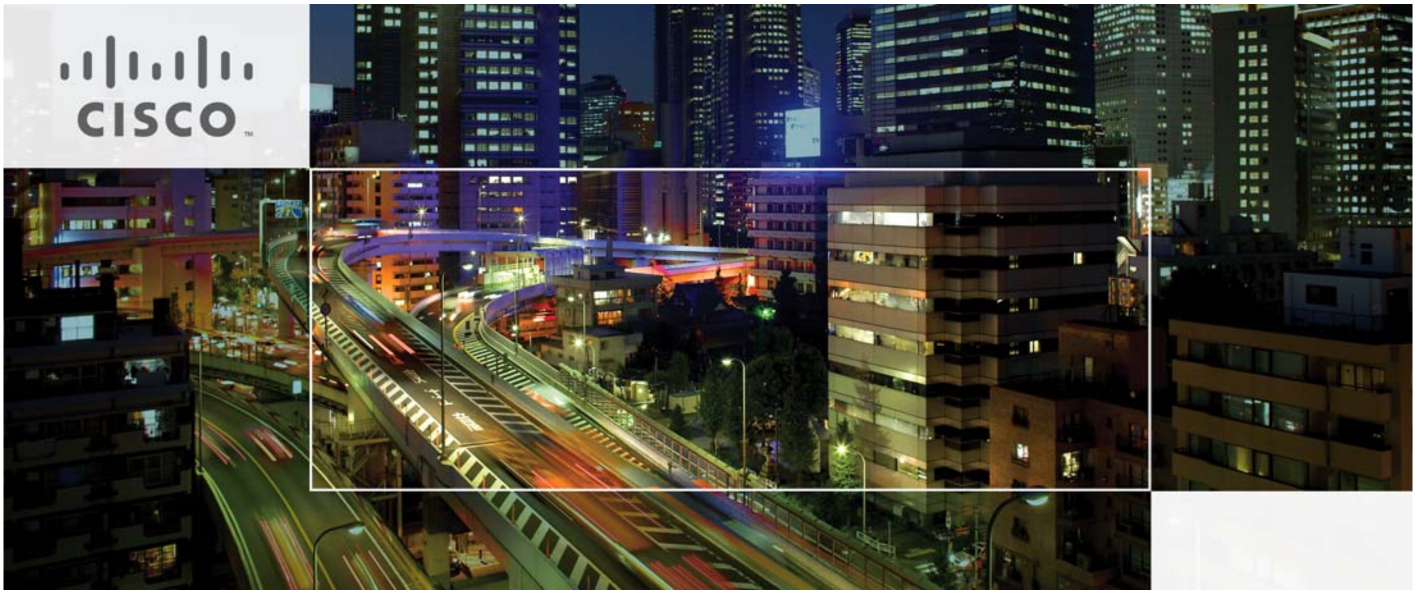


Cisco Cyber Threat Defense v2.0

Design Guide

Last Updated: July 23, 2015

Cisco Validated Design

CONTENTS

Cisco Cyber Threat Defense v2.0 Design Guide

Introduction

Goal of this Document

In recent years, security professionals have become increasingly aware that while traditional defenses remain effective against many common threats, they are vulnerable against more advanced and determined adversaries. In fact, a defense with an efficacy rate of 98 percent or even 99 percent would be considered to be among the very best available—but that still means that one or two percent of all attacks succeed in bypassing the defense, and these include the most dangerous threats. Worse yet, because most defenses typically focus on the network perimeter, they are often blind to dangerous attacks that succeed in entering.

The Cisco Cyber Threat Defense (CTD) solution provides an integrated and validated architecture for defense in depth against these modern, advanced threats. Key to Cisco's approach is the centrality of the network infrastructure to provide visibility and control throughout the enterprise, not just at the front door. Only a holistic, integrated, architectural approach can provide comprehensive coverage over the entire attack lifecycle continuum: before, during, and after a compromise. This design guide introduces a major update, version 2.0, to the Cisco Cyber Threat Defense solution.

Intended Audience

This document is intended for, but not limited to, security architects, system architects, network design engineers, system engineers, field consultants, advanced services specialists, and customers who want to understand how to deploy a robust, distributed security architecture to address today's advanced threats; with the continued flexibility to operate virtualized and physical workloads; and who function in traditional modes or have migrated towards cloud operational models. This document also leverages additional complementary solutions that are documented in separate design and deployment guides. This document assumes that the reader is familiar with the basic concepts of IP protocols, quality of service (QoS), high availability (HA), and security technologies. This document also assumes that the reader is aware of general system requirements and has knowledge of enterprise network and data center architectures.

Corporate Headquarters:
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

Executive Summary

The initial version of the Cisco Cyber Threat Defense solution was introduced in 2013, with a Cisco Validated Design (CVD) to bring together NetFlow telemetry from the Cisco network infrastructure, the Cisco Identity Services Engine (ISE) for user and device identity, and the StealthWatch System through a partnership with Lancope, Inc. to provide network behavior analysis and threat detection in the interior of the network.

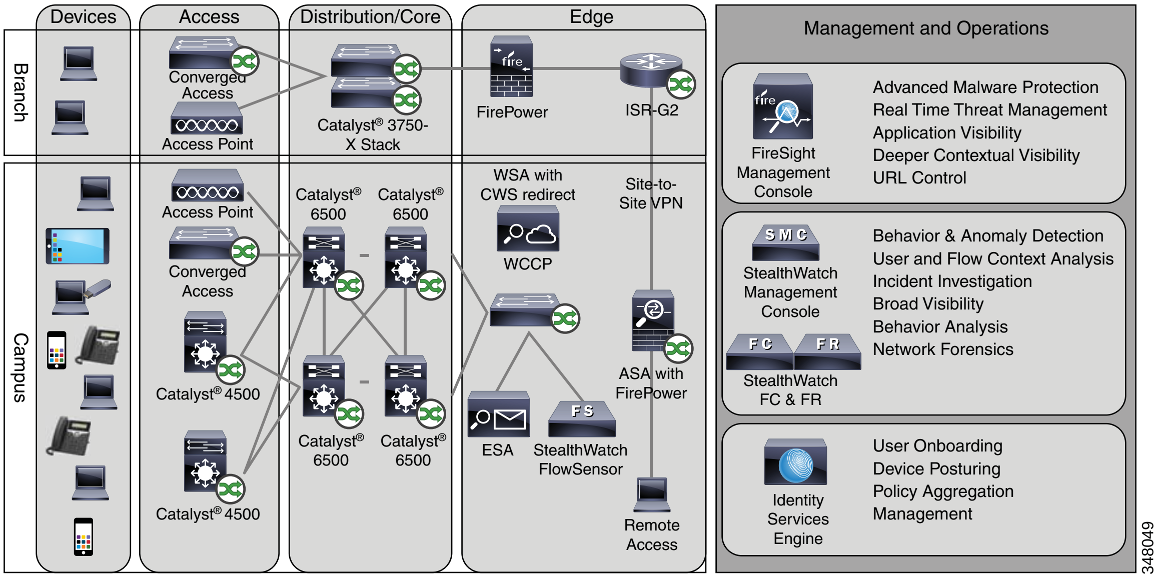
This document introduces a major update and expansion of the Cisco Cyber Threat Defense architecture, building on the network visibility from the previous solution. Version 2.0 incorporates industry-leading Next Generation Intrusion Prevention Systems (NGIPS) and Advanced Malware Protection (AMP) components from Cisco's acquisition of Sourcefire in late 2013, now known as Cisco FirePOWER. The solution includes other elements of the Cisco security portfolio such as email and web content security, and the ability to leverage the network itself as a policy enforcement point through the use of Cisco TrustSec and Endpoint Protection Services in ISE version 1.3.

Solution Overview

Solution Design

Figure 1 shows the high-level architecture of the Cisco Cyber Threat Defense solution, which builds on best practices from other CVDs without going into too much detail about deployment of the network edge. The goal of the Cyber Threat Defense solution is to introduce a design and architecture that can help facilitate the discovery, containment, and remediation of threats once they have penetrated into the network interior. Data center considerations are included in the Cyber Threat Defense solution for the Secure Data Center CVD.

Figure 1 High-Level Architecture of the Cisco Cyber Threat Defense Solution



Cisco Cyber Threat Defense version 2.0 makes use of several solutions to accomplish its objectives:

- NetFlow and the Lancope StealthWatch System
  - Broad visibility
  - User and flow context analysis
  - Network behavior and anomaly detection
  - Incident response and network forensics
- Cisco FirePOWER and FireSIGHT
  - Real-time threat management
  - Deeper contextual visibility for threats bypassing the perimeters
  - URL control
- Advanced Malware Protection (AMP)
  - Endpoint control with AMP for Endpoints
  - Malware control with AMP for networks and content
- Content Security Appliances and Services
  - Cisco Web Security Appliance (WSA) and Cloud Web Security (CWS)
    - Dynamic threat control for web traffic
    - Outbound URL analysis and data transfer controls
    - Detection of suspicious web activity
  - Cisco Email Security Appliance (ESA)
    - Dynamic threat control for email traffic
    - Detection of suspicious email activity
- Cisco Identity Services Engine (ISE)
  - User and device identity integration with Lancope StealthWatch
  - Remediation policy actions using pxGrid

Thinking Beyond the Perimeter

It seems that almost daily there are new reports of cyber-driven data breaches and thefts against large corporations, enterprises, and governments. In many of these events, the enterprise was targeted directly by the attacker and the initial breach happened months before the discovery of the theft. In the past, the industry has (over-) used the term Advanced Persistent Threat (APT) to refer to these types of attacks. Although the term has been overused by the industry to the point that its meaning has almost been lost, it is necessary to revisit the definition:

> *Advanced Persistent Threat (APT): An adversary that possesses sophisticated levels of expertise and significant resources that allow it to create opportunities to achieve its objectives by using multiple attack vectors (e.g., cyber, physical, and deception).* [NIST IR 7298 Rev 2]

Parts of this definition are worthy of special note. The first and most important is that it refers to an *adversary*, meaning an APT is not just a single attack, exploit, or piece of malware, but rather an entire campaign that the determined attacker brings to bear against a targeted entity. The second is that the attacker uses more than one method, so even successfully blocking several attempts at compromise does not necessarily prevent the attacker from trying other methods until one of them succeeds.

Therefore, protecting against an APT involves defending against a determined, human attacker who uses logic and technology to infiltrate the targeted entity and achieve an objective. Defending against such an attacker requires awareness that the attacker will at some point be able to breach the perimeter and gain an operational footprint on the network. The main objective of the Cisco Cyber Threat Defense solution is to instrument the interior of the network to assist the security operator in discovering the presence of attackers that have gained an operational footprint on the network interior.

Indicators of Compromise

The nature of the APT and the modern threat has significantly varied the time between the initiation of the attack and the final execution phases. As a result, a new approach to attack discovery is needed. Previously, most legacy threat systems generated thousands of alerts based on the approach to detect all attack-like activity. A new approach is to leverage indicators of compromise (IOC).

> *Indicator: A sign that an incident may have occurred or may be currently occurring.* [NIST SP 800-61]

By taking the position that the determined adversary or motivated attacker is going to bypass the security perimeter and gain an operational footprint on the network, the IOC approach is to get the correct amount of tooling in place to not only discover the presence of the attacker but also their mode of operation. Elements of such an approach, which would make up a broader, yet more exact, analysis of what was being seen, could include the following questions:

- What is this attack (such as a known type or category)?
- What are the attack specifics, such as how it is/was executed? What may have changed on the target endpoint, and so on?
- Where did the attack originate?
- How was hostility determined?
- What is the target? Host? User?
- What other systems/users has this device contacted?
- What is the targeted application or data?
- Does the target have a chance to be impacted by this event?
- Is this a new issue or was it delivered via an outside source, such as bring-your-own-device (BYOD)?
- Is the attacking host currently in the network or outside the network?
- What was/is the root cause?
- Can the system identify immediately how many hosts or network devices may be vulnerable to this threat?
- If this attack is blocked, how can the system determine whether it is a false positive or true positive?

To achieve an advanced indication of compromise capability, events must be correlated from the following:

- Malware activities
- Intrusion detections
- Network connections
- Network file trajectories
- Device trajectories

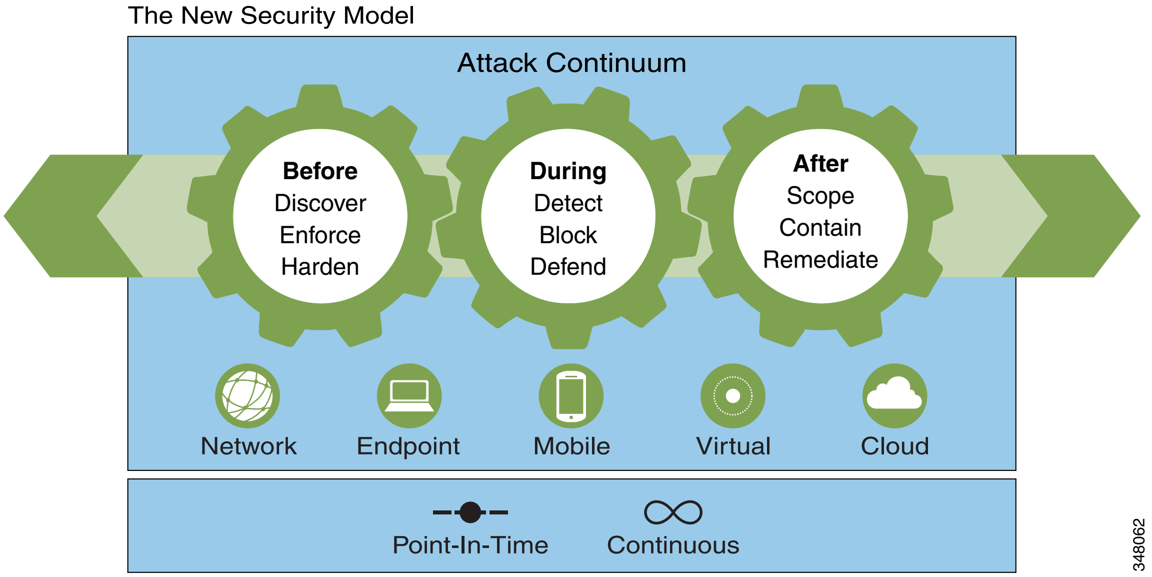
- Device network flows, including but not limited to lateral movements, parent-child relationship, or context

The goal is for all of the above to be correlated with network, endpoint, application, and user context. The resultant data set provides the unique ability to provide IOCs throughout the entire network that are accurate enough to be confidently and immediately actionable.

