Making Everything Easier!™

Blue Coat Second Edition

Big Data Security

FOR DUMMIES®

A Wiley Brand

Learn to:

- Harness the power of Big Data to detect advanced threats
- Collect digital evidence and streamline incident response
- Integrate Big Data Security into your existing security fabric

Brought to you by

BLUE COAT®

Steve Piper, CISSP

About Blue Coat

Blue Coat is a leader in enterprise security, providing on-prem, hybrid, and cloud-based solutions for protecting web connectivity, combating advanced threats, and responding to security breaches. Blue Coat is a global market leader in securing connection to the web and counts nearly 80 percent of the Global Fortune 500 as its customers.

For additional information, please visit **www.bluecoat.com.**

Big Data Security

FOR DUMMIES®
A Wiley Brand

Blue Coat Second Edition

by Steve Piper, CISSP

FOR DUMMIES®
A Wiley Brand

Big Data Security For Dummies®, Blue Coat Second Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

ISBN 978-1-119-13154-0 (pbk); ISBN 978-1-119-13155-7 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Publisher's Acknowledgments

Some of the people who helped bring this book to market include the following:

Development Editor: Kathy Simpson

Project Editor: Jennifer Bingham

Acquisitions Editor: Amy Fandrei

Editorial Manager: Rev Mengle

Business Development Representative: Karen Hattan

Special Help from Blue Coat Systems: Alan Hall, Ajay Uggirala, Greg Mayfield, Michael Rosen, Ellen Roeckl, Charles Tucker

Production Editor: Antony Sami

Table of Contents

Introduction

These days, the skill of modern threat actors has eclipsed the defensive capabilities of traditional cybersecurity tools. Major breaches consistently occur in the networks of the most technically savvy enterprises and government agencies — and even in security vendors' own networks.

Over the years, network security teams have implemented a wide range of set-it-and-forget-it tools that attempt to block threats based on signatures and traffic behaviors. Although these tools provide some defensive cover, they may also provide a false sense of security, as high-profile organizations continue to suffer new and ingenious attacks.

Security-conscious organizations are turning to Big Data Security in their fight against cybercrime. By aggregating and analyzing all available network intelligence — including raw packets, flow data, and files — these solutions can uncover advanced threats that traditional security defenses can't possibly detect. At the same time, they allow organizations to streamline their incident response efforts in ways never thought possible.

If you're responsible for securing your organization's network or responding to security incidents, you can't afford to miss this book.

About This Book

This book defines Big Data Security and its underlining security analytics and malware analysis technology in a way that any IT security professional can understand. I set the stage for Big Data Security by describing modern malware techniques and discussing why traditional security defenses aren't enough. Then I review the various form factors of Big Data Security solutions, explore their basic and advanced features, and describe how they're typically deployed. This is followed by a summary of Big Data Security use cases, methods for integrating security analytics into your existing security infrastructure, and a list of ten buying criteria for Big Data Security.

Foolish Assumptions

In preparing this book, I've assumed a few things about you, the reader:

- You work in the IT security field for a corporation, government agency, or services firm.
- You have foundational knowledge of computers and computer networking.
- You're responsible for securing your employer's IT systems and data and/or investigating security incidents.
- You often read about major data breaches and fear that your organization may be the next victim.

Icons Used in This Book

This book uses the following icons to indicate special content.

You won't want to forget the information in these paragraphs.

A Tip icon points out practical advice that can help you craft a better strategy, whether you're planning a purchase or setting up your software.

Look out! When you see this icon, it's time to pay attention. You won't want to miss this cautionary information.

Maybe you're one of those highly detailed people who really need to grasp all the nuts and bolts — even the most techie parts. If so, these tidbits are right up your alley.

Beyond the Book

Although this book contains lots of useful information about Big Data Security solutions, there's only so much ground I can cover in 72 pages. If you'd like to find out more about this topic, as well as review the features and benefits of enterprise-class security analytics solutions, go to `www.bluecoat.com`.

Chapter 1

Surveying the Cyberthreat Landscape

In This Chapter

- Reviewing the latest trends in cyberthreats
- Exploring modern malware and advanced hacking techniques
- Dissecting the costs of a successful data breach

With cyberthreats growing more sophisticated every day, hacking for kicks is a thing of the past. Nowadays, hackers can be well funded (or even state sponsored), highly motivated, and trained in advanced techniques to evade even best-of-breed security defenses.

Before I explain why traditional cybersecurity defenses simply can't keep up (see Chapter 2) and why leading IT security teams are turning to Big Data Security for answers (see Chapter 3), I discuss the cyberthreats that virtually all commercial and government organizations face — and the costs of failing to detect those threats.

What's the Risk?

Several reputable organizations monitor cyberthreat trends and the effects of those threats on organizations. Among these organizations is the Verizon RISK (Response, Intelligence, Solutions and Knowledge) Team, which publishes a highly regarded annual Data Breach Investigations Report. (To download a free copy of the report, connect to `www.verizonenterprise.com/DBIR`.) In 2015, Verizon analyzed more than 79,000 security incidents resulting in 2,122

confirmed data breaches that occurred in the previous year. This analysis yielded some interesting statistics:

- Approximately 85 percent of the incidents stemmed from external agents.
- In 60 percent of cases, attackers are able to compromise an organization within minutes
- The top three classes of attacks pertained to miscellaneous errors (inadvertent data disclosures; 29.4 percent), crimeware (25.1 percent), and insider misuse (20.6 percent).
- 23 percent of recipients now open phishing messages and 11 percent of them click on attachments

Who's at Risk?

