Infrastructure Security Architecture for Effective Security Monitoring

GIAC (GCIA) Gold Certification

Author: Luciana Obregon, lucianaobregon@hotmail.com
Advisor: Mark Stingley
Accepted: December 2, 2015

Abstract

Many organizations struggle to architect and implement adequate network infrastructures to optimize network security monitoring. This challenge often leads to data loss with regards to monitored traffic and security events, increased cost in new hardware and technology needed to address monitoring gaps, and additional Information Security personnel to keep up with the overwhelming number of security alerts. Organizations spend a lot of time, effort, and money deploying the latest and greatest tools without ever addressing the fundamental problem of adequate network security design.

This paper provides a best practice approach to designing and building scalable and repeatable infrastructure security architectures to optimize network security monitoring. It will expand on four network security domains including network segmentation, intrusion detection and prevention, security event logging, and packet capturing. The goal is a visual representation of an infrastructure security architecture that will allow stakeholders to understand how to architect their networks to address monitoring gaps and protect their organizations.
1. Introduction

The biggest challenges that Information Security departments face is identifying the critical assets that makes an organization unique, locating these assets on the network, and building security defenses around them while maintaining functionality. This lack of knowledge has driven organizations to implement a "uniform protection" approach to security. With uniform protection, every system on the network is treated as equally important and must be equally monitored for malicious activity and signs of compromise (SANS Institute, 2013). This approach drastically increases costs with regards to the technology and infrastructure required to scale security monitoring, the people needed to review the large volumes of collected data, and the improvement of processes to streamline incident response activities. In most cases, monitoring each and every system on the network is not only financially unfeasible, but also gives an organization a false sense of security.

In an information-centric approach to defense-in-depth, an organization identifies its most valuable data, building layers of defense around it to protect its confidentiality, integrity, and availability (SANS Institute, 2013). Assuming that an organization has already identified and classified its most critical data, the next obstacle to overcome is to architect a network infrastructure with security in mind to systematically protect and monitor the systems that store, process, and transmit the critical data.

A key tenet in Information Security is "prevention is ideal, but detection is a must" (Dr. Eric Cole). We cannot protect what we cannot see, and to increase visibility in those areas of the network that are critical to the business the first step is to segment the network into security zones. Network segmentation is a fundamental component of an information security strategy; it reduces the likelihood of a compromise from spreading, increase visibility into network traffic, and is the foundation for building a secure network. Without network segmentation, an attacker inside the network can access everything. Once a network is adequately segmented, security controls can be distributed across the secure zones to reduce the risk of compromise, closing monitoring gaps and increasing the visibility of network activity.
2. Logical Network Security Segmentation

1.1. Overview

A network segment, also known as a network security zone, is a logical grouping of information systems in an enterprise network. An enterprise network is divided into manageable network segments to reduce the scope of compliance, limit data exfiltration, and reduce the attack surface (Palo Alto Networks, n.d.).

A network security zone has a well-defined perimeter and strict boundary protection, and the systems within it are susceptible to similar types of cyber threats (Government of Canada, 2007). Each zone has different security requirements depending on the systems hosted within. For instance, an end-user workstation hosted in the Enterprise Zone will not be given the same level of protection as a human resources database that stores Personal Identifiable Information (PII) in the Restricted Zone. However, there are general security guidelines that all zones should adhere to (Government of Canada, 2007):

- Identification and classification of groups of systems and resources
- Each zone has one discrete point of entry defined by a stateful inspection firewall
- Only business required traffic is allowed to leave or enter a zone
- Inbound and outbound zone traffic is filtered and monitored at the perimeter

Establishing a small number of network security zones with clearly defined security requirements limits the complexity and removes ambiguity when selecting a zone for new systems while meeting most business requirements (Government of Canada, 2007).

1.2. Rationale

Network segmentation is part of a defense-in-depth strategy with the following goals (Palo Alto Networks, n.d.):

- Limit the scope of regulatory compliance
- Reduce data exfiltration
- Reduce attack surface

Luciana Obregon, lucianaobregon@hotmail.com
• Compartmentalize systems

• Increase availability

Systems that are subject to regulatory compliance such as PCI or HIPAA can be compartmentalized into a logically isolated zone to limit the scope of compliance and therefore, reduce costs and time needed to complete tedious audit processes (Palo Alto Networks, n.d.)

The attack surface of the systems within a zone can be significantly reduced by exposing a limited number of services through the zone's perimeter and implementing rigorous access controls to limit access to specific groups of users. Additionally, if a breach occurs, an attacker would have to compromise access to the all of the outer zones before getting to the zone where the critical data is stored, reducing the likelihood of data exfiltration and increasing the availability of critical systems (SecureArc, n.d.).