A Security Model that Leverages Network Integrated Security

Because of the nature of the determined attacker, new tools and technologies are needed to develop a comprehensive response to the threats affecting the enterprise. This must be done using a model that minimizes complexity and helps protect the business assets in a continuous fashion, while addressing the changes in business models, such as any-to-any. The security system should be integrated directly into the network fabric to maximize its efficiency and capability, while minimizing the risk normally associated with adding disparate, non-network-aware security controls. To design such a system, a new model is needed to ensure that this integration can properly take place, especially in the data center where the margin of error is so small. The new model shown in Figure 2 is a useful reference when developing a comprehensive security solution for any type of network. This new model showcases a key component known as the attack continuum, which identifies each of the critical mechanisms and processes that are integral to the complete security system.

Figure 2 New Security Model



This model addresses the threat problem by looking at the actions you must take before, during, and after an attack, as well as across the broad range of attack vectors such as endpoints, mobile devices, data center assets, virtual machines, and even in the cloud. Where most security solutions tend to address the threat at a point in time, it is important to look at it as a continuous cycle.

Before an Attack

Context-aware security is required to defend against context-aware attackers. Organizations are fighting against attackers that have more information about their infrastructures than the defenders trying to protect them. To achieve information superiority over attackers and defend before an attack occurs, organizations need total visibility of their environment including, but not limited to, physical and virtual hosts, operating systems, applications, services, protocols, users, content, and network behavior. Defenders need to understand the risks to their infrastructure, based on target value, legitimacy of an attack, and history. If defenders do not understand what they are trying to protect, they will be unprepared to configure security technologies for defense. Visibility needs to span the entirety of the network, including endpoints, email and web gateways, virtual environments, mobile devices, and the data center. From this visibility, actionable alerts must be generated so that defenders can make informed decisions.

During an Attack

Relentless attacks and blended threats do not occur in a single point of time. They are an ongoing activity and demand continuous security. Traditional security technologies can evaluate an attack only at a point in time, based on a single data point of the attack itself. This approach is no match against advanced attacks.

Instead, what is needed is a security infrastructure based on the concept of awareness; one that can aggregate and correlate data from across the extended network with historical patterns and global attack intelligence to provide context and discriminate between active attacks, exfiltration, and reconnaissance versus simply background activity. This evolves security from an exercise at a point in time to one of continual analysis and decision-making. If a file that was thought to be safe passes through but later demonstrates malicious behavior, organizations can take action. With this real-time insight security, professionals can employ intelligent automation to enforce security policies without manual intervention.

After an Attack

To address the full attack continuum, organizations need retrospective security. Retrospective security is a big data challenge, and a capability that few are able to deliver. With an infrastructure that can continuously gather and analyze data to create security intelligence, security teams can automatically identify IOCs, detect malware that is sophisticated enough to alter its behavior to avoid detection, and then remediate.

Compromises that would have gone undetected for weeks or months can be rapidly identified, scoped, contained, and remediated. This threat-centric security model lets organizations address the full attack continuum across all attack vectors and respond at any time, all the time, and in real time.

Note The Cyber Threat Defense 2.0 Solution's primary focus is on the "During" and "After" stages of the Attack Continuum. Additional security solutions can be found on the Cisco DesignZone website.

Solution Components

NetFlow

NetFlow is a key element of the original version of the Cisco Cyber Threat Defense solution, and continues to play a vital role in this second-generation update.

NetFlow is embedded instrumentation within Cisco IOS software that characterizes network operations by examining connection data. Standardized through the RFC process in the IPFix protocol, variants of NetFlow are available in network equipment from such vendors as Arista, Citrix, Huawei, Juniper, Palo Alto, and a variety of open source Linux operating systems distributions.

NetFlow is a Cisco application that measures IP network traffic attributes of a traffic flow. A flow is identified as a unidirectional stream of packets between a given source and destination as it traverses the Cisco device. NetFlow was initially created to measure network traffic characteristics such as bandwidth, application performance, and utilization. NetFlow has traditionally been used for billing and accounting, network capacity planning, and availability monitoring.

NetFlow is a reporting technology. As a NetFlow-enabled network device processes traffic, the device gathers data about the traffic flow and reports (or exports) the data to a defined collector. Older versions of NetFlow exported data only after the connection closed. Later, NetFlow implementations added the capability of defining one or more expiry timers (active or inactive) or conditions (connection complete or cache full). The nature of NetFlow reporting has tremendous security applications including the ability to provide non-repudiation, anomaly detection, and investigative capabilities.

NetFlow has gone through many versions since it was first introduced, as shown in Table 1. Fixed export format versions (1,5,7,8) are not flexible or adaptable, and each new version contains new export fields that are incompatible with the previous version. NetFlow Version 9 completely separates the collection and export process and allows the customization of the NetFlow collection.

Table 1 NetFlow Versions

Version	Status
1	Original; similar to v5 but without sequence numbers or BGP info
2	Never released
3	Never released
4	Never released
5	Fixed format; most common version in production
6	Never released
7	Similar to v5 but does not include AS interface, TCP flag, and ToS information; specific to Cisco Catalyst 6500 and 7600
8	Choice of 11 aggregation schemes; never gained wide use in the enterprise
9	Flexible, extensible export format to enable support of additional fields and technologies
IPFIX	Similar to v9 but standardized and with variable length fields
Flexible NetFlow	Flexibility and scalability of flow data beyond traditional NetFlow. The ability to identify and classify over 1000 applications via Cisco Application Visibility and control (AVC).

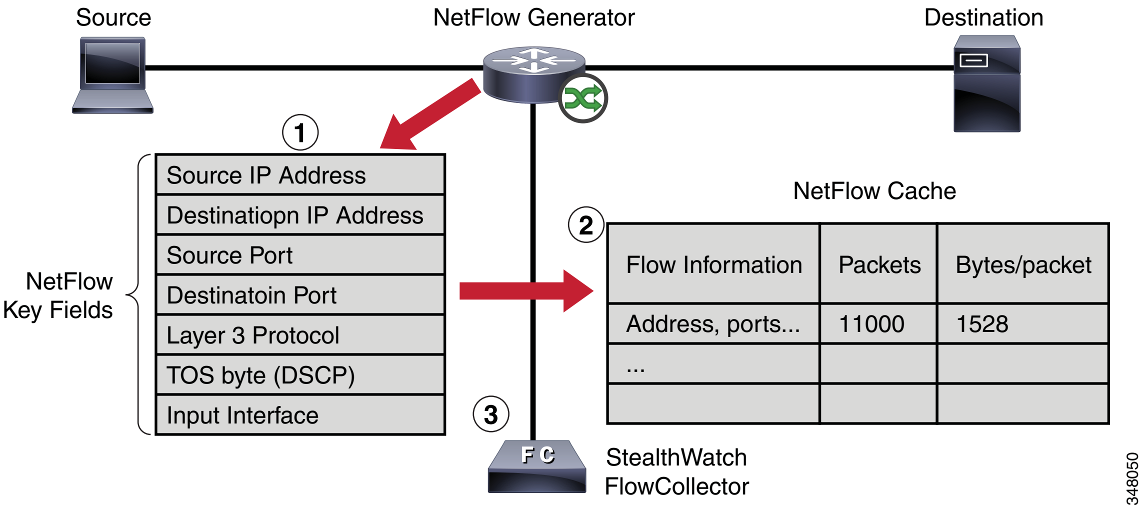
The Cisco Cyber Threat Defense solution takes advantage of the customization capability of the Flexible NetFlow Feature in Cisco IOS, allowing for customizable NetFlow Version 9 records. Using this approach, the CVD for the Cisco Cyber Threat Defense solution has defined NetFlow records for each solution device to maximize the security monitoring potential of each device by collecting packet fields such as TCP flags, time-to-live (TTL) values, protocol, and application name using Next Generation Network-Based Application Recognition (NBAR2) and Cisco AVC. Many of these fields are not available in older versions of the NetFlow protocol. Without these fields, some of the advantages offered by some of the finely tuned detection algorithms used as part of the Cisco Cyber Threat Defense solution would be lost or minimized.

The latest iteration of Cisco-developed NetFlow is Flexible NetFlow. Flexible NetFlow extends NetFlow version 9 capabilities to help customer determine how to optimize resource usage, plan network capacity, and identify the optimal application layer for quality of service (QoS). Flexible NetFlow plays a vital role in network security by detecting denial-of-service (DoS) attacks and network-propagated worms.

Tip **Best Practice:** Use the Cisco IOS Flexible NetFlow feature wherever possible.

Figure 3 illustrates NetFlow operation on a Cisco device.

Figure 3 NetFlow Operation



Flow Information	Packets	Bytes/packet
Address, ports...	11000	1528
...		

1. As data traverses a NetFlow-capable device (NetFlow Generator), the device connection table is mined and NetFlow key data fields are extracted.
2. The key fields are used to identify and correlate new flow data with existing flows and build new entries in the NetFlow cache, a database of flows maintained on the device. In addition to the key fields, the Cisco device collects additional configured collection fields, such as TCP flags, byte counters, and start and end times, and stores this information in the NetFlow cache entry for this flow.
3. When the flow terminates or a timeout event occurs, a NetFlow Protocol Data Unit (PDU), known as a flow record, is generated and exported to a flow collector.

Lancope StealthWatch System

The Lancope StealthWatch system, available through Cisco, is a purpose-built, high-performance network visibility and security intelligence solution. Through the collection, aggregation, and analysis of NetFlow data, along with other contextual data sources such as identity data from Cisco ISE, system-specific data such as syslog and Simple Network Management Protocol (SNMP), and application data via NBAR2 and Cisco AVC, the Lancope StealthWatch system helps security operations staff gain real-time situational awareness of all users, devices, and traffic on the network. The Lancope StealthWatch system also allows security operations staff to quickly and effectively respond to threats before, during, and after a security incident by providing real-time continuous forensics and a view into all network traffic.

The Lancope StealthWatch system, as a component of the Cisco Cyber Threat Defense solution, consists of several individual components, connected as shown in Figure 4 and Table 2.

Figure 4 Lancope StealthWatch System

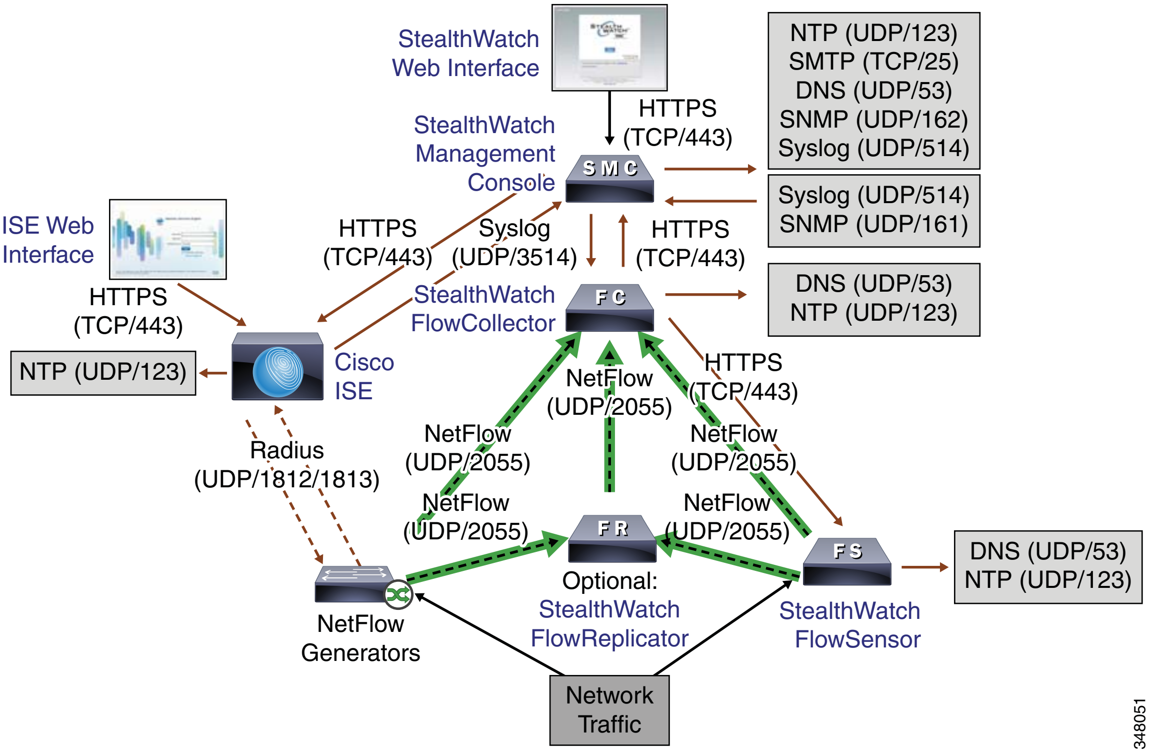


Table 2 Lancope StealthWatch Components

Component	Description
StealthWatch Management Console	Manages, coordinates, and configures all StealthWatch appliances to correlate security and network intelligence across the enterprise. Retrieves authenticated session information from the Cisco ISE to correlate flow and identity.
StealthWatch FlowCollector	Serves as a central collector for flow data generated by NetFlow-enabled devices. The StealthWatch FlowCollector monitors, categorizes, and analyzes network traffic to create comprehensive security intelligence at both the network and host level.
StealthWatch UDP Director (also known as FlowReplicator)	Aggregates NetFlow, syslog, and SNMP information in a single, high-speed appliance. This high-speed UDP packet replicator gathers essential network optimization and security information from multiple locations, and then forwards this information in a single data stream to one or more StealthWatch FlowCollector appliances.
StealthWatch FlowSensor	Passively monitors all host and server communications and network traffic statistics, translating them into flow records, which are sent to FlowCollectors.
StealthWatch FlowSensor VE	A virtual appliance designed to run inside a virtual server. The FlowSensor VE passively monitors intra-VM traffic, translating it into flow records, which are sent to FlowCollectors.

