THE SOC ALERT MANAGEMENT PLAYBOOK

How to Manage Alerts

for Security Operations and Incident Response

Bricklayer AI™

WHAT IS ALERT MANAGEMENT?

Alert management is the process of assessing and prioritizing security alerts to determine their urgency, potential impact, and then responding to and reporting on each alert for future learning.

It's a critical function in a Security Operations Center (SOC) that helps security teams efficiently manage and take action on the large volume of alerts generated by various security tools and systems.

HOW IS ALERT MANAGEMENT DIFFERENT FROM ALERT TRIAGE?

ALERT MNGMNT

ALERT TRIAGE

Alert triage is the key first step in alert management. SOCs everywhere are being overrun with alerts, and it is impossible to keep up. It's tempting to only focus on triaging those alerts: putting them in one place, prioritizing them, and having your team focus on the highest priority alerts.

As threats multiply and get more advanced, this becomes harder and harder to scale. Even if you do triage your alerts well, then the bottleneck becomes how to appropriately take action and document each alert for future learnings.

Alert management focuses on the entire lifecycle of every alert instead of just the first (but important!) triage step.

HOW DOES ALERT MANAGEMENT WORK, AND WHAT IS THE GOAL?

8 Key Steps of Alert Management:

1

Initial assessment

Quickly evaluating the basic details of an alert to determine its potential severity and authenticity.

2

Prioritization

Ranking alerts based on predefined criteria to ensure the most critical issues are addressed first.

3

Context gathering

Collecting additional information to understand the full scope and implications of the alert.

4

Validation

Determining if the alert is a true positive or false positive.

5

Categorization

Classifying the alert type (e.g., malware, unauthorized access, policy violation) for appropriate handling.

6

Escalation decision

Deciding whether the alert requires escalation to higher-tier analysts or incident response teams.

7

Initial response

Taking immediate actions, like initiating containment, within the analyst's authority to mitigate potential threats.

8

Documentation

Recording findings, actions taken, and recommendations for future reference.

Goals of Alert Management

GOALS OF ALERT MANAGEMENT

- Reduce response time for critical security incidents
- Minimize the impact of false positives on team resources
- Ensure proper allocation of security resources based on threat severity
- Provide a systematic approach to handling a high volume of alerts
- Improve overall security posture by identifying patterns and trends in alerts

Effective alert triage is crucial for maintaining a strong security posture, enabling SOC teams to focus on the most significant threats while efficiently managing the constant flow of security information.

WHAT TYPES OF ALERTS SHOULD YOUR TEAM TRIAGE AND WHAT IS THE PREVALENCE?

Tier 1 SOC analysts, whether humans or AI agents, will encounter a variety of alerts from different sources. Your team will see:

1 Endpoint alerts

Endpoint alerts are often the most common type that Tier 1 SOC analysts handle. This is because endpoints (like user workstations and servers) are frequently targeted by attackers and generate a high volume of security events. Common endpoint alerts include: Malware detections, Suspicious process executions, Unauthorized software installations, and Unusual login attempts

2 SIEM alerts

Security Information and Event Management (SIEM) systems aggregate and correlate data from multiple sources, so SIEM alerts can be quite common and diverse. Typical SIEM alerts might include: Multiple failed login attempts, Unusual account activity, Data exfiltration attempts, and Correlation of suspicious activities across different systems

3 Network alerts

While perhaps less frequent than endpoint alerts, network alerts are still common and critical. You might see network alerts for: Unusual outbound connections, Port scanning activities, DDoS attempts, or Traffic to known malicious IP addresses

4 Cloud alerts

As more organizations move to the cloud, cloud-related alerts are becoming increasingly common. Cloud alerts often involve: Unauthorized access attempts to cloud resources, Unusual API calls, Misconfigured cloud services, and Data sharing outside the organization

Among these, endpoint and SIEM alerts are often the most frequent for many organizations, with endpoint alerts potentially being the single most common type. This is because endpoints are numerous, directly used by employees, and often the first point of compromise in many attacks.