1.3. Network Security Zones

This paper introduces the following network security zones and the corresponding trust levels (SecureArc, n.d.), (Government of Canada, 2007):

• Internet Zone - No Trust
• External DMZ - Low Trust
• Enterprise Zone - Medium Trust
• Extranet Zone - Medium Trust
• Internal DMZ - High Trust
• Management Zone - Highest Trust
• Restricted Zone - Highest Trust

1.1.1. Internet Zone

The Internet Zone includes the Internet, the Public Switched Telephone Network (PSTN), and any Internet Service Provider (ISP) public backbone networks (Government of Canada, 2007). The Internet Zone is the least trusted zone. It is an extremely inhospitable zone where anonymous threat actors live. This zone is typically outside of the control of the organization.
1.1.2. External DMZ

The External DMZ houses systems that require exposure to the Internet. This zone proxies access between systems in the Enterprise Zone and the Internet. All traffic should be funneled through systems in the External DMZ to reach Internet resources. The systems deployed in this zone should be tightly controlled and hardened to reduce the attack surface. Typical systems that live in this zone are:

- External web servers
- E-mail gateways
- FTP servers
- Web proxy servers
- Remote access services

1.1.3. Enterprise Zone

The Enterprise Zone is where end-user systems reside, including end-user workstations, printers, and VoIP Phones. Endpoint protection is a critical control in this zone to limit the exposure of end-user systems to malware.

1.1.4. Extranet Zone

The Extranet Zone houses connections with highly trusted 3rd party business partners and can be an extension of the Enterprise Zone. Nonetheless, it is recommended that traffic between the Enterprise and Extranet Zones is monitored and filtered at the zone's perimeter to allow only business approved traffic to enter and leave the zone. Systems in the Extranet Zone are outside of the control of the organization and will typically not abide by the organization's security policies. Therefore, it is important to perform a 3rd party risk assessment before establishing connectivity to understand their security posture and possibly strengthen perimeter defenses (SecureArc, n.d.).

1.1.5. Internal DMZ

The Internal DMZ mediates access between systems in the Enterprise/Extranet Zones and Restricted Zone. Internal application servers typically live in this zone. End-users must authenticate before gaining access to the data hosted in the Restricted Zone.

Luciana Obregon, lucianaobregon@hotmail.com
1.1.6. Restricted Zone

The Restricted Zone houses business-critical systems and large repositories of sensitive information (Government of Canada, 2007). A breach of confidentiality, integrity, or availability of any system in this zone could negatively impact an organization's competitive advantage, reputation, or share price. This zone should have the highest level of protection to ensure that those attacks that are most likely to succeed against the systems within are detected. Systems that live in this zone typically include

- User database servers
- Human Resources database servers
- Financial Database servers
- Intellectual property database servers

1.1.7. Management Zone

The Management Zone houses administration and monitoring systems such as performance servers, configuration management servers, log management servers, jump hosts, and security management systems (SecureArc, n.d.). Systems in this zone are prime targets of attacks given their ability to access any other zone in the enterprise. Users of systems in this zone have higher privileges that allow them to perform their job duties.

The number of Management Zones should be kept to a minimum while satisfying the requirements of all IT staff. For example, there should be one Management Zone for Network Administrators, one for Database Administrators, and a different one for Systems Administrators. Communication among Management Zones should be prohibited unless it addresses a specific business requirement. Additionally, communication between the Management Zones and the Internet should be restricted to only those destinations, ports, and protocols required to download patches or software upgrades.
1.4. Rules of Communication

The network security zone model uses the concept of "trust" as the foundation. Each zone is assigned a trust level. Trust increases from the outer zone to the most inner one that stores the organization's critical data. Communication is only allowed between systems in adjacent zones. Security controls exist between each zone, such as stateful inspection firewalls, intrusion prevention and detection systems, data loss prevention, and rigorous access controls. Security controls implemented inside a zone enable the detection of malicious activity between systems within a zone. (SecureArc, n.d.).

Traffic directionality can also be taken into account when defining the rules of communication among zones. This paper assumes that bi-directional traffic is allowed between adjacent zones. However, some organizations may require that communication between zones only originates from a lower trust zone to a higher trust zone or vice versa.