Next Generation Intrusion Prevention System

Cisco's Next-Generation IPS (NGIPS) solution provides a construct of essential yet progressive capabilities for advanced threat prevention. Cisco NGIPS integrates security intelligence through real-time contextual awareness and security automation functions into the network fabric. Cisco NGIPS also leverages contextual information and awareness to provide details such as network activity, operating systems, applications, and users to assess and mitigate threats, provide consistency and process in responses, and reduce an organization's security expenditures. As the evolution of the threat landscape progresses, it is important to not only capture, but to understand attacks and forensics capabilities, in addition to leveraging reporting and high-level analysis and alerting functions. The advent of such understanding and mitigation solutions available in the Cisco NGIPS solution now provides the ability to not merely detect attacks, but also to prevent them.

The evolution of environmental factors such as space, power, operational administration, and efficiency has prompted the emergence of this multifunction security device, which provides deployment flexibility and an ability to understand the dynamics of said environments.

The following attributes are key to the Cisco NGIPS solution:

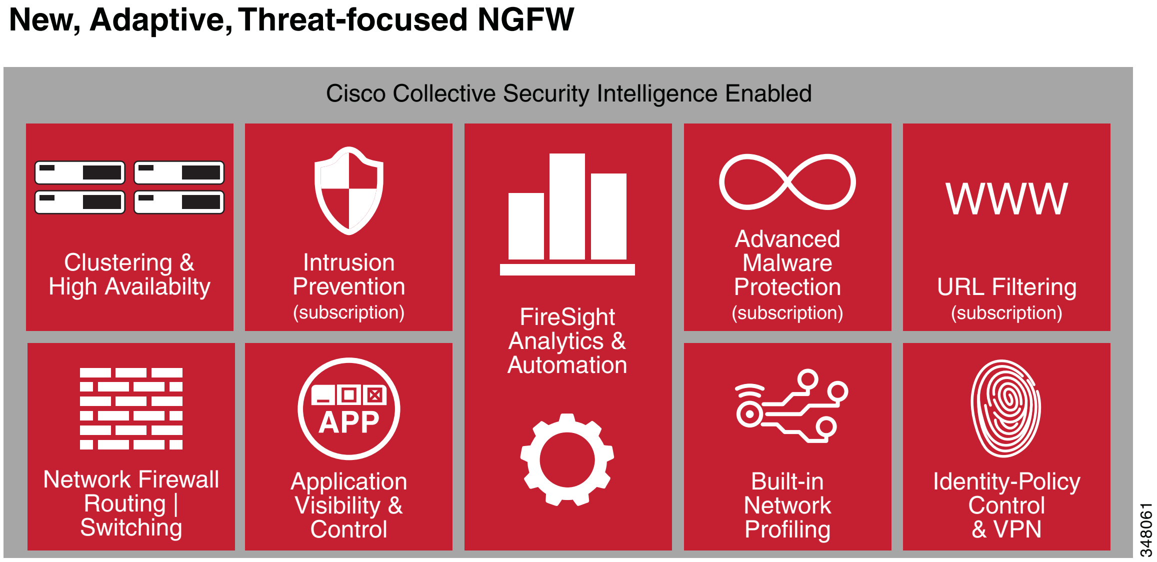
- Ability to identify, monitor, and inspect a wide range of client applications and activities, leveraging Cisco Application, Visibility, and Control (AVC), while enforcing policy against such information.
- Readily available access to various types of contextual data within and external to the environment (for the ability to reference externally available factors/information for correlation) to identify network behavior, user identity, network resources, attack trends and vectors, traffic profiles, and much more.
- Ability to sit in-line, but not disrupt network operations.
- Ability to support vulnerability and threat-centric signatures and vectors.
- Ability to support content awareness and data loss prevention in transit traffic regardless of protocol. This includes URL filtering and the ability to inspect and classify inbound and outbound files and attachments ranging from executable files to PDF and office files.
- Ability to support intelligence gathering and use context awareness, or the ability to bring information in from various sources to make more effective decisions with regard to blocking/filtering and/or warning on transit traffic.

FirePOWER

Cisco ASA with FirePOWER Services is the industry's first adaptive, threat-focused Next-Generation Firewall (NGFW) that delivers integrated threat defense for the entire attack continuum by combining the industry-leading firewall capabilities of the Cisco ASA with industry-leading Cisco FirePOWER threat and advanced malware protection.

The Cisco ASA continues to support stateful inspection, NAT, VPN, routing services, and many others, yet now it integrates Cisco FirePOWER Services in the packet flow path. The FirePOWER Services function handles URL filtering, AVC, threat protection via the NGIPS service module, and Advanced Malware Protection (AMP). Together, the Cisco ASA and FirePOWER Services provide a fully integrated, threat-centric NGIPS/NGFW solution, as illustrated in Figure 5.

Figure 5 NGIPS/NGFW Solution


New, Adaptive, Threat-focused NGFW
Cisco Collective Security Intelligence Enabled
Clustering & High Availabilty
Intrusion Prevention (subscription)
FireSight Analytics & Automation
Advanced Malware Protection (subscription)
WWW
URL Filtering (subscription)
Network Firewall Routing | Switching
APP
Application Visibility & Control
Built-in Network Profiling
Identity-Policy Control & VPN
348061

Cisco Advanced Malware Protection

Cisco AMP allows users to gain continuous visibility and control to defeat malware across the extended network and the full attack continuum: before, during, and after an attack.

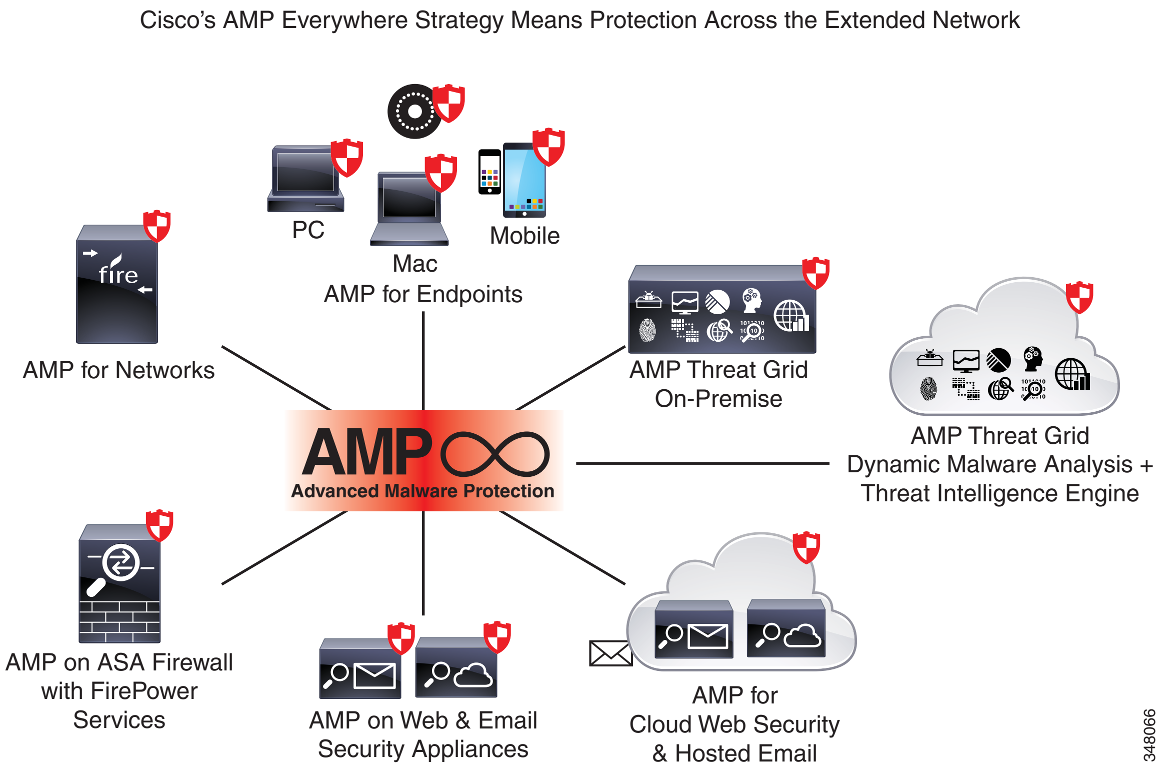
- Before—Prevents known malware, policy-violating file types, and communications from entering your extended network.
- During—Continuously analyzes files and network traffic for threats that evade your first lines of defense
- After—Quickly and efficiently understands, scopes, contains, and remediates an active attack

Cisco AMP provides protection across the broadest range of attack vectors, and can be deployed as:

- A network-based solution, integrated into dedicated Cisco ASA firewall and Cisco FirePOWER network security appliances
- An endpoint solution for PCs, Macs, mobile devices, and virtual environments
- An on-premise, private cloud, virtual appliance built for high-privacy environments
- An integrated feature in Cisco Cloud Web Security or Cisco Web and Email Security Appliances

Figure 6 shows the level of protection provided by Cisco AMP.

Figure 6 Cisco Advanced Malware Protection



Cisco AMP takes full advantage of the vast cloud security intelligence from the Cisco Talos Security Intelligence and Research Group as well as from the Cisco Collective Security Intelligence Ecosystem to deliver advanced protection. Cisco AMP also integrates with Cisco AMP Thread Grid dynamic malware analysis and threat intelligence technology, enhancing its capabilities to aggregate and correlate data to identify advanced and evasive cyber threats.

Cisco AMP provides continuous analysis and retrospective alerting with the following features:

- File Reputation—Analyzes file payloads inline as they traverse the network, providing users with the insights required to automatically block malicious files and apply administrator-defined policies using the existing Cisco Web or Email Security user interface and similar policy reporting frameworks.
- File Sandboxing—Uses a highly secure sandbox environment to analyze and understand the true behavior of unknown files traversing the network. This allows AMP to glean more granular behavior-based details about the file and combine that data with detailed human and machine analysis to identify a file's threat level.
- File Retrospection—Solves the problem of malicious files that have passed through perimeter defenses but are subsequently deemed a threat. All point-in-time detections are less than 100 percent. Rather than operating at a point in time, File Retrospection provides continuous analysis, using real-time updates from AMP's cloud-based intelligence network to stay abreast of changing threat levels. As a result, AMP helps identify and address an attack quickly, before it has a chance to spread.

AMP for Networks

Cisco AMP for Networks can run on any FirePOWER appliance, including FirePOWER Services for ASA. It provides malware detection, both in real time and retrospectively, for common file transfer mechanisms; including HTTP, SMTP, IMAP, FTP, and NBT.

AMP for Content

Email and web access is a required offering that blends both business/corporate use with personal use. There is no way administrators can block access to all non-business use because of the blended nature of today’s Web 2.0 offerings and the mobile use of end user devices. The power of AMP is also available for WSA, ESA, and CWS. By leveraging the full context of mail or web flows, as well as IP reputation and file-based reputation, the content security gateways can enforce and have visibility of blended attacks at the network edge.

AMP for Endpoints

Cisco AMP for Endpoints can be installed on Microsoft Windows, Mac OS X, and Android mobile devices, to provide superior threat defense on the endpoints themselves.

FireSIGHT Management Center

The FireSIGHT Management Center (FMC) provides visibility into network-connected devices combining threat information from FirePOWER and AMP with contextual information about devices including physical and virtual hosts, operating systems, applications, services, protocols, users, geolocation information, content, and network behavior. In addition, FireSIGHT provides access to intelligence sources and information and the ability to incorporate automated functions to increase operational efficiency by providing the ability to correlate events and intelligence data to make informed decisions for your environment. As the management console and database repository for the NGIPS and AMP solutions, the FMC provides a centralized point of event and policy management for security operations. It also allows the ability to automatically aggregate and correlate the various data generated by Cisco ASAs with FirePOWER Services and Cisco FirePOWER physical and virtual appliances deployed throughout the network.