This question has a simple, disturbing answer: everyone. Virtually every enterprise and government agency is a target for cyberattacks, right now, at this very moment. What's worse, most analysts estimate that at least 95 percent of large organizations are already infected — maybe even yours.

Successful cyberattacks make headlines every day around the globe. Following are just a few high-profile attacks that occurred in 2014 and 2015.

Commercial breaches:

- **AOL** (April 2014): Hackers compromised the email addresses, postal addresses, address-book contacts, and encrypted passwords of more than 2 million email users.
- **JPMorgan Chase** (July 2014): Hackers stole names, addresses, phone numbers, and email addresses for approximately 76 million households and 7 million small businesses.
- **Home Depot** (September 2014): Malware installed on POS devices in 2,200 stores resulted in the theft of up to 56 million customers' credit-card information.

- **Sony Pictures** (November 2014): This cyberattack was allegedly perpetrated by North Korea in response to a comedic movie called *The Interview.* The attack resulted in the theft and distribution of unreleased Sony Pictures films, confidential emails, security certificates, and more.
- **Anthem** (February 2015): Cybercriminals compromised nearly 80 million customer records held by the second-largest health insurer in United States.

Government-agency breaches:

- **European Central Bank** (July 2014): The agency announced that a database linked to its public website had been compromised, resulting in the theft of personal data related to people registering for events.
- **U.S. Postal Service** (September 2014): Data for more than 800,000 employees and 2.9 million customers was compromised, including names, dates of birth, Social Security numbers, and mailing addresses.
- **Oregon Employment Department** (October 2014): Hackers exposed the names, birth dates, Social Security numbers, and other personally identifiable information of approximately 850,000 job seekers.
- **U.S. National Oceanic and Atmospheric Administration** (October 2014): Hackers broke into U.S. weather systems and disrupted satellite transmissions.
- **White House** (October 2014): The unclassified Executive Office of the President computer network was breached, resulting in temporary outages and loss of connectivity for users.
- **U.S. State Department** (November 2014): Two weeks after the White House data breach (see the preceding item), the State Department revealed that its unclassified email network had been compromised.

My favorite website for viewing data-breach summaries is Information Is Beautiful (`www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks`.).

Trending Cyberthreats

When it comes to cyberthreats, the only constant is change. In this section, I review some of the most common trends in the ever-changing cyberthreat landscape.

Shifting attackers

Hacking has changed dramatically over the past half century. In the 1970s and 1980s, *phone phreaking* (unauthorized manipulation of telephone switching equipment, primarily to place free long-distance phone calls) was the craze. The 1983 movie *WarGames* introduced the general public to computer hacking, and the legend of hackers as heroes was born. In the 1990s, widespread Internet access gave admittance to a new generation of hackers, enabling them to crack software copy protection and deface websites primarily for bragging rights.

Today, you can classify the motives that drive most hackers in four general categories: cybercriminals, state-sponsored hackers, hacktivists, and malicious insiders. I discuss all four categories in the following sections.

Cybercriminals

Cybercriminals have one driving motive: to hack for some sort of profit. They may hack to steal credit-card numbers in bulk (sometimes by the millions) for sale in underground markets, or they may run operations (also quite profitable) that steal Facebook, Twitter, and/or email account credentials. Some cybercriminals develop hacking tools such as Trojans and offer services to prospective criminal entrepreneurs, contributing to a black economy estimated at billions of dollars.

Cybercriminals bent on espionage may break into corporate networks to gain access to proprietary data, which they could ransom back to the company or sell to the victim's competitors. Still others create *malware*, code designed to harm a computer system, and use this malware to exploit vulnerabilities from within.

State-sponsored hackers

Perhaps the most notable shift in the hacking community over the past decade has been the emergence of *state-sponsored*

hackers: hackers employed by nations to break into other nations' government and/or commercial computer systems to achieve political objectives.

These acts are often called *cyberwarfare* (or electronic warfare). Cyberwarfare pertains to stealing data, committing espionage, and crashing computer systems, and it could result in loss of life. In theory, a state-sponsored hacker could penetrate the network security defenses of a nuclear power plant, causing a widespread nuclear meltdown.

Hacktivists

Hacktivism is the use of computers to protest or to promote political ends. Unlike state-sponsored hackers (see the preceding section), hacktivists attack public websites to deface them, redirect traffic away from them, crash them (through denial-of-service attacks called *web sit-ins*), or steal confidential data. Two high-profile hacktivist groups are Anonymous and LulzSec.

Malicious insiders

A *malicious insider* is a current or former employee, contractor, or other person who misuses his or her authorized access to an organization's computing resources to damage the confidentiality, integrity, or availability of data or information systems. Malicious insiders typically are disgruntled employees, although some are motivated by profit.