However, it's important to note that the prevalence of alert types can vary significantly based on your organization's:

- Security stack and tools
- Industry-specific threats
- Current attack trends
- Security program maturity

HOW DO YOU MANAGE EACH ALERT EFFECTIVELY?

1 Initial Assessment

- Identify alert source (endpoint, network, SIEM, cloud)
- Determine alert severity based on predefined criteria
- Check for any related alerts or incidents

2 Context Gathering

- Collect relevant logs and data
- Identify affected assets (users, devices, systems)
- Review recent activities on affected assets

3 Threat Intelligence Correlation

- Check IoCs against threat intelligence feeds
- Look for known attack patterns or signatures
- Assess potential impact based on current threat landscape

4 Alert Validation

- Determine if the alert is a true positive or false positive
- For true positives, assess the current status (ongoing or resolved)
- For false positives, document for tuning purposes

5 Impact Analysis

- Evaluate potential damage or data loss
- Identify any regulatory or compliance implications
- Assess business impact (e.g., service disruption, reputational damage)

6 Containment Check

- Verify if automated containment measures were triggered
- Determine if immediate manual containment is necessary
- Isolate affected systems if required and approved through appropriate management approval chain if applicable

7 Escalation Decision

- Decide if the alert requires escalation to Tier 2/3 or incident response team
- If not escalating, document reasoning clearly

8 Initial Response Actions

- Implement immediate mitigation steps within your authority
- Update relevant stakeholders as per protocol
- Document all actions taken

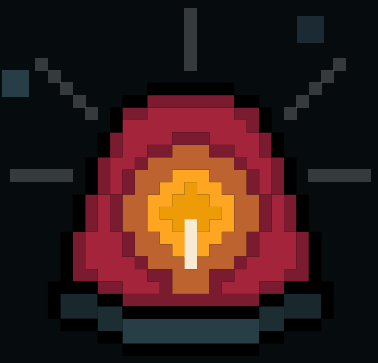
9 Logging and Reporting

- Record detailed notes of the triage process
- Update ticket/ case management system
- Prepare initial report if required

Feedback Loop

- Identify any gaps in detection or response
- Suggest improvements to alerting mechanisms or processes
- Contribute to knowledge base for future reference

HOW DO YOU PRIORITIZE ALERTS?



While following these steps, prioritize based on:

- **Severity of the alert**
- **Potential impact on critical assets**
- **Scope of the potential incident**
- **Current threat landscape**

Remember, the specific order might vary slightly depending on your organization's protocols and the nature of the alert. Always be prepared to adapt your approach based on the unique characteristics of each situation.

HOW DO YOU DETERMINE ALERT SEVERITY?

Predefined criteria for determining alert severity are crucial for consistent and efficient alert management. These criteria often vary between organizations but generally follow similar principles, either with four or five severity classifications. Here's a typical four tier severity classification with examples:

SEVERITY 1
Critical

- Active breach or unauthorized access to critical systems
- Widespread malware outbreak affecting multiple systems
- DDoS attack causing significant service disruption
- Detected exploitation of a critical zero-day vulnerability
- Confirmed data exfiltration of sensitive information

SEVERITY 2
High

- Multiple failed login attempts on privileged accounts
- Malware detection on a critical server
- Unusual outbound traffic to known malicious IP addresses
- Unauthorized changes to firewall rules or security configurations
- Phishing attempt targeting executives or high-value employees

SEVERITY 3
Medium

- Suspicious process execution on non-critical systems
- Anomalous user behavior (e.g., accessing unusual resources)
- Moderate increase in failed login attempts
- Detection of vulnerability scanning activity
- Unauthorized software installation on endpoints

SEVERITY 4
Low

- Single instance of malware blocked on a non-critical endpoint
- Isolated policy violations (e.g., accessing blocked websites)
- Non-critical system patches or updates failing
- Minor misconfigurations in cloud services
- Low-level port scans that were successfully blocked