Multiple zones of the same type can coexist to address different business requirements. For example, there can be two Restricted Zones, one storing financial database servers while the other one human resources database servers. Among zones, traffic is limited to specific resources.
3. Security Event Logging

1.1. Overview

A log is a record of events occurring in a computer system or network that triggers a notification, adding it to a local system file or forwarding it to a centralized log management infrastructure for further processing and analysis (National Institute of Standards and Technology, 2006). It records what is happening in an organization, and it is a prime resource for troubleshooting and supporting business goals.

Log management is the process of generating, gathering, transmitting, storing, analyzing, and disposing of event logs from disparate sources (National Institute of Standards and Technology, 2006).

Security event logging allows organizations to (OWASP, 2015):

- Monitor, identify and report security incidents and policy violations;
- Comply with industry regulatory requirements and internal policies;
- Maintain audit trails to support forensic investigations;
- Monitor system and network performance;

Luciana Obregon, lucianaobregon@hotmail.com
• Maintain records of user activity;
• Troubleshoot problems or unusual conditions.

1.2. What to Log

Most organizations follow a uniform protection approach that treats all enterprise systems as equally important (SANS Institute, 2013). This approach leads to the collection of event logs from all hosts, network devices, and applications across an enterprise network, increasing cost in storage, personnel, and network infrastructure.

Before configuring a system or application to generate vast amounts of irrelevant logs, an organization should analyze the content of each source of log data to determine if collecting the data will enhance visibility into security events and contribute to overall risk reduction.

Logs can be categorized as follows (National Institute of Standards and Technology, 2006):

• Security logs
• Operating system logs
• Application logs

Security logs assist in the detection of malicious and anomalous network activity and support forensic investigation efforts. At a minimum, an organization should be collecting security logs from the following category of systems (National Institute of Standards and Technology, 2006):

- Host-based protection software;
- Intrusion detection and prevention systems (IDS/IPS);
- VPN or remote access systems;
- Web proxy servers;
- Vulnerability management software;
- Authentication servers;
- Routers and layer 3 switches that contain access control lists;
- Firewalls;
Operating system logs assist in the investigation of suspicious activities around a particular system. An audit log records both, successful and failed login attempts, account modifications, file access attempts, use of privileges, and security policy changes (National Institute of Standards and Technology, 2006).

Some applications inherently generate log files and support network protocols such as SYSLOG or SNMP to transfer the logs to a centralized log collector. Other applications use the logging capabilities of the operating system in which they are installed.

1.3. Log Management Architecture

The log management architecture presented in this section consists of three tiers (National Institute of Standards and Technology, 2006):

- Tier I: Log generation
  - Includes the systems, networks, and applications that generate log data.

- Tier II: Log analysis and storage
  - Consists of the log servers, also known as log collectors, which receive the log data from Tier I.

- Tier III: Log monitoring
  - Includes administrative consoles used to monitor and review the log data.

Figure 3 represents a baseline infrastructure for log management. The Management Zone provides protection for log management components against eavesdropping and attacks that could lead to the log data being tampered with or deleted altogether.

The log generating devices in Tier I forward log records to a load balancer in the Management Zone. The load balancer’s function is to distribute the data received from systems in Tier I across the log collectors in Tier II, providing redundancy and fault tolerance. The volume of logged data, log retention period, and availability requirements should be taken into consideration when deciding the number of log collectors to deploy. Each log collector receives a portion of the total data, and together the log collectors hold all the data, making the architecture more flexible, scalable and redundant.
The log monitoring servers in Tier III maintain access to the full set of log data by performing searches across all log collectors in Tier II. Security Event Information Management (SEIM) systems reside in tier III. The logs can be stored locally on the log collector servers, on removable media, or on a storage area network (SAN).

The architecture presented in Figure 4 goes a step further by compartmentalizing the log data per zone. Dedicated log collectors deployed per zone allow the systems in Tier I to forward log data to the corresponding load balancer supporting the zone. This approach can reduce search time and the scope of compliance by segregating regulated data into a dedicated log management sub-system.
4. Network Intrusion Detection and Prevention Systems

1.1. Overview

Intrusion Detection Systems (IDS) inspect network traffic to identify signs of malicious activity and policy violations, enabling organizations to respond before a threat actor causes significant harm to IT systems. Intrusion Prevention Systems (IPS) have the same capabilities of IDS but go a step further by attempting to react to a detected threat to prevent it from being successful (National Institute of Standards and Technology, 2007).

The biggest challenge with IDS and IPS technologies is to place the sensors accurately across the network to protect critical assets and tune them to detect those attacks that are most likely to succeed while reducing the volume of alerts.