The central management capabilities of FireSIGHT allow it to centrally manage all network security and operational functions, including event monitoring, analysis, incident prioritization, and reporting.

Content Security Controls

With today’s mobile characteristic of end devices as well as the blended use of said devices for both business and personal use, Email and web content security are critical components in the threat defense architecture. Different than other applications and devices in security, where it is applicable to lock down what devices/users can communicate with specific end points, content security has to allow a level of ubiquitous access to different servers and users. The criminal ecosystem depends upon this level of access. While many papers and guides have been written and disseminated around the value of content security on inbound flows, this section focuses on outbound.

Web Security

Advanced attackers are taking advantage of new attack vectors such as mobile devices, web-enabled and mobile applications, and web browsers to accomplish their mission. In this new environment, anyone within your organization can be attacked at any time and any location. The following are two reasons why:

- The web is a popular attack vector for criminals; attackers are organized, and web-born tactics are insidious. Watering hole attacks conceal malware on member-based sites, phishing scams target individuals with personal details, and botnets take control of victims' devices. It is not a small number of sites that pose a threat: 93 percent of customer networks access websites that host malware, according to the Cisco Annual Security Report. Web-based attacks are ever changing, harder to detect, and more damaging than ever.
- Without proper control, your own users put your business at risk. Your branch offices, individual employees, and even guest users can consume excess bandwidth, throttling software-as-a-service (SaaS) application use and other high-priority business functions. They may also access content outside of your acceptable use policies such as social media, Internet videos, and personal SaaS applications, creating a shadow IT infrastructure outside the governance of corporate IT and outside the protection of traditional security solutions.

Security is not just about building bigger walls and adding more one-off solutions. To stop criminals and control usage, you need a solution that fits your infrastructure, grows with your business, and immediately adapts to changing threats. It must provide the most up-to-date malware defense and offer tools for managing a breach when it occurs.

The following are key components of the Web Security portfolio deployed in Cisco CTD version 2.0:

- Advanced malware protection of inbound flows and tracking of threats specific to sandboxing and retrospection. This follows the same description of the other AMP components, but currently AMP scans for outbound flows only. Where this makes sense for CTD is the discovery of possibly infected clients after retrospection. A clean list of all clients that have interacted with malicious code and where the content has come from provides details on patient zero as well as the source of the infection.
- Signature-based security. Anti-virus scanning for outbound flows easily allows administrators to discover infections and the possibility of their user base being used as part of a distributed attack.
- Behavioral analysis of outbound flows to C&C sites can be easily discovered with L4TM. By scanning for all outbound TCP and UDP flows, the WSA can monitor and block these communications to C&C servers. Combined with other aspects of Cisco CTD version 2.0, this critical data can assist in the discovery of IOCs.

Email Security

The email threat landscape contains increasingly sophisticated advanced threats and targeted attacks. Mass spam campaigns and unsafe email attachments are no longer the primary security concerns. By scouring social media websites, attackers now find information on intended victims and contrive spear phishing emails. These attacks use personal information and social engineering tactics, often tied to global news events, to deceive users with malicious links serving up malware. There are more opportunities for attacks than ever before. Employees once checked text-based email from a workstation behind a company firewall, but today they interact with rich HTML messages from multiple devices, both anytime and anywhere. Ubiquitous access creates new network entry points that blur the lines of historically segmented security layers.

Cisco Email Security (CES) is the leader in this evolved threat landscape. Cisco has the highest capture rate and the lowest false positive rate available for spam. Cisco also has the industry's only proven zero-hour antivirus solution, providing protection from brand-new viruses in less than 60 minutes. This is couples with Cisco's more traditional scanning engine called Defense in Depth for spam and virus. The same solution also scans your outbound email to help comply with data loss prevention compliance requirements. CES can automatically scan your outbound email for social security numbers, credit card numbers, patient health information, and more. It automatically encrypts those messages so the recipient has to authenticate to open and view, providing immediate compliance. Cisco also offers the best performance and the lowest total cost of ownership. Cisco, with its complete security architecture, is focused on the future: all of Cisco's security technologies work together to enhance security for your company.

The following are key components of the Email Security portfolio deployed in Cisco CTD version 2.0:

- Advanced malware protection of inbound flows and tracking of threats specific to sandboxing and retrospection.
- Signature-based security, including anti-spam and anti-virus scanning for outbound flows.
- Behavioral analysis of outbound flows to C&C sites or flows as part of an outbound spam attack. Monitoring the outbound flows against a baseline of what is normal or expected can help administrators quickly understand where infections have occurred and can help clean up.

TrustSec and the Identity Services Engine

Cisco ISE is the market-leading security policy management platform that unifies and automates access control to proactively enforce role-based access to enterprise networks and resources, regardless of how a user chooses to connect—wired, wireless, or VPN.

Cisco ISE uses Cisco Platform Exchange Grid (pxGrid) technology to share rich contextual data with other Cisco platforms as well as integrated partner ecosystem solutions. This makes it easier than ever to add features to identify, mitigate, and remediate security threats across the network. Overall, secure access control is centralized and simplified to securely deliver vital business services, enhance infrastructure security, enforce compliance, and streamline service operations.

Cisco TrustSec simplifies the provisioning and management of secure access to network services and applications. Compared to access control mechanisms that are based on network topology, Cisco TrustSec defines policies using logical policy groupings, so secure access is consistently maintained even as resources are moved in mobile and virtualized networks. De-coupling access entitlements from IP addresses and VLANs simplifies security policy maintenance tasks, lowers operational costs, and allows common access policies to be applied to wired, wireless, and VPN access consistently.

Cisco TrustSec classification and policy enforcement functions are embedded in Cisco switching, routing, wireless LAN, and firewall products. By classifying traffic based on the contextual identity of the endpoint versus its IP address, Cisco TrustSec enables more flexible access controls for dynamic networking environments and data centers.

At the point of network access, a Cisco TrustSec policy group called a Security Group Tag (SGT) is assigned to an endpoint, typically based on that endpoint's user, device, and location attributes. The SGT denotes the endpoint's access entitlements, and all traffic from the endpoint carry the SGT information. The SGT is used by switches, routers, and firewalls to make forwarding decisions. Because SGT assignments can denote business roles and functions, Cisco TrustSec controls can be defined in terms of business needs and not underlying networking detail.

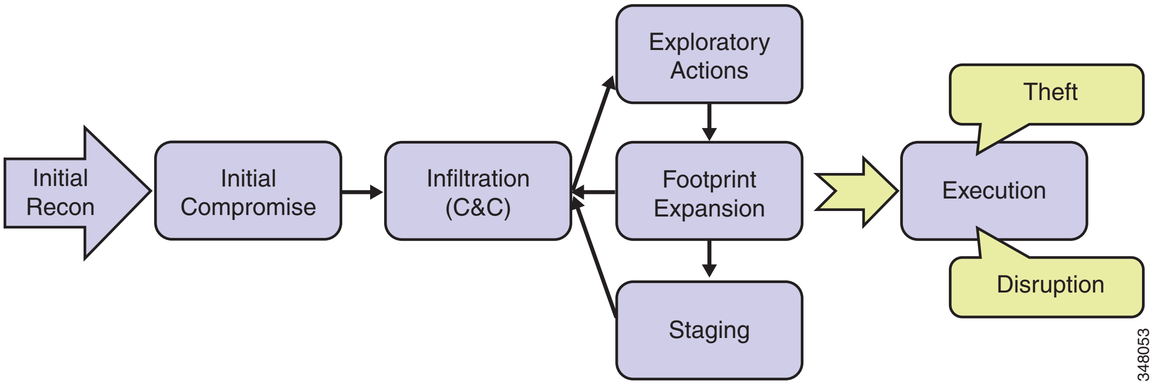
Operating under the Presumption of Breach

Despite all the best security controls a determined, motivated attacker can still gain an operational footprint on the network to execute towards their objective. The attacker holds many advantages because they know their objective, whereas the defender must be able to discover the attacker's presence, a non-trivial task. This section first dissects the attack lifecycle before describing how the defender can operationalize data out of the components of the Cyber Threat Defense solution to discover the presence of an attacker and protect critical assets.

Dissection of the Attack Lifecycle

The attack lifecycle focuses primarily on the steps that an attacker takes from the very beginning until they are able to execute against their objective. All of the steps have both technical and non-technical means of executions, and some of the initial steps can happen entirely outside of the visibility of the victim organization. Each of the steps is shown in Figure 7.

Figure 7 Attack Lifecycle



Initial Reconnaissance

In this step, the attacker gathers information about the target organization. This might consist of leveraging public information, using social media to find and target employees, determining what technologies are in use at the target organization, and general preparation for the initial attack.

Initial Compromise

In this step, the attacker gains an operational footprint inside the victim organization. The compromise might use either known malware or a zero-day, custom-crafted piece of malware, or it might be accomplished using non-technical methods such as social engineering. In any case, at the completion of this step, the attacker has successfully infiltrated the organization.

Infiltration

At the completion of the previous step, the attacker has successfully infiltrated the organization. This step is about being able to maintain that point of presence. For example, if in the previous step the attacker has been able to successfully retrieve a username and password, in this step the attacker opens a remote connection to the organization. If the initial compromise has been the installation of a piece of malware on a personal computer belonging to an employee, this step is the opening up of a command-and-control (C&C) channel with that piece of malware.

Exploratory Actions

At this point the attacker begins to locate resources inside the organization that are relevant to their objective. What the attacker does here depends on the objective and the method in which they were able to infiltrate the organization. Often, the attacker begins scanning the network interior to find other resources that might be vulnerable or hosting data the attacker is interested in exploiting. In this step, the attacker might also seek to recover legitimate credentials such as usernames and passwords.

Footprint Expansion

Until this point, the attacker still has only one infiltration point into the organization; for example, a single malware-infected machine or a single set of stolen credentials. In this step, the attacker expands their point of presence to include multiple infiltration points and/or multiple points of presence inside the organization. The return to the Infiltration stage in Figure 7 represents the remote connection to the new resource.

Staging

In this step, the attacker prepares for the final execution phase. What happens here depends on the attackers' objective. If the goal is to obtain or steal data, this step might be the slow collection of data from different resources inside of the organization. If the objective is disruption, this step might be ensuring that they have an operational footprint on all necessary targets. The return to the Infiltration stage in Figure 7 represents the remote connection or credentials the attacker is using as their Infiltration point.

Execution

This is the final step of the attack, in which the attacker has decided that their mission is complete and their goals and objectives have been met. Attacker objectives tend to fall into one of two categories: theft of data or disruption of activities.

Building Resiliency into the Network

Understanding the steps of the attack lifecycle is the first step in being able to build resiliency into the network. The goal is to have the right instrumentation in place to quickly identify the presence of the attacker and manage the attack. This section explores the various components of the Cisco Cyber Threat Defense solution and how they are used to identify and manage the different steps of the attack lifecycle.

Blocking Known Attacks using NGIPS

The first step in building resiliency into the network is to deploy the correct mix of technologies to control access to the network. Accomplishing this requires adhering to other Cisco best practices and validated designs to construct secure access controls and network edges, both for the campus and data center.

The Cisco Cyber Threat Defense solution recommends following the best practices and principles contained within the following CVDs, found on www.cisco.com/go/designzone:

- Cisco Threat Management with Next-Generation IPS
- Firewall and IPS Technology Design Guide

Note Note that some design guides may not yet have been updated to use FirePOWER in place of the legacy Cisco IPS.

- Cisco TrustSec 2.0 Design and Implementation Guide

Adherence to the above creates a network that has sufficient access controls and threat management solutions in place to limit the attack surface such that it is more difficult for an attacker to successfully launch known exploits against an organization. The following is achieved:

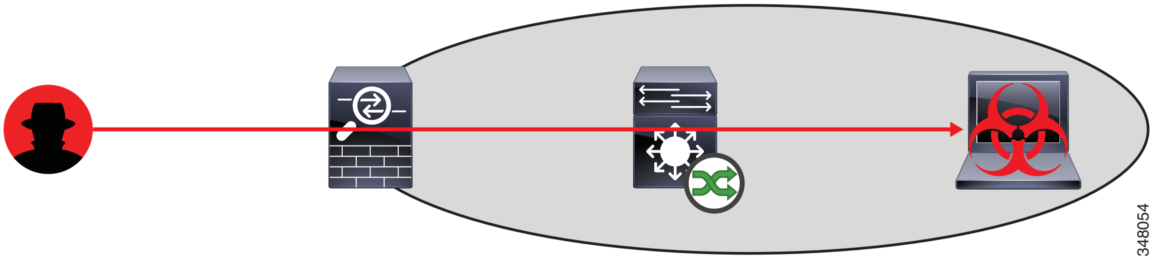
- NGIPS blocks known malware
- ESA blocks known Simple Mail Transfer Protocol (SMTP)-borne malware
- WSA blocks HTTP-borne malware
- ISE performs posture checking and policy enforcement for devices joining the network

It is also assumed for this guide that the network and access edge is in place and operational to facilitate telemetry to support other use cases for the Cisco Cyber Threat Defense solution.

Detecting Command-and-Control

Operationally, an attacker's C&C channel is a communication channel from a host on the network interior (an inside trusted and authorized zone) to a host on the outside. This channel can exist in many forms, such as a communication channel from a malware-infected host to a C&C server or a VPN connection using stolen credentials. Figure 8 illustrates conceptually what is meant by a C&C channel: a Layer 3 or 4 connection, between a resource located on the network interior and the attacker's control point outside of the network.