Broader reach, bigger risk

Every host on your network is vulnerable to an attack. I don't care how secure you think your network is: If a human can access it via the network, any host can be compromised. In other words, every host *can* be compromised. Think of your network as being one large attack surface. The larger and more geographically dispersed it gets, the easier it is to penetrate.

Aside from the natural growth of your network due to the expanding nature of your business or government agency, several networking trends are causing your network's attack surface to expand at a rapid pace.

Social media

The recent tidal wave of social media applications has created new opportunities — and new risks — for business and government agencies. Popular social media sites such as Facebook, Twitter, and LinkedIn introduce risks that IT departments must evaluate and mitigate.

Social media sites continue to be conduits for malware. A well-intentioned user may connect to Facebook during lunch, click a seemingly innocuous link, and infect her computer with malware so new that none of her organization's perimeter- or host-based defenses can detect it.

Virtualization

Mainstream adoption of virtualization platforms such as VirtualBox, VMware, and Xen has reached critical mass. Nearly every large organization uses virtualization to host mission-critical applications in the data center.

Virtualization, however, poses a few inherent risks that don't apply to physical hosts:

- IT can't natively inspect traffic between virtual machines (VMs) without using a specialized tool.
- Many VMs go unprotected because IT hasn't yet budgeted for virtual security protection.
- New virtual hosts are frequently pushed into production without the knowledge (or approval) of IT security — a problem commonly known as *VM sprawl.*

Cloud computing

The explosion of *cloud computing* (applications delivered as services over a computer network; also known as Software As a Service or SaaS) has caused new concerns for IT security. Whether applications are deployed via a public cloud, a private cloud, or a hybrid, unless proper security measures are taken, data can be breached just as easily through a cloud architecture as it can through a traditional computer network.

When working with a cloud vendor, be sure to ask about its IT security defenses. Although security falls on the vendor's shoulders, simply assuming that the vendor has implemented appropriate security measures may be a costly mistake.

Mobile devices

Many people consider mobile devices, such as smartphones and tablets, to be the next frontier for cyberattacks. This trend, coupled with the Bring Your Own Device (to work) movement, causes immense concerns for IT. It's already a challenge for IT to secure all organization-owned devices; it's even harder and more complicated to secure all privately owned devices that connect to the network.

As a result, attacks on mobile devices — especially those targeting the Android and iOS operating systems — are growing rapidly in number and sophistication.

Encrypted communications

The use of encrypted communications is growing rapidly. According to information-security research firm NSS Labs, over 35 percent of enterprise traffic today is encrypted, and encryption is growing rapidly. Although encryption helps secure communications, encrypted sessions also pose a security risk for organizations that can't inspect encrypted traffic for threats (see the nearby sidebar "Viewing network blind spots with SSL visibility appliances").

Viewing network blind spots with SSL visibility appliances

With over one-third or more of network traffic being encrypted in organizations today, failing to inspect encrypted traffic for cyberthreats and data exfiltration is a risk that no enterprise can afford. Fortunately, innovative new SSL (Secure Sockets Layer) visibility appliances mitigate this risk.

An *SSL visibility appliance* is a purpose-built, high-performance device that inspects and decrypts both ingress and egress SSL (and Transport Layer Security, or TLS) traffic, based on policy, then passes it to a sequence of inline and/or passive network security devices, and re-encrypts the benign traffic before forwarding it to its intended destination — all with a negligible increase in packet latency.

Most security devices have little-to-no visibility and control of SSL encrypted traffic — or suffer significant performance degradation if they do. But an SSL visibility appliance

(continued)

(continued)

complements and enhances these solutions within the security infrastructure by exposing advanced threats and data exfiltration without compromising network throughput or device performance.

Sophisticated cyberattackers commonly cloak their subversive activities within encrypted SSL channels. When a host is infected with malware, attackers use SSL to connect with their command and control (C&C) servers and then issue commands to the infected host. When a server containing the targeted data of interest is compromised, data is often exfiltrated within SSL sessions.

SSL-encrypted traffic causes network blind spots that too many organizations ignore. Fortunately, SSL visibility appliances enable a more holistic encrypted traffic management solution that adheres to privacy and policy regulations while shedding light on these blind spots — helping organizations improve security, maintain data privacy, and reduce risk.

For more information on SSL visibility appliances, check out *Encrypted Traffic Management For Dummies, Blue Coat Special Edition* (Wiley), written by yours truly.

Modern malware

Over the past decade, malware has grown in both volume and sophistication. This section describes some of the latest trends in modern-day malware.

Custom malware

Perhaps the most significant trend in recent years is the increase in targeted attacks. After a cybercriminal identifies his target and his data of interest, he customizes malware that targets common vulnerabilities. This malware sails past traditional signature-based defenses.

Drive-by downloads

A *drive-by download* occurs when an unsuspecting user visits a compromised website and malware begins downloading in the background without the user even noticing. The user doesn't have to click a thing to become infected. Rather, she just needs an unpatched vulnerability (typically, within the web browser) that corresponds to the downloaded exploit.