Factors influencing severity classification

1 Asset criticality

Alerts on mission-critical systems are typically higher severity

2 Potential impact

The possible damage if the threat is realized

3 Scope

Number of affected systems or users

4 Threat actor

Known sophisticated threat groups may increase severity

5 Data sensitivity

Alerts involving sensitive data are often higher priority

6 Attack stage

Later stages of the cyber kill chain are usually more severe

7 Time sensitivity

Some alerts require rapid response to prevent escalation

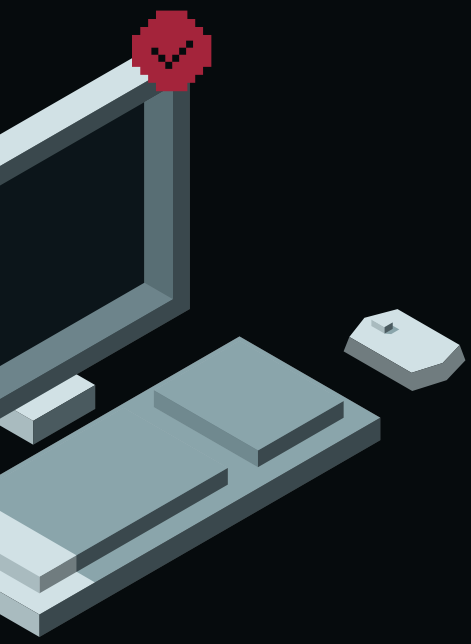
It's important to note that these criteria should be:

- Clearly defined and documented
- Regularly reviewed and updated
- Aligned with the organization's risk tolerance and business objectives
- Flexible enough to account for context-specific factors

WHAT CRITERIA IS USED TO DETERMINE IF AN ALERT IS A TRUE POSITIVE OR A FALSE POSITIVE?

Determining whether an alert is a true positive or a false positive is a crucial skill in alert management.

Here are the key criteria and methods used to make this distinction:



CONTEXT ANALYSIS

Examine the full context of the alert, including:

- Source and destination of the activity
  - Time and frequency of the event
  - Associated user accounts or systems
- Compare with normal behavior patterns for the affected asset

LOG CORRELATION

- Cross-reference the alert with logs from other security tools and systems
- Look for supporting evidence in network logs, endpoint logs, and application logs
- Check if multiple sources corroborate the suspicious activity

THREAT INTELLIGENCE

- Compare indicators of compromise (IoCs) with known threat intelligence
- Check if IP addresses, URLs, or file hashes are associated with known malicious activities
- Assess if the observed behavior matches known attack patterns

ENVIRONMENTAL KNOWLEDGE

- Consider scheduled maintenance or authorized changes that might trigger alerts
- Check if the activity aligns with expected business operations
- Verify if recent system or network changes could cause false alarms

HISTORICAL DATA

- Review historical alerts and incidents for similar patterns
- Check if this type of alert has been previously identified as a common false positive

REPRODUCIBILITY

- Attempt to reproduce the conditions that triggered the alert in a safe environment
- Verify if the same actions consistently trigger or fail to trigger the alert

ASSET CRITICALITY AND VULNERABILITY

- Assess if the affected asset is a likely target based on its role and data
- Check if the asset has known vulnerabilities that align with the alert

ALERT LOGIC EXAMINATION

- Review the specific conditions and thresholds that triggered the alert
- Assess if the alert logic is overly sensitive or prone to false positives

USER VERIFICATION

- When appropriate and safe, contact the user or system owner to verify the activity
- Check if the action was intentional and authorized

FORENSIC ANALYSIS

- For high-severity alerts, perform a quick forensic triage on the affected system
- Look for artifacts that confirm malicious activity (e.g., malware remnants, unauthorized changes)

BEHAVIOR ANALYSIS

- Examine the sequence of events before and after the alert
- Assess if the overall behavior is consistent with a genuine security threat

FALSE POSITIVE PATTERNS

- Be aware of common causes of false positives in your environment
- Check if the alert fits known patterns of false positives specific to your tools or infrastructure

Remember, the goal is to efficiently determine the alert's validity without spending excessive time on benign events. As your human and AI agent team learns, they will quickly improve their ability to spot true positives and common false positives in your specific environment.