Figure 8 C & C Channel



To detect and defend against C&C channels and to remediate the attacker's infiltration point, it is first necessary to examine the channel itself, understand the various ways it can manifest itself, and what can be done to detect the channel from a network and protocol perspective. The following is a partial list of sample items to analyze when detecting C&C channels:

- Countries
- Applications
- Uploads/downloads ratio
- Time of day
- Repeated connections

- Beaconing (repeated dead connections)
- Long-lived flows
- Known C&C servers
- Web requests/URLs
- Suspicious user login activity

The technologies in place as part of the Cisco Cyber Threat Defense solution help to automate some of the items above, provide behavioral or anomalistic detection of C&C channels, and can also provide visibility to assist a security operator in manually discovering C&C activity in their environment. The following sections describe Cisco Cyber Threat Defense solution technologies and how they help detect and possibly remediate C&C activity.

Lancope StealthWatch

Leveraging NetFlow data from network and perimeter devices, visibility into all traffic flows entering and leaving the network can be obtained. Several behavior- and anomaly-driven algorithms analyze network traffic to identify covert channels based on suspicious traffic flows. The StealthWatch Labs Information Center (SLIC) threat feed also provides a list of known C&C servers to the StealthWatch System to generate alarms when there is communication detected to these servers. Note that known C&C servers can be either an IP address or a URL.

Additionally, a record of all communication is maintained, allowing a forensic look-back to identify covert channels and compromised devices when new information and IOCs are discovered by security operators.

Cisco Web Security Appliance

The Cisco WSA can provide automated, inline blocking of HTTP and HTTPS traffic destined to low reputation web servers through URL inspection as the traffic passes through the appliance. Many times, low reputation web servers can be known botnet C&C servers.

Cisco FirePOWER

Cisco FirePOWER can provide automated, inline blocking of traffic destined to low reputation web servers through URL inspection as it passes through a NGIPS sensor. Many times, low reputation web servers can be known botnet C&C servers.

Cisco Cloud Web Security (CWS) with Cognitive Threat Analytics (CTA)

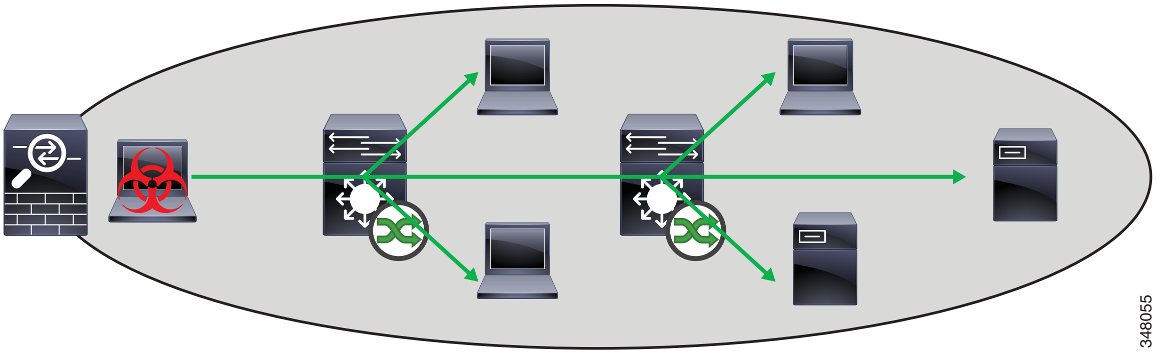
The CTA service in Cisco CWS Premium provides analysis and anomalistic detection of C&C channels through the monitoring of URL activity. This service helps identify C&C servers that might be previously undiscovered or are being brought to bear in a targeted manner against a single organization.

Defending Against Internal Reconnaissance

> *Reconnaissance: Research, identification and selection of targets, often represented as crawling Internet websites such as conference proceedings and mailing lists for email addresses, social relationships, or information on specific technologies.* [Lockheed Martin, “Intelligence-Driven Computer Network Defense”]

From the viewpoint of the network interior, an attacker’s activity on these actions manifest themselves as attempts to identify other resources of interest through the use of Layer 3 and 4 protocols. Figure 9 conceptually illustrates the idea behind internal reconnaissance.

Figure 9 Internal Reconnaissance



Examples of activities include sending Internet Control Message Protocol (ICMP) echo-request messages to random internal IP addresses, and broad scanning of network address blocks to identify devices with open services or ports. While these reconnaissance techniques can be quite noisy, and the more clever and targeted attackers are operating in a more "low and slow" manner, in general the methodology is the same: using the existing point of presence, the attacker attempts to locate other resources on the network using network protocols.

The following is a partial list of sample items that can be analyzed to identify reconnaissance activity on the network interior:

- High number of flows
- High client byte ratio
- One-way or unanswered flows
- Flows within the subnet/logical group
- Flows to non-existent IPs
- Flow patterns
- Abnormal behavior such as connection attempts to a secure data center
- Suspicious flow types such as ICMP

The technologies in place as part of the Cisco Cyber Threat Defense solution help to automate some of the items above, provide behavioral or anomalistic detection of internal reconnaissance activity, and provide visibility to assist a security operator in manually discovering internal activity in their environment. The following sections describe Cisco Cyber Threat Defense solution technologies and how they help prevent, detect, and possibly remediate internal reconnaissance activity.

Lancope StealthWatch

As discussed previously, by leveraging NetFlow data from network devices throughout all layers of the network—access, distribution, core, and edge—Lancope StealthWatch provides complete visibility into all traffic flows in the network. This visibility allows a metadata record to be maintained of every communication that traversed a network device. This aggregated data can be analyzed to identify hosts with suspicious patterns of activity. Lancope StealthWatch has a specific "Reconnaissance" alarm category with several different algorithms watching behavior and identifying suspicious activity. In addition, Lancope StealthWatch supports historical look-back of retrospective activities to allow the operator to investigate a host following the discovery of an IOC in another system.

Cisco NGIPS

The Cisco NGIPS can be used to detect and block specific applications used in scanning activity at the major segmentation points where it is deployed. For example, the Cisco NGIPS can be used to block ICMP messages between the campus network and the data center.

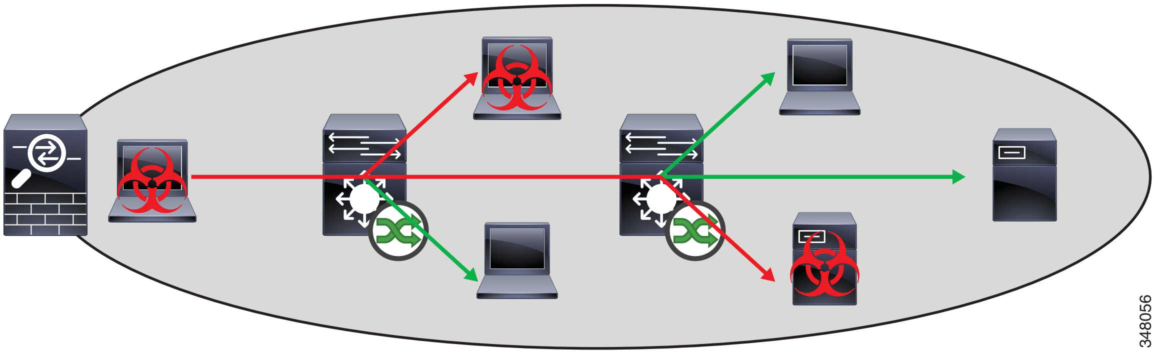
Cisco TrustSec

Leveraging policy, security group-driven segmentation can limit the effectiveness of internal reconnaissance activity as well as reveal activity. By implementing policy limiting peer-to-peer network traffic between hosts in the same security group to business-critical applications, the network can actively prevent network-level reconnaissance activity from occurring.

Defending Against Internal APT Propagation

Once additional or target resources are identified, an attack can choose to leverage malware to spread their operational footprint to the target. Figure 10 illustrates this spread conceptually; operationally, malware spread appears as a Layer 4 communication between two hosts where a volume of data is being sent to the target host. The target host might then begin to exhibit suspicious activity, such as network reconnaissance and even malware propagation.

Figure 10 Internal Propagation



Examples of malware propagation include:

- Self-propagating malware strains that scan for hosts susceptible to a specific vulnerability and then exploit that vulnerability
- An attacker with legitimate credentials that connects to a remote resource and then transfers and runs an executable

Defending against internal malware propagation requires a combination of analyzing network traffic and file analysis and tracking. The following is a partial list of sample items that can be analyzed to identify internal malware propagation:

- High number of flows
- High client byte ratio
- Connections within the subnet/host group
- Flow patterns
- Abnormal behavior
- Known vulnerabilities (rule matching)

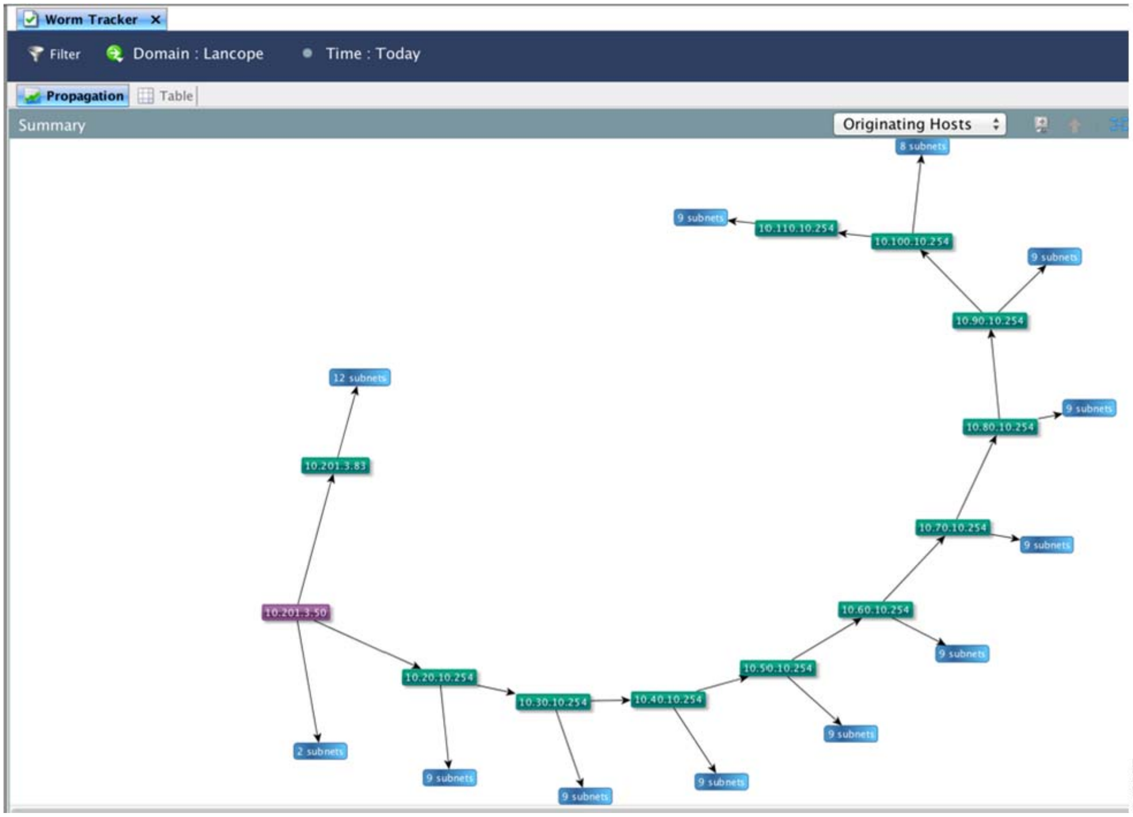
- File trajectory and movements
- Non-adherence to protocol
- Tunneled flows
- Executable analysis

The technologies in place as part of the Cisco Cyber Threat Defense solution help to automate some of the items above, provide behavioral or anomalistic detection of malware propagation, and provide visibility to assist a security operator in manually discovering malware activity in their environment. The following sections describe Cisco Cyber Threat Defense solution technologies and how they help prevent, detect, and possibly remediate internal malware propagation.

Lancope StealthWatch

As discussed in previous sections, Lancope StealthWatch provides complete visibility into all traffic flows in the network, providing a metadata record of every communication that traverses a network device. This aggregated data can be analyzed to identify hosts with suspicious patterns of activity. Lancope StealthWatch has a specific alarm category for malware spread with several different algorithms watching behavior and identifying suspicious activity. There is also a feature known as the worm tracker (shown in Figure 11) that tracks the spread of malware through its suspicious activity across the network. In addition, Lancope StealthWatch supports historical look-back of retrospective activities to allow the operator to investigate a host following the discovery of an IOC in another system.