Watering-hole attacks

Like drive-by downloads (see the preceding section), a *watering-hole attack* involves a website infected with malware. Two things distinguish these attacks from drive-by downloads, however:

- Watering-hole attacks may not incorporate drive-by downloads. Some require the user to click on a link.
- Watering-hole attacks are targeted. Cybercriminals compromise specific websites that they believe to be frequented by the people they want to target, such as sites specific to a particular industry conference or tradeshow.

Ransomware

Ransomware is a type of malware that compromises a computer system by encrypting personal files on the hard drive or restricting access to the master boot record or partition table. After the system is infected, the perpetrator of the attack demands a ransom payment. The files are rendered inaccessible or even deleted if the user fails to pay, and even making payment doesn't necessarily ensure that the criminal will uphold her end of the bargain and release the files.

A simpler, less-severe type of ransomware, called *scareware,* consists of bogus antivirus or clean-up tools that claim to have detected high-risk cyberthreats on the victim's computer and demand money to remove them.

Spearphishing

Spearphishing is an email-spoofing cyberattack that targets a specific organization for political or financial gain. Unlike phishing emails, which are opportunistic in nature (the same phishing email is sent to thousands of people), spearphishing emails are handcrafted, typically created after the attacker researches his target on social media sites such as LinkedIn and Facebook. Spearphishing messages often appear to be sent from someone known to the victim, such as a company executive, corporate IT, or a person from the victim's address book.

Spearphishing attacks typically contain malware that enables the attacker to take control of the victim's computer — malware such as a remote-access Trojan (RAT). The RAT "phones home" to the attacker's C&C server and awaits

further instructions. For this reason, higher level corporate executives are frequently targeted.

Spearphishing is widely viewed as the most common way for attackers to initiate advanced persistent threats (discussed later in this chapter).

Zero-day attacks

A *zero-day attack* is an attack on an unknown operating system or application vulnerability, so named because the attack occurs on *day zero* of awareness of the vulnerability. Thus, the developers of the operating system or application have zero days to patch the vulnerability.

Although the frequency of cyberattacks associated with zero-day vulnerabilities is low, such attacks are effective because they often go undetected for long periods. Also, when these attacks are detected, patching the vulnerability and remediating the damage still takes days or even weeks.

Advanced persistent threats

Advanced persistent threats (APTs) are sophisticated cyberthreats that use a variety of intelligence-gathering techniques to access sensitive information, such as infected media, supply-chain compromise, and social engineering. Attacks by individual hackers usually aren't called APTs, because individuals rarely have the resources to be both advanced and persistent, even if they're intent on gaining access to a specific target.

APTs define sophisticated tactics and techniques involving reconnaissance of an attack target, as well as the methods employed by the attackers, to penetrate the target network. APT malware is designed to remain undetected for extended periods.

By their very nature, APTs involve far more effort and research on the part of the attacker than common spam campaigns do, and they take place over a protracted period. APTs employ both social engineering and technological attacks to target specific (usually privileged) users of networks. These attacks may rely on a combination of previously stolen private information and public information about a business or government entity to identify those users as potential entry points on the targeted network.

Attackers' motives may fall into any of the four general categories described in "Shifting attackers" earlier in this chapter, but these attacks are most commonly associated with government or industrial espionage for political or economic gain. The information that a criminal targets in an APT, such as a database of credit-card numbers, may not be easy to identify or quantify after the fact. Also, hacktivists have been known to engage in long-term attacks against perceived enemies or groups that are hostile to their interests.

APTs are extremely difficult to detect because they use *low and slow* tactics to avoid detection. The attacker may gain access to the network by means of social engineering techniques to deliver malware to users of targeted systems. After the malware is installed, the attacker commonly attempts to map the network from the inside to identify other potentially vulnerable hosts. When such hosts are identified, attackers attempt to access them, capturing information over an extended period and remaining undetected. Stolen information sent back to the criminal operation may be batched into small chunks and delivered at random intervals — sometimes over an encrypted channel.

To stand a chance of mitigating the advanced threat techniques described in this section, organizations must leverage technologies specifically designed to detect them. Two such technologies are malware analysis appliances (described in the nearby sidebar "Defeating advanced threats with malware analysis appliances) and security analytics platforms (introduced in Chapter 3 and the core subject of this book).

Defeating advanced threats with malware analysis appliances

One of the most innovative advancements in network security technology over the past half decade is the malware analysis appliance. This device is designed to analyze suspicious files and URLs for the presence of malware.

As I mention earlier in this chapter, advanced malware is customized. Altering just a few bytes of code makes a file unique enough to bypass traditional signature-based threat defenses. A malware analysis appliance is designed to execute

(continued)

(continued)

suspicious files within the safety of a Windows-based virtual machine called a *sandbox*. By observing the automated execution and manipulation (see Chapter 2) of suspicious files and URLs, the malware analysis appliance assigns a risk score, helping security analysts identify and prioritize potentially successful host breaches.

