, Joanmi Bardera, Peter Vestergaard, and Rafael Goldfarb. Many thanks for all their insightful input.

Dean Farrington