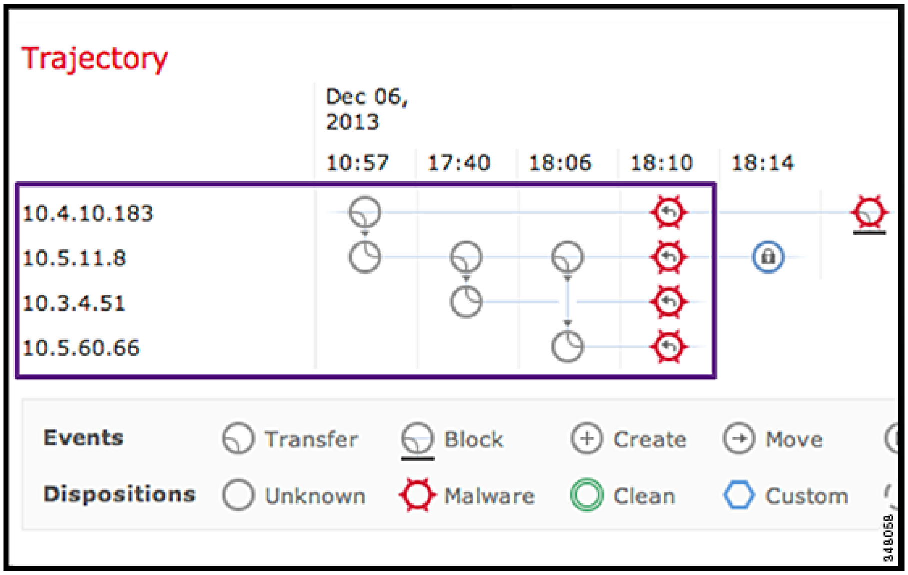
Figure 11 Tracking Suspicious Behavior with StealthWatch



Advanced Malware Protection

Cisco AMP for Endpoints provides file-matching analysis to identify suspicious files when they are transferred onto an endpoint. AMP can provide automated blocking of suspicious files as well as the ability to track the spread of a file throughout the network using a feature known as File Trajectory, as shown in Figure 12.

Figure 12 File Trajectory



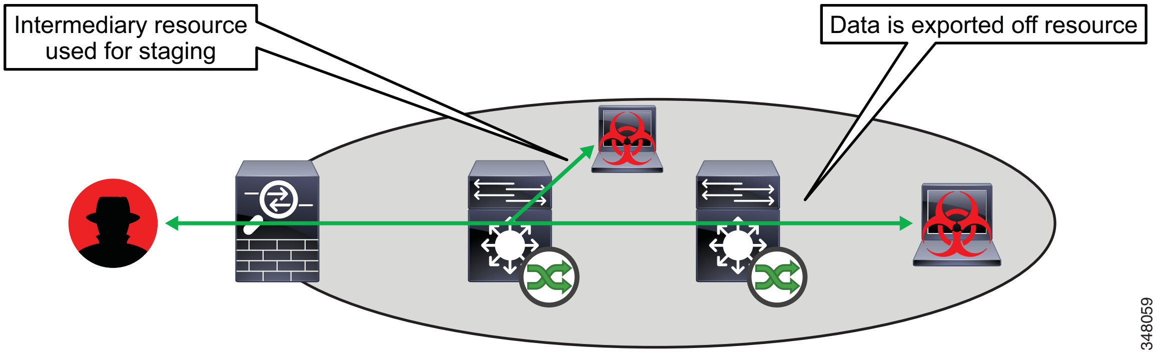
Cisco TrustSec

Leveraging policy, security group-driven segmentation can provide network-level protection against possible lateral movement of an attack. Policy can be implemented to limit network-level communication between user devices without impacting user-to-server traffic.

Defending Against Data Loss or Exfiltration

Many times the attacker's final objective is the theft of data from the target organization. This data can come in many forms such as trade or state secrets, intellectual property, or customer information, including customer credit card data. Operationally, data loss can be viewed as a transfer of data using a transport or application layer protocol from the inside network to the outside network. An intermediary step, known as data staging or data hoarding, can also occur in which the attacker uses an intermediary device to temporarily store the data before exfiltrating it, as shown in Figure 13.

Figure 13 Data Exfiltration


Intermediary resource used for staging
Data is exported off resource
348059

Defending against data loss is a significant challenge for most organizations because of the disparity of data (the difficulty in recognizing the data of value) as well as the multiple possible attack vectors to reach that data, possibly because of a geographically diverse attack surface such as a large retail company with multiple point-of-sale terminals. As the final step in the attack lifecycle, it is important to note that many times this detection comes after some data has already been exfiltrated, so not only is it important to ensure systems are operational to detect an attack before a data loss event has occurred, but also to have the right technology in place to retroactively analyze the theft. As a result, defending against data loss is a multiple-step challenge and may require multiple, different pieces of technology to discover the data loss event.

The following is a partial list of example items that can be analyzed to identify data loss:

- Historical data transfer behavior
- Applications
- Time of day
- Countries
- Amount of data (single and in aggregate)
- Time frames
- Asymmetric traffic patterns
- Traffic between functional groups

The technologies in place as part of the Cisco Cyber Threat Defense solution help to automate some of the items above, provide behavioral or anomalistic detection of data theft activity, and provide visibility to assist a security operator in manually discovering data theft activity in their environment. The following section describes how Lancope StealthWatch helps prevent, detect, and possibly remediate data exfiltration.

Lancope StealthWatch

The complete visibility provided Lancope StealthWatch, as previously described, allows a metadata record to be maintained of every communication that traverses a network device, allowing a forensic audit trail to exist in the event of a data exfiltration event. Lancope StealthWatch also automates the analysis of many of the above considerations and has behavior and anomaly alarm categories for both data hoarding and data exfiltration events. The creation of host groups supports the ability to monitor the movement of data to other parts of the network (as is discussed in the next section), supporting the ability to proactively monitor when data is being leaked from a critical server to a suspicious location. In addition, Lancope StealthWatch supports historical look-back of retrospective activities to allow the operator to investigate a host following the discovery of an IOC in another system.

Continuous Monitoring of Critical Assets

Not all network-connected devices have the same value. Some devices are critical to business processes and/or operations, or may house sensitive data. These devices, hosts, and users require special attention when monitoring their use and potentially misuse. The essential activity of monitoring the network activity of these critical assets helps identify any suspicious or malicious activity and accelerates the incident response and containment process, minimizing the risk exposure of an organization.

Every organization has different priorities for their “crown jewels” and what constitutes a critical asset. The approach to identifying these assets can be a judgment based on the risk and perceived value of the targets and their information.

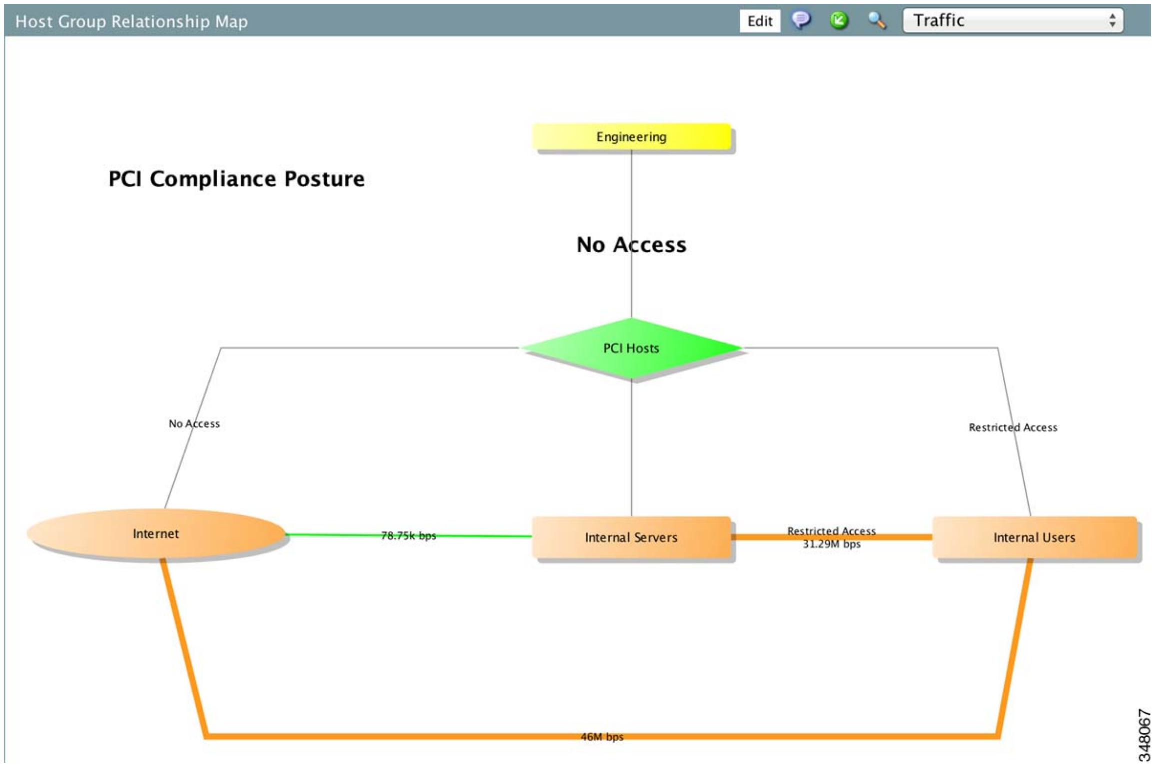
The following is a list of examples that can help in identifying the “crown jewels”:

- Business impact analysis—Systems that are directly related to the operation of the business
- Revenue impact—Systems that drive orders and delivery
- Expenses impact—Systems that manage contractual obligations
- Legal requirements—Statutes and contracts that highlight which information and systems to protect (for example, compliance requirements)
- Sensitivity profile—Systems that access privileged or restricted information
- Risk profile—Systems that because of their nature are underprotected; for example, legacy systems
- Visibility profile—Highly visible systems that can prove potentially embarrassing to the organization if attacked

The technologies that comprise the Cisco Cyber Threat Defense solution can provide active monitoring of all critical assets. The broad visibility provided by NetFlow and Lancope StealthWatch allows an audit trail of all end-to-end communication of network-connected devices. The deep visibility provided by the Cisco NGIPS and FMC allows deep visibility and heightened relevance into any malicious files that are being used by or against critical assets.

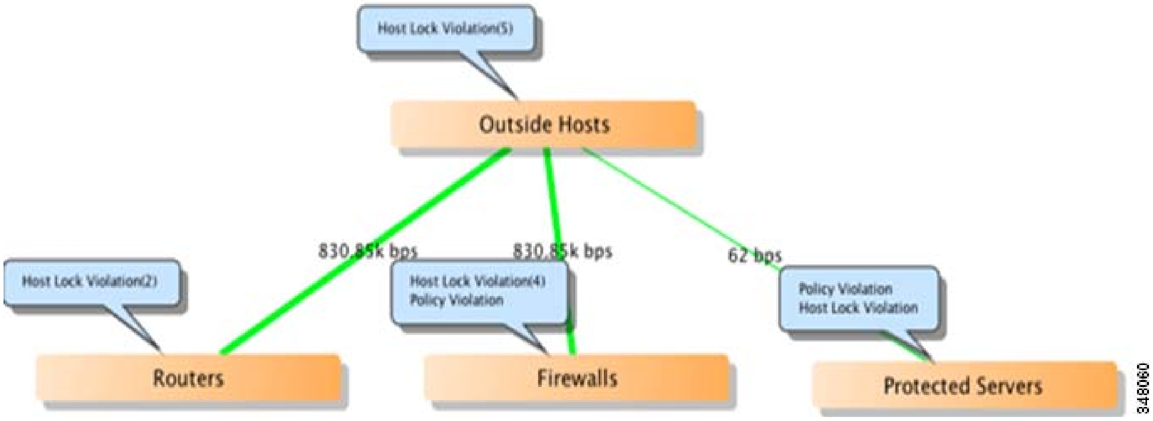
Figure 14 illustrates the use of host groups and the custom map features of Lancope StealthWatch to actively monitor the zones of the organization relevant to Payment Card Industry (PCI) compliance regulations. In this example, host groups for each major component of the PCI zones are defined, relationship properties and policies between the zones defined and mapped, and custom rules developed to alert when traffic is seen on the network that violate the policy.

Figure 14 Segmentation Monitoring with NetFlow and Lancope StealthWatch



Because the network infrastructure itself can also be attacked, active policy monitoring can also be put in place to monitor the communications of the network infrastructure, as shown in Figure 15.

Figure 15 Monitoring the Network Control Plane with NetFlow and Lancope StealthWatch



Design Considerations

NetFlow and the Lancope StealthWatch System

Lancope StealthWatch System

Add StealthWatch FlowSensors (Optional)

Where NetFlow generation is not possible from the network equipment, the Lancope StealthWatch FlowSensor and FlowSensor VE can be used to translate the communications into flow records. This enables networking equipment not specified in this guide to participate in deployments of the Cisco Cyber Threat Defense solution version 2.0. Additionally, the StealthWatch FlowSensor can be used to add packet-level application identification and performance metrics for key areas of the network. Perform the following steps when considering adding a StealthWatch FlowSensor to a Cisco Cyber Threat Defense solution version 2.0 deployment.

Procedure

Step 1 Choose a StealthWatch FlowSensor.

When choosing a StealthWatch FlowSensor, consider the expected traffic profile of the monitoring point, because the FlowSensor must be able to process the level of traffic being sent to it. As with any other NetFlow generation device in the Cisco Cyber Threat Defense solution version 2.0, Cisco recommends that the FlowSensor be deployed as close to the access layer as possible.

Table 3 lists the StealthWatch FlowSensor appliance models and their specifications. The processing capacity shown is the sustained rate supported. The FlowSensor can handle short bursts beyond the listed capacity. Like all NetFlow generators, the volume of NetFlow traffic generated by the StealthWatch FlowSensor varies based on the monitored traffic profile.