1.2. IDS/IPS Infrastructure

The components listed below typically comprise an IDS/IPS infrastructure (National Institute of Standards and Technology, 2007):

- Sensor
  - IDS and IPS sensors that monitor and analyze network activity.
- Management server
  - Centralized hardware or software product that receives information from all the sensors on the network and performs data analysis. It provides a centralized point of access for all security events detected by the sensors, correlates these events and provides reporting capabilities.
- Database server
  - Stores the security events.
- Console
  - Interface used by security analysts to administer the sensors, apply software updates, and monitor and analyze security events.
- Spanning port
A spanning port makes a copy of all the traffic traversing specific switch ports or VLANs and sends it to the sensor. A spanning configuration is the easiest and cheapest way of getting the traffic to the sensors; however, it has many limitations (Cisco, n.d.):

- Switches have a limited number of SPAN ports, typically two.
- Reconfigurations of the spanning port can cause the sensors to stop capturing and monitoring critical traffic.
- Increase data loss due to an oversubscribed spanning port or overloaded switch backplane.
- SPAN port configurations do not guarantee 100% view of network traffic.

- Network TAP

A network tap deployed inline between the sensor and the network itself decreases the risk of dropped packets by using the existing signals to reconstruct the traffic flows (SANS Institute, 2015). A network tap is a scalable solution when multiple monitoring tools are required to capture the same traffic; the monitoring tools can be connected directly to the network tap without impacting network traffic. The tap must fail-open in the event of power loss or malfunction.

- IDS Load balancer

An IDS load balancer is a passive device that aggregates and distributes the traffic received from a spanning port or TAP traffic across multiple sensors (I.J. Intelligent Systems and Applications, 2014). When the output rate of monitored traffic exceeds the throughput of a single sensor, an IDS load balancer can be used to decrease the number of dropped packets and increase visibility.

The components discussed in this section connect to each other through a Management Zone. The sensor's management interface, management servers, database servers, and consoles are all attached to the Management Zone, allowing direct access to authorized users.
1.3. Zone-based Deployment Models

Correct placement of sensors significantly enhances visibility while reducing the number of alerts and, therefore, the efforts to identify and investigate intrusions and violations.

In this section, a zone-based approach is followed to place the sensors at strategic locations that house business-critical assets as well as at junction points where the level of exposure is high, such as an External DMZ or Extranet Zones.

1.1.1. Restricted and Management Zones

These two zones should be carefully watched to detect the most sophisticated types of threats, paying close attention to insider threats. Given the sensitivity of these two zones, both intra-zone and inter-zone traffic should be monitored.

Intra-Zone Traffic Monitoring:

The motivation for intra-zone monitoring is to detect lateral movement between systems inside the zone, fraudulent activities executed by end-users exploiting existing trust relationship between systems, or a worm outbreak.

Deployment Model 1

If the aggregated throughput of the monitored traffic does not exceed the sensor's throughput, an IDS sensor can be directly connected to a spanning port configured in the zone's core switch. The spanning port should mirror traffic from those VLANs that have been designated for critical systems and send a copy of this traffic to the IDS monitoring port. The spanning port can be configured to send a copy of all received traffic, transmitted traffic, or both. If both, received and transmitted traffic is sent to the sensor doubling the throughput and potentially increasing the packet drop rate.
Deployment Model 2

An alternative to deployment model 1 consists in connecting a SPAN TAP to the spanning port of the zone's core switch. The SPAN TAP is a passive device that receives an identical copy of the network traffic traversing the switch and replicates it to the sensors that directly connect to it. Unlike a regular network TAP, the SPAN TAP is not "inline", and unlike an IDS load balancer, each sensor connected to the SPAN TAP receives all the traffic. This model provides flexibility when multiple monitoring tools are required to monitor the same traffic.

Luciana Obregon, lucianaobregon@hotmail.com
Deployment Model 3

If the aggregated throughput of the monitored traffic exceeds the sensor's throughput, an IDS load balancer can be directly connected to a spanning port configured on the zone's core switch. The spanning port should mirror traffic from those VLANs that have been designated for critical systems and send a copy of this traffic to the IDS load balancer. The IDS sensors are directly connected to the IDS load balancer and receive a portion of the total data.

Figure 7 - Intra-zone traffic monitoring using IDS load balancer

Inter-zone Traffic Monitoring:

The motivation for inter-zone monitoring is to detect malicious traffic that may have gotten passed the zone's perimeter firewall. Additionally, the sensor can act as an auditing device to ensure that the firewall's security policies are working as intended.