When evaluating malware analysis appliances, look for vendors with the following capabilities:

- Multiple detection environments, including emulation and virtualization, to detect the greatest variety of sandbox-aware malware
- Layered analysis techniques — static, dynamic, reputational, and YARA rules — applied recursively to primary file and URL samples and their secondary attack vehicles including all dropped/downloaded files, call-back IPs, and referred URLs
- Capability to customize virtual machine profiles to match corporate gold image configurations
- Open malware detection criteria, customizable pattern matching rules, and custom risk scoring
- Rich analysis artifacts available for download and inspection by security analysts
- Open API for submitting samples and retrieving results from any connected authorized system

Malware analysis appliances are widely viewed as mission-critical in the fight against advanced threats. Be sure to select a solution that not only boasts a comprehensive feature set, but also fits within your Big Data Security strategy (see Chapter 6).

The Cost of Failure

As you can well imagine, the cost of a data breach can be immense. In some cases, it can kill a company. Companies that have been victimized by large-scale data breaches also face enormous costs in the following categories:

- Investigation and forensics costs
- Customer and partner retention costs
- Public relations costs
- Lost revenue due to damaged reputation
- Regulatory fines
- Civil claims and legal fees

According to a 2015 study by Ponemon Institute (www.ponemon.org), based on interviews with representatives from more than 350 companies that experienced a data breach in the preceding year, the average total organizational cost per data breach was $3.8 million. That total equates to $154 per compromised record.

Enterprises can't afford to miss out on the considerable data protection offered by Big Data Security solutions. Simply put, acquiring such a solution is a no-brainer.

Chapter 2

Why Traditional Security Isn't Enough

In This Chapter

- Understanding basic threat detection techniques
- Reviewing traditional security defenses
- Recognizing today's advanced threats

Some security pundits suggest that the network perimeter is dead. Although I agree that it's unwise to overinvest in perimeter defenses, as long as hosts connect to the Internet through a common gateway, you'll always have a network perimeter, and you'll always have to defend it. To give the pundits credit, however, IT departments don't always invest enough resources in monitoring internal threats that bypass perimeter defenses — threats that come through mobile devices, portable media, 3G/4G connections, and more.

Because no security solution is foolproof, no matter what any vendor says, the best approach to security is *defense in depth*. This approach involves applying layers of best-of-breed endpoint security, network security, and Big Data Security (more about this in Chapter 3) solutions to detect, prevent, and analyze ongoing cyberthreats.

In this chapter, I want to make sure that you're grounded in the traditional defenses that should comprise your defense-in-depth strategy and understand how they detect threats. More important, I want you to understand why traditional security defenses sometimes fail and why you need more to keep up with today's sophisticated attacks.

Basic Threat Detection Techniques

Security products that are designed to detect and/or block cyberthreats incorporate one or more of the following basic techniques: threat signatures, blacklisting, whitelisting, and behavioral profiling. To explain these concepts, I use a real-world analogy with air travel:

- **Threat signatures:** Before you can take your baggage onboard an airplane, it's scanned by a highly sensitive explosives-detection system that looks quite similar to the CT scanners at your local hospital. This machine is equipped with sophisticated pattern-matching software designed to detect many explosive devices. In the IT world, this example is analogous to signature-based antivirus clients and network intrusion prevention systems inspecting network traffic for known malware and exploits.
- **Blacklisting:** The U.S. Transportation Security Administration maintains a no-fly list to help keep known terrorists off airplanes. Similarly, IT departments employ *blacklisting,* wherein communications with known-bad Internet hosts are flagged for investigation or blocked.
- **Whitelisting:** U.S. Customs and Border Protection has a Global Entry program that enables frequent international travelers (Americans and foreigners from select countries) to receive expedited clearance through airport immigration upon arriving in the United States. In the IT security world, this practice is known as *whitelisting,* wherein communications from specific Internet hosts are approved in advance and exempted from more-detailed inspection.
- **Behavioral profiling:** Airport law enforcement personnel are trained to observe and question travelers who act suspiciously, even after they've passed through airport security. In IT security, *behavioral profiling* involves tools that attempt to detect anomalies from a baseline of normal network traffic (as in the spread of malware).

In the next section, I review traditional security defenses that incorporate these threat-detection techniques.

Traditional Security Defenses

The traditional endpoint and network security defenses described in this section are commonly used in today's enterprise and government networks.

Endpoint security

Endpoint security enables a computing device to assume at least some responsibility for its own security. Endpoint security systems work on a client/server model: Client software is installed on each network device, and a central server (or gateway) monitors and/or manages the client software installed on the devices.

Following are a few familiar examples of endpoint security:

- Virus, malware, and spyware prevention
- Personal firewalls
- Spam filtering
- URL filtering
- Application controls
- File integrity monitoring

All the remaining security products described in this section fall into the category of network security devices.

Intrusion prevention systems

An *intrusion prevention system* (IPS) monitors network traffic for malicious activity. If an IPS is configured for *inline* operation (that is, if it's in the direct path of flowing traffic), it can block threats. If, however, the IPS is configured for *passive* operation (that is, it merely inspects copied traffic), it's capable only of providing alerts when it detects threats.

The latter mode of operation is also referred to as *intrusion detection system* (IDS) mode. Today, IDS is a mode of operation within an IPS rather than a unique product offering.