Table 3 StealthWatch FlowSensor Appliance Specifications

Model	Processing Capacity	Interface	Speed	Physical Layer	Form Factor	Power
250	100 Mbps	2	10/100/100	Copper	1 RU-short	Non-redundant
1000	1 Gbps	3	10/100/1000	Copper	1 RU-short	Non-redundant
2000	60,000	5	10/100/1000	Copper or Fibre	1 RU	Redundant
3000	120,000	1 or 2	1GB	Fibre	1 RU	Redundant

Note If the processing capacity of a single StealthWatch FlowSensor is reached, you can stack multiple FlowSensors using an appropriate Ethernet load balancer.

The StealthWatch FlowSensor VE is a virtual appliance that can be installed inside a vSphere/ESX host and used to generate NetFlow records for traffic between VMs in that host. The FlowSensor VE connects promiscuously to the virtual switches. It passively captures Ethernet frames from the traffic it observes and then creates flow records containing valuable session statistics that pertain to conversational pairs, bit rates, and packet rates. The FlowSensor VE then sends these records to the

StealthWatch FlowCollector. Table 4 describes the requirements for the deployment of the StealthWatch FlowCollector VE.

Table 4 StealthWatch FlowSensor VE Specifications

Disk Space Requirement	Flow Export Format	Minimum CPU Requirements	Minimum Memory Requirement	Interfaces
1.4 GB	NetFlow v9	2 GHz Processor	512 MB 1024 MB for application inspection	Up to 16 vNICs

Step 2 Integrate the StealthWatch FlowSensor into the network.

The StealthWatch FlowSensor must be placed in a Layer 1 or Layer 2 adjacent manner to the monitoring point. Sample deployment modes include using Test Access Ports (TAPs), Switched Port Analyzer (SPAN) ports, or a network hub. See the *System Hardware Installation Guide* on the Lancope StealthWatch Documentation CD for detailed information on how to integrate the StealthWatch FlowSensor into the network.

Choose a StealthWatch FlowCollector

The StealthWatch FlowCollector serves as a central collection and analysis point for NetFlow data generated by all NetFlow generators in the Cisco Cyber Threat Defense solution version 2.0. The choice of what number(s) and model(s) of StealthWatch FlowCollectors are needed in the solution deployment depends on the following factors:

- Decisions made in the previous sections influencing the volume of flows per second that will be reaching the StealthWatch FlowCollector
- StealthWatch FlowCollector deployment strategy
- Physical capacity of each StealthWatch FlowCollector

Procedure

Step 1 Determine the StealthWatch FlowCollector deployment strategy.

StealthWatch FlowCollectors can be deployed in a distributed or centralized manner. In a distributed deployment, FlowCollectors are deployed at multiple sites and usually placed close to the source producing the highest number of NetFlow records. This deployment has the advantage of limiting the overhead introduced by NetFlow. In a centralized deployment, all StealthWatch FlowCollectors are placed in a single data center (possibly behind a load balancer), providing the benefit of a single collection location and possibly a single IP address globally for NetFlow collection. This deployment offers advantages in environments where NetFlow generators are far apart.

There may be limitations in bandwidth between sites to consider as well (such as over a WAN). In general, a single FlowCollector should be used for as much related traffic as possible. The benefits of centralized collection diminish when the traffic is not similar.

When a particular FlowCollector receives flow data, it creates a single database entry for any duplicate flow records it receives for that flow. This de-duplication process ensures that the FlowCollector stores the flow data in the most efficient way while preserving details about each flow exporter and eliminating the reporting of inflated traffic volumes.

In an ideal implementation, every router that exports data related to a particular flow sends that data to the same FlowCollector. However, each unique host pair (or conversation) consumes additional resources on the FlowCollector. If the number of simultaneous connections gets too high, flow records are purged from memory. Take care during deployment planning to ensure that each FlowCollector has sufficient resources to keep state on all active conversations without purging records until after the conversations have been idle for some time.

Tip **Best Practice:** All NetFlow records belonging to a flow should be sent to the same StealthWatch FlowCollector.

Step 2 Performance considerations.

Each StealthWatch FlowCollector can support a minimum guaranteed flow volume, as listed in Table 5. However, also consider the following factors in the selection of a StealthWatch FlowCollector for the Cisco Cyber Threat Defense solution version 2.0:

- Exporter count—Number of NetFlow generation devices that each StealthWatch FlowCollector can accept.
- Data rate—Rate of fps that the StealthWatch FlowCollector is receiving.
- Host count—Number of hosts (both inside and outside the network) for which the StealthWatch FlowCollector can maintain state. Cisco recommends that the number of inside hosts not exceed 60 percent of the host count value.
- Flow storage—Amount of granular flow data required for a particular location on the network.

Note A system that approaches both the maximum number of exporters and the maximum data rate for a particular chassis may suffer from performance problems. For example, an estimated 10–20 percent reduction in the maximum data rate may occur at the maximum number of exporters.

Table 5 StealthWatch FlowCollector Appliance Specifications

Model	Flows per Second	Exporters	Hosts	Storage
StealthWatch FlowCollector 1000	Up to 30,000	Up to 500	Up to 250,000	1.0 TB
StealthWatch FlowCollector 2000	Up to 60,000	Up to 1000	Up to 500,000	2.0 TB
StealthWatch FlowCollector 4000	Up to 120,000	Up to 2000	Up to 1,000,000	4.0 TB

Table 6 lists the support for a StealthWatch FlowCollector VE based on the amount of reserved memory and the number of CPUs for the VM.

Table 6 StealthWatch FlowCollector VE Specifications

Flows per second	Exporters	Hosts	Reserved Memory	Reserved CPUs
Up to 4500	Up to 250	Up to 125,000	4 GB	2
Up to 15,000	Up to 500	Up to 250,000	8 GB	3
Up to 22,500	Up to 1000	Up to 500,000	16 GB	4
Up to 30,000	Up to 1000	Up to 500,000	32 GB	5

Choose a StealthWatch Management Console

The StealthWatch Management Console (SMC) manages the entire StealthWatch System installation and is licensed by the number of FlowCollectors that are connected to it and the total volume of flows monitored across the entire system.

Table 7 shows the SMC models and the number of StealthWatch FlowCollectors they can support.

Table 7 SMC Appliance Specifications

SMC Model	Maximum FlowCollectors	Size	Storage	Memory
SMC 1000	5	1 RU	1.0 TB	8 GB
SMC 2000	25	2 RU	2.0 TB	16 GB

Table 8 lists the number of FlowCollectors and concurrent users based on reserved memory and CPUs that the SMC VE can support.

Table 8 SMC VE Specifications

FlowCollectors	Concurrent Users	Reserved Memory	Reserved CPUs
1	2	4 GB	2
3	5	8 GB	3
5	10	16 GB	4

Note If a high number of host groups and monitored interfaces is expected in the deployment, a higher-performance SMC should be considered, because the amount of data being sent to the SMC can increase in these deployments.

Choose a StealthWatch FlowReplicator (Optional)

The StealthWatch FlowReplicator receives or monitors UDP packets and generates copies of those packets to send to one or more new destinations, modifying the packets as they traverse the appliance to appear as though they came from the original source. Each FlowReplicator comes with two active interfaces: one is assigned an IP address for management, monitoring, and generation of packet copies; and the other can be put into promiscuous mode for monitoring.

Each FlowReplicator is rated for a certain volume of input and output in terms of packets per second (pps). Each is tested against a generation of two to three copies per packet, but can support more destinations if required. Table 9 lists the StealthWatch FlowReplicator models and specifications.

Table 9 StealthWatch FlowReplicator Appliance Specifications

FlowReplicator Model	Processing Capacity	Physical Layer	Form Factor	Power	Fault Tolerant
1000	10,000 pps input 20,000 pps output	Copper	1 RU-short	Non-redundant	No
2000	20,000 pps input 60,000 pps output	Copper or Fiber	1 RU	Redundant	Yes

Note If the physical limits of the appliance are exceeded and too many copies are being generated for the link, packets are dropped.

NetFlow

Consult Table 10 for Cisco platform details.

Table 10 Platform Details

Platform	Hardware Details	Software Details	NetFlow Details	In CTD version 1?
Catalyst 3K-X	3560-X/3750-X with SM	IOS 15.0.(2)SE7	FNF (v9)	Yes
Catalyst 3850/3650	3850/3650	IOS-XE 3.3.5SE	FNF (v9)	Partial (3850 in 1.1.2)
Catalyst 4500	Sup7-E Sup8-E	IOS-XE 3.4.5SG IOS-XE 3.3.2XO	FNF (v9)	Partial (Sup7-E)
Catalyst 6500	Sup2T	IOS 15.0.(1)SY7a	FNF (v9)	Yes
Catalyst 2960-X (NetFlow Lite)	2960-X	IOS 15.0.(2)EX	NetFlow Lite (sampled V9)	No
ISR G2	2901, 2911	IOS 15.(3)M4	FNF (v9)	Yes
ASR 1000	ASR 1001/1002F	IOS-XE 3.10.xS	FNF (v9)	Yes
ASA 5500	ASA 5505, 5510	ASA 9.0.4	NSEL (v9)	Yes
ASA 5500-X with FirePOWER Services	ASA 5515-X, 5545-X	ASA 9.3.2 FirePOWER 5.3.1	NSEL (v9)	No
NetFlow Generation Appliance (NGA)	NGA 3240	1.0.2	FNF (v9)	Yes (3140)
UCS VIC	VIC 1240/1280/1225	2.2(2e)	FNF (v9)	No

Next Generation Intrusion Prevention System

FirePOWER Deployment Options

Designing secure networks requires a well-planned security policy that can be effective only if the underlying network is reliable and available. The Cisco ASA has long supported high availability and load balancing configurations to guarantee service.

Cisco ASA with FirePOWER Services is supported in the following deployment modes:

- Active/Standby—For locations where high availability is the primary concern
- Clustering—For locations where the following considerations apply:
  - Asymmetry is a concern
  - High Availability (HA) is required

  - Horizontal performance scaling is needed
- Multi-context—For separation of policy by logical and physical interfaces

Note FirePOWER Services on the Cisco ASA are supported in transparent and routed (L3) modes for both single and multiple Cisco ASA high-availability deployments.

Active/Standby for High Availability

HA using an Active/Standby model is supported on Cisco ASA 5500-X and Cisco ASA 5585-X platforms. Active/Standby defines one device in an HA-pair as the primary (active) and the other as the secondary (standby). To quickly facilitate a failover, the state of the Cisco ASA, including connection information, is shared between the firewalls. This is known as stateful failover. However, FirePOWER state information is not shared between FirePOWER modules in the failover pair.

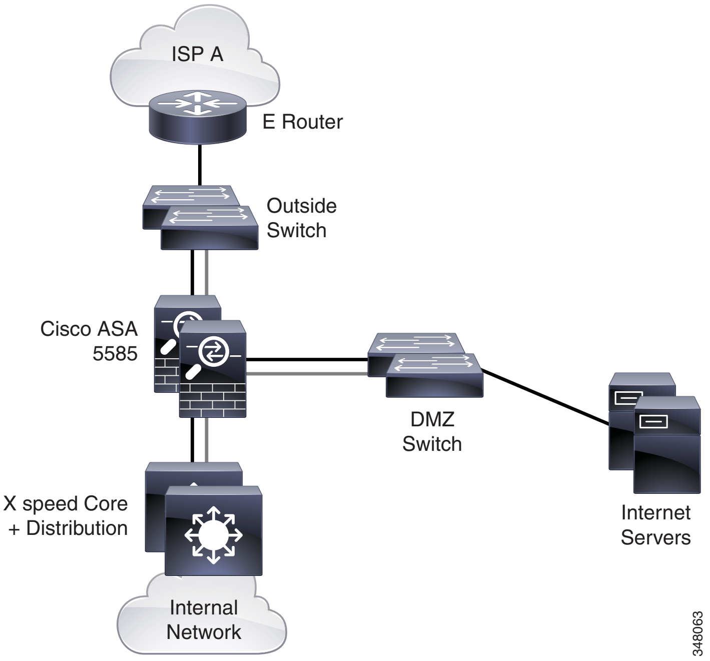
The FirePOWER Services module is inserted into the Cisco ASA packet processing path after normalization. The TCP/IP normalization service identifies abnormal packets that the Cisco ASA can act on when they are detected.

Figure 16 shows a typical Active/Standby deployment.:

- Available on ASA 5500-X and ASA 5585-X
- L2 transparent or L3 routed deployment options
- Failover link
- ASA provides valid, normalized flows to FirePOWER module
- State-sharing between firewalls for high availability

Note State sharing does not occur between FirePOWER modules

Figure 16 Typical Active/Standby Deployment



Clustering Deployment

For scalability, a cluster may contain up to 16 Cisco ASA 5585-X firewalls with FirePOWER Services Modules. The Cisco ASA 5500-X supports a cluster of two units.

