Exam Cram Notes: Data Sources for Investigations

1. Overview

In the context of security investigations and incident response, data sources play a critical role in identifying, understanding, and mitigating security threats. The ability to gather and analyze relevant data from various sources enables investigators to piece together the sequence of events that led to the incident, identify the scope of the attack, and prevent future occurrences. Data sources for investigations are typically logs, alerts, and data from security tools and systems.

---

2. Key Data Sources for Security Investigations

A. System and Network Logs

Logs from various systems and networks provide essential information about what is happening within the environment. These logs capture a range of activities, including user access, system errors, network traffic, and more. Key types include:

1. **Operating System Logs (OS Logs)**
    - **Examples:** Windows Event Logs, Linux syslogs, and audit logs.
    - **Purpose:** Track user activity, system events, security incidents (e.g., failed logins, privilege escalations), and hardware issues.
    - **Critical for Investigations:** Can provide crucial evidence for unauthorized access, suspicious activity, and malware infection.
2. **Network Traffic Logs**
    - **Examples:** NetFlow, packet capture logs (PCAP), firewall logs.
    - **Purpose:** Record information about network traffic between devices and services.
    - **Critical for Investigations:** Network logs can help detect communication with malicious IP addresses, data exfiltration, and unauthorized connections.
3. **Application Logs**
    - **Examples:** Web server logs, database logs, application event logs.
    - **Purpose:** Capture events occurring within applications or services.
    - **Critical for Investigations:** Can be used to track user actions, identify faulty application behavior, and detect anomalous activity like injection attacks or privilege abuse.
4. **Firewall and IDS/IPS Logs**
    - **Examples:** Intrusion detection/prevention system logs (Snort, Suricata), firewall logs (Palo Alto, Cisco ASA).
    - **Purpose:** Monitor and block malicious network traffic based on security rules.
    - **Critical for Investigations:** Logs from firewalls and intrusion detection/prevention systems (IDS/IPS) show unauthorized access attempts, abnormal traffic patterns, and network breaches.

B. Security Information and Event Management (SIEM) Data

SIEM systems aggregate, correlate, and analyze logs from multiple sources across the network and endpoints. They provide a centralized platform for monitoring and responding to security events.

1. **Data Aggregation**
    - **Example:** Logs collected from servers, endpoints, network devices, and security appliances are aggregated into a SIEM system.
    - **Purpose:** Helps in detecting correlations between different logs and identifying potential incidents.
    - **Critical for Investigations:** Allows analysts to correlate various events, providing a holistic view of the attack and aiding in faster detection and response.
2. **Event Correlation**
    - **Purpose:** Combines individual events (e.g., failed login attempts) into a meaningful security event (e.g., a brute force attack).
    - **Critical for Investigations:** Simplifies complex data and highlights important events or patterns, reducing investigation time.
3. **Alerting and Reporting**
    - **Purpose:** SIEMs provide real-time alerts based on preconfigured rules or detected anomalies.
    - **Critical for Investigations:** Alerts help investigators prioritize incidents and start the investigation promptly.

C. Endpoint Data

Endpoints, such as desktops, laptops, servers, and mobile devices, generate data that provides critical insights into potential security incidents. Endpoint data can include information on user activities, installed software, and communication with external systems.

1. **Endpoint Detection and Response (EDR) Data**
    - **Examples:** Data from EDR solutions like CrowdStrike, SentinelOne, and Carbon Black.
    - **Purpose:** Tracks the behavior of applications and processes on endpoints, including file modifications, new executables, and network connections.
    - **Critical for Investigations:** Provides a detailed view of what's happening on the endpoints and allows investigators to trace malicious activities.
2. **File System Data**
    - **Purpose:** File creation, modification, and access logs can help identify unauthorized data access, ransomware, or malware-related activities.
    - **Critical for Investigations:** Can help trace malware or unauthorized file manipulation.
3. **Process and Memory Dumps**
    - **Purpose:** Information about processes running on a machine, including any abnormal behaviors (e.g., suspicious processes running with high privileges).

- **Critical for Investigations:** Aids in identifying malicious or unauthorized processes that could be part of an attack.