IPS solutions incorporate thousands of signatures to detect known and unknown threats to operating systems and applications. Better IPS offerings also incorporate anomaly-based detection methods (see "Network behavior analysis" later in this chapter) and stateful protocol analysis.

Next-generation firewalls

A *next-generation firewall* (NGFW) is a multifunction security appliance that incorporates firewall, IPS, and application control processes into a unified network security platform. Enterprises are turning to NGFWs to reduce network security costs — and, perhaps more importantly, to implement full-featured application control, a key component of modern NGFW offerings.

Secure email gateways

A *secure email gateway* (SEG) monitors inbound email traffic for spam, viruses, malware, and other threats. It also monitors outbound email traffic for sensitive data such as credit-card, Social Security, and bank-account numbers.

Secure web gateways

A *secure web gateway* (SWG) monitors inbound web traffic for malware and other cyberthreats. It also performs URL filtering on outbound traffic to restrict web browsing to safe, IT-approved websites. Better SWG solutions offer granular policy-based control of web-based applications, such as instant messaging, multiple-player games, peer-to-peer applications, and Voice over IP (VoIP) applications. They also offer strong user authentication, data loss prevention (DLP), content caching, bandwidth management, and more.

Data loss prevention systems

Data loss prevention (DLP) is designed to detect (and in some cases prevent) potential breaches by inspecting data in use,

in motion, and at rest. DLP solutions can be configured to search for two types of data:

- **Described:** Data formatted with a fixed schema, such as credit-card and Social Security numbers
- **Registered:** Data such as a specific document, spreadsheet, block of text, or database record

Today, security systems such as IPS, NGFW, SEG, and SWG incorporate basic DLP capabilities in that they search for described data. Organizations that are highly sensitive to data leaks, however, should consider implementing an enterprise-class DLP solution.

Network behavior analysis

Network behavior analysis (NBA) is a sophisticated network security system that baselines normal network traffic to detect anomalies, such as advanced persistent threats and advanced targeted attacks (see Chapter 1). It processes network flow records (such as NetFlow, cFlow, jFlow, sFlow, and IPFIX) generated by routers and switches, and then applies sophisticated algorithms to those records to uncover potential threats.

NBA solutions can also monitor for internal cyberthreats — threats that may be carried right through the office front door on laptops and other mobile devices.

Malware analysis appliances

As I imply in Chapter 1, malware analysis appliances are critical components of any sensible defense-in-depth strategy. These devices are designed to detect custom malware and other advanced threats that slip past traditional signature-based defenses using dynamic (behavioral) analysis. Better sandboxes add static code analysis, reputational analysis, and custom YARA rules to create a complete picture of malicious traits, indicators, and behaviors.

Although mainstream adoption of malware analysis appliances is fairly recent, sandbox technology has been around for more than a decade. Unfortunately, the bad guys have

found clever ways to circumvent sandboxes (see "Evading sandbox detection" later in this chapter), keeping sandbox vendors on their toes.

Security information and event management

Security information and event management (SIEM) systems frequently serve as central points for managing and analyzing events from network security devices across large distributed enterprises and government agencies. A SIEM has two primary responsibilities: aggregating events and logs from network devices and applications, and using that intelligence to uncover network problems. SIEMs are useful for uncovering threats within the data they're configured to receive, but that data represents only a fraction of the data available to Big Data Security solutions.

Don't confuse SIEMs with log management solutions, which don't correlate security intelligence. Log managers merely aggregate logs to satisfy regulatory compliance or to provide a convenient means of querying logs.

Advanced Threat Techniques

As I say at the start of this chapter, no cybersecurity solution is foolproof. Following are advanced threat techniques that experienced threat actors employ to bypass run-of-the-mill traditional security.

Customizing malware

The easiest and most common way to evade traditional signature-based defenses is to customize the malware associated with each threat campaign. If the file's signature (typically represented by an MD5 checksum or hash) is unique, it's likely to sail past your legacy defenses as though they weren't there.

Exploiting zero-day vulnerabilities

Arguably, the most effective way for an attacker to compromise a host is to create malware designed to exploit a zero-day vulnerability. Although uncovering a never-before-seen vulnerability is no trivial exercise, it's a highly effective approach, as every host associated with the unknown vulnerability can be compromised with relative ease.

Hiding within SSL traffic

As I mention in Chapter 1, threat actors often cloak their subversive activities within encrypted sessions, which unfortunately go uninspected by most security devices. Such activities include compromising hosts with malware from infected websites, controlling infected hosts through command-and-control (C&C) servers, and exfiltrating compromised data.

Employing multistage, multivector attacks

When they target an enterprise or government agency, savvy threat actors know that there's more than one way to skin a cat. Instead of targeting one person with a spearphishing email, an attacker targets dozens or even hundreds of people through multiple attacks: spearphishing, drive-by downloads, watering-hole attacks, and more (see Chapter 1).

Each attack entails multiple stages, from initial point of entry to network reconnaissance to data exfiltration. Attackers are very good at covering their tracks by uninstalling any malware that was used during the attack.