WHEN SHOULD MY SOC TEAM CONTAIN A HOST?

Containment is used as an appropriate response to incoming alerts when there's a high likelihood of an active threat that could spread or cause further damage if not quickly isolated. Often containment requires a management approval chain to be followed.

In this case, it is critical to ensure you are following appropriate communication protocols with your stakeholders and following any regulatory reporting timelines.

Here are specific scenarios and considerations for when containment is the right approach:

CONFIRMED MALWARE INFECTION

- When malware is detected on a system and there's risk of it spreadingto other devices
- Especially critical for ransomware or worms that can rapidly propagate

SUSPECTED ACCOUNT COMPROMISE

- If there are signs of unauthorized access or suspicious activity on a user account
- Containment might involve disabling the account or limiting its privileges

DATA EXFILTRATION IN PROGRESS

- When there's evidence of ongoing unauthorized data transfer
- Containment could involve blocking specific network connections or isolating affected systems

ACTIVE LATERAL MOVEMENT

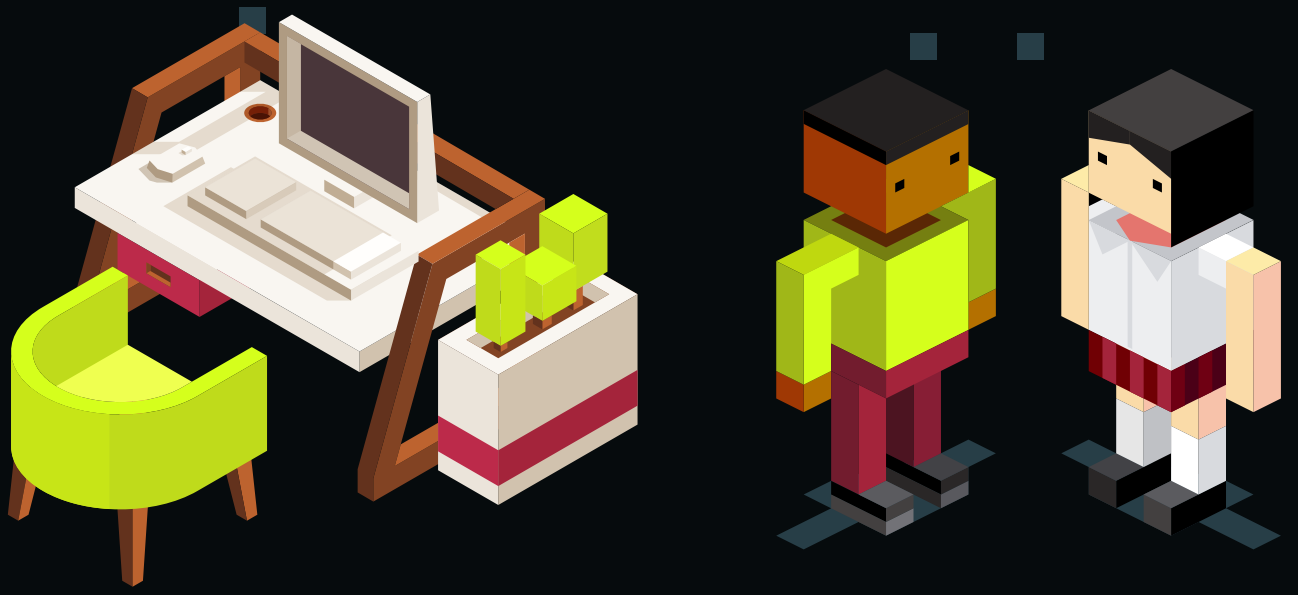
- If there are indicators of an attacker moving between systems within the network
- Containment aims to limit the attacker's ability to access additional resources