Deployment Model 4

A network TAP deployed "inline" between the zone's core switch and perimeter firewall monitors traffic entering and leaving the zone. The IDS sensor that connects directly to the TAP's monitoring port receives all the traffic that travels through the TAP. This deployment model eliminates dropped packets, increases visibility, and provides flexibility and scalability to add additional monitoring tools to watch the same type of traffic.

Luciana Obregon, lucianaobregon@hotmail.com
Deployment Model 5

Alternatively, an IPS sensor can be deployed inline between the zone's core switch and the perimeter firewall. Before selecting this deployment model, a rigorous risk assessment should be performed. Deploying an IPS sensor could have the following negative effects on network performance:

- Increased latency could negatively impact those applications that are time sensitive.
- False positives could block legitimate traffic, rendering critical systems unstable or even unusable.
1.1.2. Enterprise Zone

*Deployment Model 6*

Arguably, the most vulnerable point in the Enterprise Zone is where remote users or teleworkers access internal resources through a VPN infrastructure. An IDS sensor should be deployed between the VPN gateway and the zone's core switch so that the IDS sensor can watch unencrypted traffic between remote users and internal resources (Cisco, n.d.). Similar to deployment model 5, an inline IPS sensor can be deployed instead of an IDS sensor.

*Figure 10 – Monitoring unencrypted VPN traffic using IDS*

*Figure 11 – Monitoring unencrypted VPN traffic using IPS*
1.1.3. Extranet Zone, Internal and External DMZ

Traffic should be monitored at the ingress point into the Extranet Zone, Internal DMZ, and External DMZ to ensure that:

• Malicious traffic originated at lower trust zones attempting to reach systems in higher trust zones is detected.

• Command and control traffic beaconing out to external systems is identified and potentially stopped.

• Any large volume of data attempting to leave higher trust zones is detected.

• Auditing measures are in place to certify that firewall policies are properly implemented and functioning as expected.

A combination of the deployment models described above can be used to monitor traffic at the zone’s perimeter.

5. Packet Capture

Full packet capture, when available, is the best resource for incident response activities to analyze and provide evidence of what transpired before, during, and after a breach (SANS Institute, 2015).

Full packet capture allows organizations to troubleshoot network problems, conduct forensic investigations, examine security problems, debug applications and protocol implementations, and learn how protocols behave (Wireshark, n.d.),

Packet sniffing devices capture raw network packets from the network wire and display them in a console or save them for future analysis (Snort, 2015).

Packet sniffer appliances have a minimum of two interfaces, one for management and one for monitoring. Isolating the management interface in a Management Zone ensures that authorized individuals with a "need to know" have access to the captured data.

The organization's security policy should address not only the retention period but also the type of data captured, both of which can be used to determine the amount of storage required to meet the demands as well as where to place the packet sniffers across the enterprise network.

Luciana Obregon, lucianaobregon@hotmail.com
1.1. Infrastructure

1.1.1. Intra-zone Traffic Sniffing:

A spanning port configured on the zone's core switch can be used to send a copy of the traffic traversing the switch to the sniffing device. The spanning port configuration should be identical to the one used for traffic monitoring. The same limitations discussed in section 4.3.1 apply with regards to throughput and dropped packets.

Alternatively, a packet sniffer can be plugged directly into a monitoring port in a SPAN TAP as shown in Figure 13.

![Figure 12 – Intra-zone sniffing using spanning port](image1)

![Figure 13 – Intra-zone sniffing using SPAN TAP](image2)

1.1.2. Inter-zone Traffic Sniffing:

To capture inter-zone traffic, a packet sniffer connected to a monitoring port in the network TAP captures traffic entering and leaving a zone, as shown in Figure 14. A network TAP is the most effective method to capture network traffic because it eliminates dropped packets and has no impact on network performance.
6. Conclusion

This paper presented a systematic approach to architecting network infrastructures with security in mind, making great emphasis in the optimization of security event monitoring. It focused on four network security areas including network segmentation, security event logging, network intrusion detection and prevention, and packet capturing.

The concept of network segmentation and security zones introduced the foundation in the overall defense-in-depth strategy. Without network segmentation, everybody can access every system on the network, which may be very convenient from an end-user standpoint but it gives an attacker a clear advantage and an avenue to compromise critical systems with very little effort.

Security event logging, network-based intrusion detection and prevention, and traffic sniffing were introduced, and a visual representation was presented to demonstrate how these components can be architected together to increase visibility, reduce the volume of alerts, reduce costs, and increase return on investment.
7. References


Luciana Obregon, lucianaobregon@hotmail.com