Leveraging domain-generation algorithms

Domain-generation algorithms are algorithms used in various malware families to generate large numbers of Internet domain names that can be used for hosting C&C servers. Although such throwaway domains typically are short-lived — active

for only days or sometimes weeks — they're effective against solutions that incorporate domain-name blacklists.

Domain-generation algorithms were popularized in 2008 by the family of Conficker worms, which ultimately generated more than 50,000 new domain names per day.

Evading sandbox detection

Detecting advanced malware is a game of cat and mouse. The good guys invented a way to evaluate suspicious files within the safety of a sandbox. Then the bad guys evolved their malware to evade sandbox detection through the following techniques:

- Suppress malicious payload if the operating environment exhibits the attributes of a virtual machine (VM).
- Suppress malicious payload until a mouse click is observed.
- Suppress malicious payload for minutes or hours, or until a future date and time has been reached.
- Suppress malicious payload until a specific action has been achieved (such as scrolling down to page 5 of a file).
- Suppress malicious payload if the operating environment is bereft of files and folders generally associated with a user.
- Present dialogs or multistep installers to determine if a real user is present on the system.

Fortunately, leading modern malware analysis appliances overcome these payload suppression techniques by masking virtual machine indicators, simulating mouse clicks, responding to dialogs, navigating through installers, detecting and bypassing sleep calls, and allowing for the creation of realistic files and folder directory structures.

Don't assume that all malware analysis appliances are alike. Before signing the sales order, be sure to discuss with your chosen vendor how its malware analysis offering circumvents sandbox evasion techniques.

Chapter 3

Understanding Big Data Security

In This Chapter

- Defining Big Data in the context of information security
- Contrasting limited- and full-visibility Big Data Security solutions
- Exploring security intelligence and analytics

To succeed in information security, you must come to terms with a universal truth: No matter how hard you try or how much money your organization spends, your network *will* be compromised at some point. (In fact, it probably already has been compromised!)

The security defenses that I discuss in Chapter 2 offer at least some protection against a wide variety of threats, but they're simply not enough in today's world. Luckily, you already have what you need to improve your situational awareness, provide context, and make your existing defenses more effective.

I'm talking, of course, about your organization's Big Data.

What Is Big Data?

Big Data refers to collections of data sets so large and complex that they're awkward to work with when you use traditional database management and analysis tools. Big Data challenges include capturing, storing, searching, sharing, analyzing, and visualizing large data sets.

Big Data isn't unique to information security. It's applicable to myriad use cases, including scientific discovery, economic analysis, business intelligence, counterterrorism, and fraud detection. Because this book is about Big Data Security, I limit my interpretation to the realm of information security, starting with common Big Data sources.

Internal sources

Typical internal sources of Big Data include the following:

- All IP traffic flowing across your network, including web traffic, email, and file transfers
- Network flow records (such as NetFlow, cFlow, jFlow, and sFlow) from network routers and switches
- VM-to-VM (virtual machine to virtual machine) IP traffic on VMware, Xen, and other virtualization platforms
- User account directories, such as Microsoft Active Directory and LDAP
- Detonation and behavioral analysis result feeds from malware analysis appliances

External sources

Typical external sources of Big Data include the following:

- Cyberthreat and reputation feeds, such as Emerging Threats, Google Safe Browsing, Malware Domain List, SANS Internet Storm Center, SORBS (Spam and Open-Relay Blocking System), VirusTotal, and other spam or IP address blacklists
- IP geolocation services, such as Digital Envoy, Geobytes, MaxMind, and Quova
- Website intelligence services, such as DomainTools, Robtex, and the global domain registry database

In the next section, I delve into commercially available solutions that can harness these sources.

What Is Big Data Security?

Big Data Security refers to any computer-based *solution* (combination of hardware and software) that captures and analyzes some or all Big Data sources to uncover and mitigate cyberthreats.

Some commercial Big Data Security solutions incorporate most of the Big Data sources mentioned in the preceding section into one highly scalable platform. Solutions of this type — called *full-visibility solutions* — are ideal for uncovering unknown or hard-to-detect cyberthreats and for performing forensic analysis after a cyberattack occurs.

Other Big Data Security solutions — called *limited-visibility solutions* — incorporate only some of the Big Data sources and are best suited to uncovering known or easy-to-detect threats.

Limited-visibility solutions

In this section, I provide examples of solutions that offer a relatively limited view of a network.

Security information and event management

A reliable *security information and event management* (SIEM) system is critical to the success of any enterprise information security organization. SIEMs complement existing IT security investments — firewalls, intrusion prevention system (IPS) and network access control (NAC) appliances, data loss prevention (DLP) solutions, secure web gateways (SWGs), endpoint security solutions, and so on — by aggregating and correlating log and event data in a central location.

SIEMs are particularly useful for aggregating security events to a single customizable dashboard so that security professionals don't have to monitor multiple consoles for alerts. SIEMs also help organizations affected by regulatory standards meet certain log-aggregation and monitoring obligations.

A SIEM can be classified as a Big Data Security solution, but its visibility is limited to the data that it's fed, such as log, flow, and endpoint event data.

Network behavior analysis

A *network behavior analysis* (NBA) solution can be classified as a Big Data Security solution because it aggregates, profiles, and inspects massive numbers of flow records (such as NetFlow, jFlow, sFlow, and cFlow) from network routers and switches.