EXPLOITATION OF CRITICAL VULNERABILITIES

- When a system with a known, exploitable vulnerability is showing signs of compromise
- Containment prevents further exploitation while patches are prepared

INSIDER THREAT ACTIVITY

- If there's suspicion of malicious activity by an insider
- Containment might involve restricting access to sensitive systems or data



DDOS ATTACK

- When a Distributed Denial of Service attack is detected
- Containment could involve traffic filtering or temporarily isolating affected services

UNAUTHORIZED CHANGES TO CRITICAL SYSTEMS

- If unexpected or unauthorized modifications to important systems are detected
- Containment prevents further changes while the situation is investigated

BREACH OF SEGMENTATION

- When there's evidence that network segmentation has been compromised
- Containment reinforces boundaries between network segments

IOT DEVICE COMPROMISE

- If compromised IoT devices are detected, especially in industrial or healthcare settings
- Containment isolates these devices to prevent them from being used in larger attacks

WHAT SHOULD MY SOC TEAM CONSIDER WHEN CONTAINING A HOST?

1 Speed vs. Accuracy

Balancing the need for quick action with ensuring you're not causing unnecessary disruption

2 Business Impact

Assessing how containment actions will affect business operations

3 Scope of Containment

Determining whether to contain at the network, system, or account level

4 Evidence Preservation

Ensuring containment actions don't destroy valuable forensic evidence

5 Regulatory Requirements

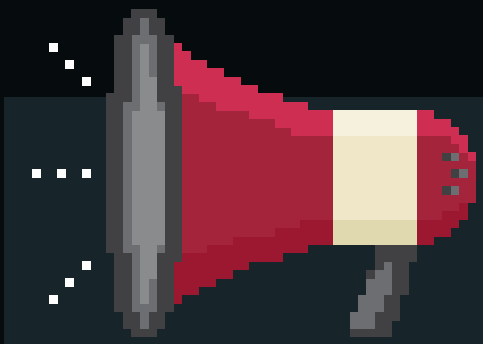
Considering any compliance obligations that might influence containment decisions

6 Attacker Awareness

Implementing containment in a way that doesn't alert the attacker to detection

7 Recovery Preparation

Using containment time to prepare for system restoration and further investigation



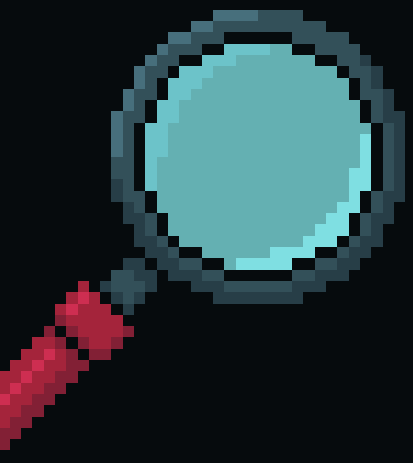
It's important to note that containment is often a preliminary step in the incident response process. It's typically followed by more thorough investigation, eradication of the threat, and recovery procedures.

As always with host containment, it is critical to ensure that the appropriate management approval process is being followed.

HOW SHOULD MY SOC DETERMINE THE SCOPE OF AN INCIDENT?

Determining the scope of an endpoint alert is crucial for effective incident response.