D. Cloud Data

As more organizations move to the cloud, cloud infrastructure and services generate unique data that is important for security investigations.

1. **Cloud Service Provider Logs**
   - **Examples:** Logs from cloud service providers like AWS CloudTrail, Azure Activity Logs, or Google Cloud's audit logs.
   - **Purpose:** Monitor activities like account logins, resource access, and configuration changes within cloud environments.
   - **Critical for Investigations:** Cloud logs are crucial for tracking the activity of users or attackers across various cloud-based resources and services.
2. **Cloud Security Posture Management (CSPM) Data**
   - **Purpose:** Tools that monitor and manage the security configuration of cloud environments (e.g., misconfigurations, insecure access controls).
   - **Critical for Investigations:** CSPM data helps identify vulnerabilities, poor configurations, and potential access points for attackers.

E. Threat Intelligence Data

Threat intelligence feeds and external sources provide information about known threats, attack tactics, and malicious actors.

1. **External Threat Intelligence Feeds**
   - **Examples:** Threat intelligence from providers like Recorded Future, FireEye, and Open Source Threat Intelligence (e.g., MISP, OpenDXL).
   - **Purpose:** Provides actionable intelligence on known threats, indicators of compromise (IOCs), attack techniques, and tactics (e.g., MITRE ATT&CK).
   - **Critical for Investigations:** Helps analysts correlate observed activity with known threat actors and attack methods, allowing for faster identification of the attack.
2. **Indicators of Compromise (IOCs)**
   - **Examples:** IP addresses, domain names, file hashes, URLs, and other artifacts linked to known threats.
   - **Purpose:** Helps investigators identify compromised systems or data by matching IOCs against internal logs and data.
   - **Critical for Investigations:** Crucial for identifying attacks and preventing the spread of threats.

F. Forensic Data

Forensic analysis is essential for understanding the scope of an attack and preserving evidence for later use in investigations or legal proceedings.

1. **Disk Forensics**
    - **Purpose:** Analyzing hard drives and other storage devices to recover deleted files, uncover malicious artifacts, and piece together the timeline of the incident.
    - **Critical for Investigations:** Essential for deep analysis of compromised systems, recovering evidence, and tracing attacker actions.
2. **Memory Forensics**
    - **Purpose:** Analyzing the system's memory (RAM) to uncover running processes, network connections, and in-memory artifacts left by malware.
    - **Critical for Investigations:** Memory forensics can provide insights into attacks that leave little trace on the file system, such as fileless malware or advanced persistent threats (APTs).

---

3. Importance of Data Sources for Investigations

✅ **Contextual Understanding:** Data sources provide critical context for understanding the nature of the attack, the tools and techniques used, and the ultimate goal of the threat actors.
✅ **Faster Response:** Properly utilizing data sources helps to identify incidents quickly, allowing teams to respond and contain the threat promptly.
✅ **Accurate and Thorough Investigations:** Having multiple, reliable data sources allows investigators to conduct a thorough investigation by corroborating evidence from different angles.
✅ **Evidence for Legal and Compliance Purposes:** Comprehensive data sources serve as legal evidence in case of litigation or regulatory inquiries.

---

4. Best Practices for Data Collection

✅ **Centralize Log Collection:** Use SIEM systems to aggregate logs from multiple sources, making analysis more efficient.
✅ **Preserve Evidence:** Ensure that data is collected and preserved according to best forensic practices, avoiding contamination or alteration.
✅ **Correlate Data:** Link data from different sources (e.g., endpoint, network, cloud) to understand the full scope of the incident.
✅ **Monitor and Update:** Continuously monitor for new data sources and ensure that existing systems are updated with the latest threat intelligence.

---

5. Exam Focus Areas for Data Sources in Investigations

✅ **Understand the key types of data sources** used in security investigations, including logs, SIEM data, endpoint data, cloud logs, and threat intelligence.
✅ **Know the purpose of each data source** and how it contributes to the overall investigative process.
✅ **Understand how to collect and preserve data** in accordance with best practices, ensuring that data integrity is maintained for future analysis.
✅ **Familiarize yourself with tools** used to aggregate, analyze, and correlate data, such as SIEM, EDR, and forensic tools.