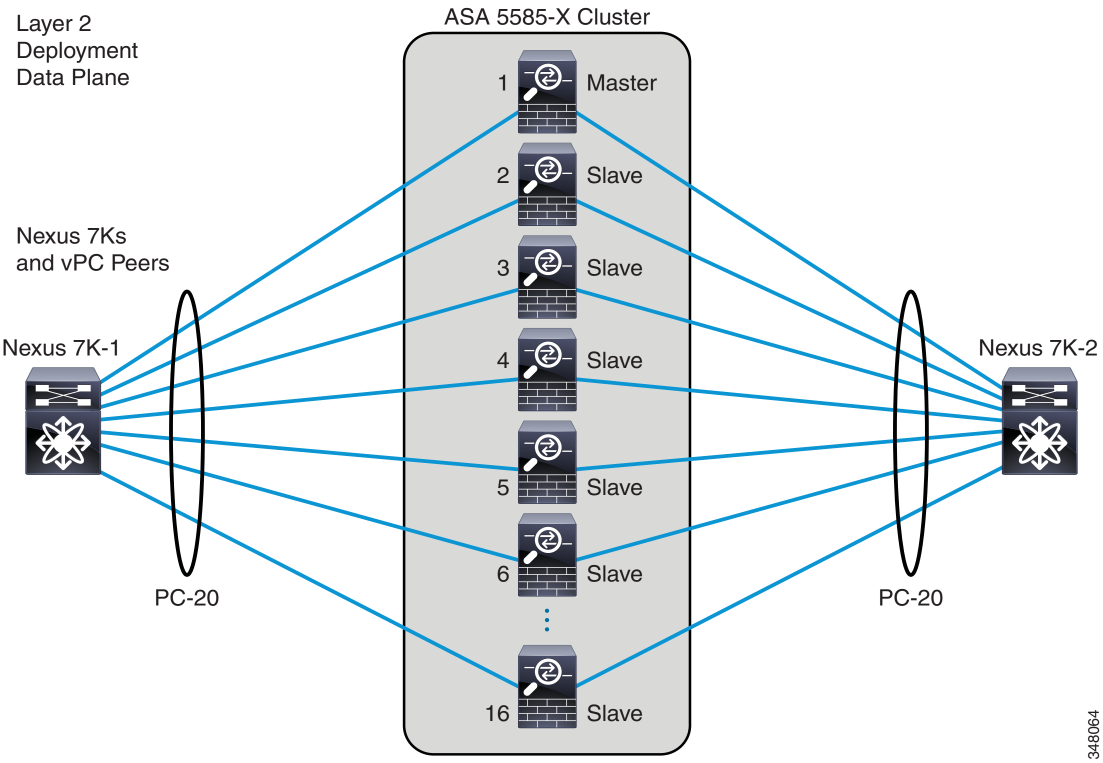
Cisco ASA clustering provides the following benefits:

- Ability to aggregate traffic to achieve higher throughput
- Scaling the number of Cisco ASA appliances into one logical firewall within the data center architecture
- True Active/Active model; when in multi-context mode, every member for all contexts of the cluster are capable of forwarding every traffic flow
- Removes asymmetry by sending packets that are received by an ASA that does not own a flow to the actual owner over the Cluster Control Link (CCL)
- Operates in either Layer 2 and Layer 3 modes
- Supports single and multiple contexts (firewall virtualization)
- Cluster-wide statistics are provided to track resource usage
- Single Cisco ASA configuration is maintained across all units in the cluster using automatic configuration sync

Note State sharing does not occur between the FirePOWER Services modules within the Cisco ASA cluster. Thus, the FirePOWER configurations are not synced within the cluster.

Figure 17 shows a typical cluster deployment.

Figure 17 Typical Cluster Deployment



Multi-Context Deployment

A common method for segmenting traffic through the Cisco ASA is to configure multiple context mode. This allows the enforcement of distinct security policies on each separate virtual security context. Some typical use cases for multi-context mode on the Cisco ASA include:

- Service providers supporting security services for many customers
- Large enterprises or college campuses that need to keep departments completely separate
- Enterprises that need to provide distinct security policies to different departments
- Networks that require more than one Cisco ASA but have budget restrictions

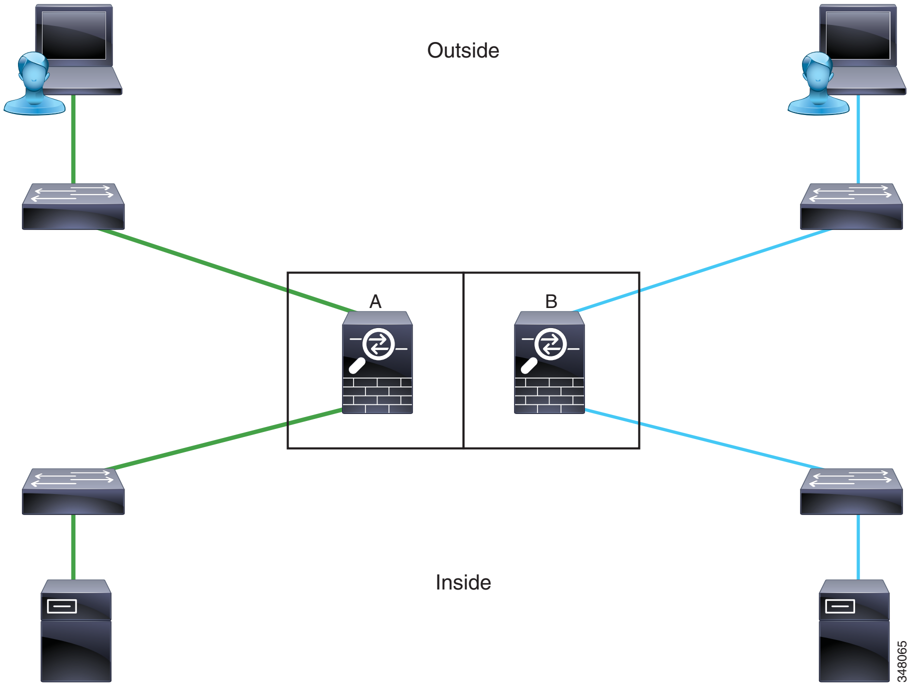
In multi-context mode, Cisco ASA interfaces and sub-interfaces are assigned to each context. When deploying FirePOWER services on the Cisco ASA in multi-context mode, the interfaces associated with each context can be grouped into security zones. Different FirePOWER policies can then be applied to each zone; for example, a policy might be applied for context A/interface outside to context A/interface inside.

Figure 18 shows a typical multi-context deployment:

- ASA can be configured in multi-context mode such that traffic going through the ASA can be assigned different policies.
- These interfaces are reported to the FirePOWER blade and can be assigned to security zones that can be used in differentiated policies.
- In this example, you could create one policy for traffic going from Context A to Context B inside, and then a different policy for Context B outside to Context B inside.

Note There is no management segmentation inside the FirePOWER module similar to the context idea inside ASA configuration.

Figure 18 Typical Multi-Context Deployment



Note Although the ASA Clustering deployment discussion is included in this document for additional context, ASA Clustering was not included in the validation. ASA Clustering has been extensively validated and documented in the Secure Data Center for the Enterprise Solution Portfolio. For more details, see the following URL:
http://www.cisco.com/c/en/us/solutions/enterprise/design-zone-secure-data-center-portfolio/index.html

FirePOWER appliances have a variety of deployment options available, including the following:

- Passive—Passive mode allows the sensor to inspect packets coming from a SPAN port on a switch or a network tap. Because the appliance is not in the packet flow (the packets are copies and not originals), there is no impact to the network, either in latency or in failure scenarios. The downside is that because the device is in passive mode, it is not working with the original packets and therefore cannot drop malicious traffic before it reaches its destination.
- Interface pairing—Interface pairing mode is the normal IPS mode where the appliance is acting like a wire and taking packets off one interface, inspecting them, and sending them down another interface, and vice versa. Because the appliance is inline, it can block malicious traffic including single packets, but the traffic must go through the appliance; therefore, some additional latency is added and failure scenarios must be accounted for. This is one reason features such as fail open interfaces are generally used.
- Virtual switch—This mode allows the sensor to act as a virtual switch and provide inspection of packets coming in off one or more virtual interfaces acting as a Layer 2 switch. This mode allows the device to drop packets much like interface pairing mode.
- Routed mode—This mode puts the device into the network as a Layer 3 device where each interface has an IP address. This is typically used when the device is acting as a firewall and applying Layer 3 and Layer 4 policy in addition to IPS, AVC, and URL policies.

Many different appliances are available with a broad range of performance levels. There are also numerous options for interface types.

The appliances can be set to Failopen or Failclosed mode depending on the desired results.

FireSIGHT Management Center

The FireSIGHT Management Center (FMC) comes in both appliance and virtual form factors. Both offer the same features but differ in performance. The appliance version can manage all types of FirePOWER sensors. There are two types of virtual FMCs: a full version that manages all sensor types, and one that manages only FirePOWER Services for Cisco ASA sensors.

High Availability is a feature available on FMC. This allows the Management Centers to share licensing and keep configurations synced. Virtual devices need to use the built-in recovery methods typical to virtual solutions.

Advanced Malware Protection

Cisco AMP provides protection across the broadest range of attack vectors and can be deployed as follows:

- A network-based solution, integrated into dedicated Cisco ASA firewall and Cisco FirePOWER network security appliances
- An endpoint solution for PCs, Macs, mobile devices, and virtual environments

- An on-premise, private cloud, virtual appliance built for high-privacy environments
- An integrated feature in Cisco Cloud Web Security or Cisco Web and Email Security Appliances

AMP for networks is a feature that can be licensed on any FirePOWER sensor. As such the design and deployment considerations are exactly the same as the FirePOWER service modules and FirePOWER appliances.

AMP for endpoints is a capability that can be installed on Windows PCs, Macs, virtual environments, and mobile devices. AMP for Endpoints not only monitors files for malicious content but can watch the actions that result from unknown files being brought onto a system. It can then block the file from spreading or causing damage on the system after it is determined to be malicious.

Content Security Controls

The Cisco content security products cover all aspects of secure gateway deployment options, including physical appliances, virtual appliances, and cloud offerings. Based on the customer's environment requirements, both email and web security offerings cover these options. Neither the Email Security nor the Web Security products are transparent inline devices; both must have some level of redirection or routing to get the payload to traverse the gateway.

Deployment Options for Email Security

The ESA functions as an SMTP Mail Transfer Agent (MTA). It is not intended as an email mailbox server such as Exchange, Lotus Notes, or other third-party data stores. SMTP routing to and from the ESA is done via DNS, using MX records based on the domain portion of an email message or defined SMTP routes based on A Records or IP addresses.

ESA hardware appliances are available in three models (C170, C380, and C680), suitable for small, medium, and large networks. The software functionality is identical on all sizes of appliance.

- ESA Virtual (ESAV) appliances are supported on VMware ESXi hypervisors, and are also available in several sizes:
  - Cisco ESAV C000v for evaluations only
  - Cisco ESAV C100v for up to 1000 mailboxes
  - Cisco ESAV C300v for up to 5000 mailboxes
  - Cisco ESAV C600v for large enterprises or service providers

As an alternative to physical or virtual ESA appliances deployed on premise, Cisco Cloud Email Security (CES) provides the same level of protection, with the service hosted in Cisco data centers. Hybrid deployments are also available.

A recommended practice is for the ESA to be the first SMTP hop on the way in and the last hop on the way out of the enterprise network. For simplicity, ESA appliances should usually be deployed with a single interface and IP address for both mail and management traffic.

Note A full discussion of ESA deployment options is beyond the scope of this document. For further details, see the Cisco Design Zone for Internet Edge at www.cisco.com/go/designzone.

Deployment Options for Web Security

WSA hardware appliances are available in three models (S170, S380, and S680), suitable for small, medium, and large networks. The software functionality is identical on all sizes of appliance.

- WSA Virtual (WSAV) appliances are supported on KVM or VMware ESXi hypervisors, and are also available in several sizes:
  - Cisco ESAV C000v for up to 1000 web users
  - Cisco ESAV C100v for up to 3000 web users
  - Cisco ESAV C300v for up to 6000 web users

Cisco CWS can be deployed as an entirely cloud-based solution or integrated together with on-premise WSA appliances. Cisco CWS Premium incorporates AMP and CTA for enhanced detection of advanced threats.

WSA appliances can be deployed using several options, including transparent redirection and explicit proxy configuration. A full discussion of WSA deployment options is beyond the scope of this document. For further details, see the Cisco Design Zone for Internet Edge at www.cisco.com/go/designzone.

TrustSec and the Identity Services Engine

Further information, including CVDs on Cisco TrustSec, ISE, and Endpoint Protection Services, is available at www.cisco.com/go/trustsec.

Conclusion

This document discusses some of the challenges in defending networks against modern, advanced threats and provides design guidance for the Cisco Cyber Threat Defense solution. This solution focuses on improving situational awareness and reducing the time required to detect and respond to threats acting inside the network, not just at the traditional network perimeter. To do this, it enables the network infrastructure to provide increased visibility and control, and incorporates Cisco's leading security technologies in an integrated design. The end result is a real improvement in the defender's ability to detect, block, and remediate threats in new ways.

References

General Security Information

- National Institute of Standards & Technology (NIST) IR 7298 Revision 2: *Glossary of Key Information Security Terms*—http://nvlpubs.nist.gov/nistpubs/ir/2013/NIST.IR.7298r2.pdf
- NIST SP 800-61 Revision 2: *Computer Security Incident Handling Guide*—http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf
- Lockheed Martin Corporation, *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains*—http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf
- Cisco Systems, *Cisco 2015 Annual Security Report*—http://www.cisco.com/go/security

- Mandiant, *APT1: Exposing One of China's Cyber Espionage Units*— http://intelreport.mandiant.com/Mandiant_APT1_Report.pdf
- CrowdStrike, *Actionable Indicators for Detection of Signs of Compromise from Target-Related Breaches*— http://www.crowdstrike.com/blog/actionable-indicators-detection-signs-compromise-target-related-breaches/index.html