Here's a structured approach to assess the scope:



INITIAL ALERT ANALYSIS

- Examine the alert details carefully
- Identify the affected endpoint(s) and user(s)
- Note the type of activity or threat detected

ASSET IDENTIFICATION

- Determine the role and importance of the affected endpoint
- Check if it's a critical system or contains sensitive data
- Identify the primary user and their access levels

TIMELINE ASSESSMENT

- Establish when the suspicious activity began
- Look for any precursor events or related alerts
- Determine if it's an ongoing issue or a past event

LATERAL MOVEMENT CHECK

- Look for signs of spread to other systems
- Check for unusual network connections from the affected endpoint
- Examine authentication logs for suspicious login attempts on other systems

DATA IMPACT EVALUATION

- Assess what data could have been accessed or compromised
- Check for any unusual data access patterns or exfiltration attempts
- Determine if sensitive or regulated data is involved

USER ACCOUNT ANALYSIS

- Check if the associated user account shows signs of compromise
- Look for unusual account activity across other systems
- Determine if it's a single-user issue or affects multiple accounts

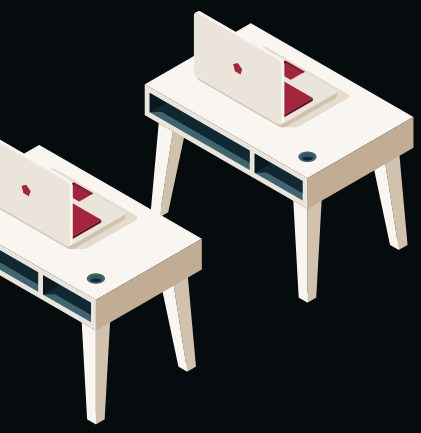
SIMILAR THREAT HUNTING

- Search for similar indicators of compromise (IoCs) across your environment
- Use threat intelligence to identify related attack patterns
- Run queries to find similar behavior on other endpoints

NETWORK TRAFFIC ANALYSIS

- Examine network logs for unusual traffic patterns from the endpoint
- Look for communication with known malicious IP addresses or domains
- Check for anomalies in data transfer volumes or destinations

LOADING...



HISTORICAL CONTEXT

- Review past alerts or incidents involving this endpoint or user
- Check if this is part of a larger trend or an isolated incident

APPLICATION AND PROCESS EXAMINATION

- Analyze running processes and installed applications on the endpoint
- Look for unauthorized or suspicious software
- Check for any tampering with security tools or system configurations

PERIPHERAL SYSTEM CHECK

- Examine systems that frequently interact with the affected endpoint
- Look for signs of compromise on connected devices (e.g., USB drives)

ENVIRONMENTAL FACTORS

- Consider any recent changes in the IT environment that might be relevant
- Check if similar alerts are triggering on systems with similar configurations

SEVERITY AND IMPACT ASSESSMENT

- Based on the gathered information, assess the potential impact
- Determine if this is a localized issue or a widespread threat

ESCALATION EVALUATION

- Decide if the scope warrants escalation to higher-tier analysts or incident response teams
- Consider if external entities (e.g., legal, PR) need to be notified based on the scope

By systematically working through these steps, you can build a comprehensive picture of the alert's scope. This approach helps in prioritizing your response and allocating resources effectively.

Remember, the goal is to quickly determine whether you're dealing with an isolated incident or a potentially larger breach.
As you gather more information, continually reassess the scope and adjust your response strategy accordingly.

HOW SHOULD I THINK ABOUT EXPANDING MY SOC TEAM WITH AI?

1 Identify Your Problem

Understanding your security gaps will help determine where to deploy AI Agents, so you can optimize your operations and drive maximum impact.

2 Use an Expert AI Agent Team

Choose trained AI agents that fill a specific operational role within your SOC which you would otherwise hire a human for. Think security analyst, intel analyst, or incident responder.

3 Train Your Agent Team on Your Procedures

Train your AI agents to search, correlate, de-dupe, or run commands that are essential for completing tasks within your SOC.

4 Run Complex Security Processes Faster

Run multi-task workflows where multiple AI agents and your human team work together to use tools and run tasks to accomplish a complex security process.

5 Scale Your SOC Confidently & Securely

Groups of autonomous AI agents and human experts can and should work together as a human + AI security team, far expanding what human-only teams can accomplish.