NBA solutions may also detect cyberthreats based on behavior profiling (see Chapter 2). To detect this type of traffic, these solutions baseline normal network traffic so that they can detect anomalies such as malware propagation or exports of large amounts of data from a rogue host. They also offer numerous benefits for network operations personnel, including the capability to troubleshoot network outages and performance degradations.

As useful as NBA solutions are, they can inspect only the traffic flow information that they're configured to receive. As a result, NBA may not provide visibility into the full traffic payload, including user and application data.

Endpoint security

Modern endpoint security platforms offer a wealth of host-based intelligence, including data on operating systems, applications, and (in some instances) missing system patches and uncorrected security misconfigurations.

You could argue that endpoint security platforms can be classified as Big Data Security solutions. In practice, however, they're limited to the hosts they're configured to monitor. Also, they offer little to no insight into what data has been compromised in a cyberattack.

Full-visibility solutions

What if you could have a Big Data Security solution that truly sees everything on your network? Imagine a giant digital video recorder that records every TV show on every channel and stores the recordings for weeks at a time. Now imagine a solution that not only records every packet that traverses your network, but also provides built-in traffic-analysis capabilities.

I'm pleased to tell you that such a solution exists, and it's available to you right now. This solution is a relatively new

category of information security technology called *security analytics*.

From this point forward, when I discuss the merits of Big Data Security, I'm referring primarily to security analytics, because it offers the greatest breadth and depth of Big Data analysis for information security.

Introducing Security Analytics

I consider security analytics to be the ultimate Big Data Security solution. By capturing every packet that traverses your network, security analytics provides an additional layer of defense by detecting threats that are virtually invisible to traditional defenses.

Unlike traditional network security devices, which inspect only a portion of your traffic and capture packets only at the point of attack, security analytics solutions capture and inspect all network traffic before, during, and after the attack.

Security analytics solutions offer several distinct features, including the following:

- Incident response and forensics
- Situational awareness
- Cyberthreat detection
- Data loss monitoring and analysis
- Verification of an organization's policy compliance
- Security assurance (always-on verification of the effectiveness of your other security tools)

Chapter 4 provides detailed explanations of these six security analytics use cases.

Security analytics is one of the fastest growing technologies in the network security industry. In the 2015 Cyberthreat Defense Report published by information security researcher (and my employer) CyberEdge Group, more than 800 IT security professionals in North America and Europe identified the network security components that their organizations plan to

acquire within the coming year. The most common response was "security analytics."

To download a complimentary copy of this report, connect to www.cyber-edge.com/2015-cdrl.

How it works

A security analytics platform is the computer-network equivalent of a closed-circuit security camera system. It's always on, recording network activity 24 hours per day, 7 days per week. Your security analytics platform never sleeps.

Security analytics solutions are designed for enterprise use, capturing and indexing data (including packet header and payload, OSI Layers 2 through 7) at wire speed, providing a complete, forensically sound record of all network activity. These solutions also perform real-time or back-in-time analysis of files, applications, flows, or packets. They must have ample storage capacity because they record and store terabytes of data for days, weeks, or even months.

Security analytics form factors

Leading security analytics providers offer solutions in three form factors: physical appliances, virtual appliances, and software. These multiple choices enable an organization to select the ideal platform for its networking environment.

Physical appliances

Most organizations choose to deploy security analytics solutions on purpose-built appliances (see Figure 3-1) selected by the vendor to achieve a given level of throughput and performance. IT departments don't need to worry about selecting appropriate hardware or installing and configuring software, because the solutions incorporate hardened preinstalled operating systems.

Virtual appliances

To maintain complete visibility of your network, I suggest that you also implement virtual security analytics appliances to inspect VM-to-VM traffic in your virtual network infrastructure (such as VMware ESX, Citrix XenServer, Windows Hyper-V,

and KVM). Otherwise, this traffic results in a blind spot on your network, and you're not leveraging your security analytics platform to its fullest potential.

Figure 3-1: Sample security analytics appliance from Blue Coat.

Software

Some organizations prefer to select the hardware that houses the security analytics software. (In fact, some government agencies are required to choose hardware from an approved-products list.) These organizations can opt to deploy security analytics software on their own hardware platforms. If so, IT personnel must install, configure, and maintain the system and ensure that it's optimized for the required performance level.

When evaluating security analytics platforms for your organization, avoid offerings that are restricted to just physical appliances because they limit your deployment options and prevent you from inspecting VM-to-VM traffic. Also look for solutions that can scale and provide a recommended 30-day window of captured traffic.

Common features

Not all security analytics solutions are created equal, so it's important to assess both the basic and advanced features that are important for your organization. Following are descriptions of some features offered by virtually all security analytics solutions.

Interactive dashboard

The security analytics dashboard (see Figure 3-2) is the primary user interface for monitoring the system and investigating potential threats. Better security analytics solutions offer

a selection of prebuilt dashboards and a library of drag-and-drop widgets so that users can customize and interact with the dashboard to their liking. A widget may contain summary data within a table, for example, or a pie, bar, or column chart.

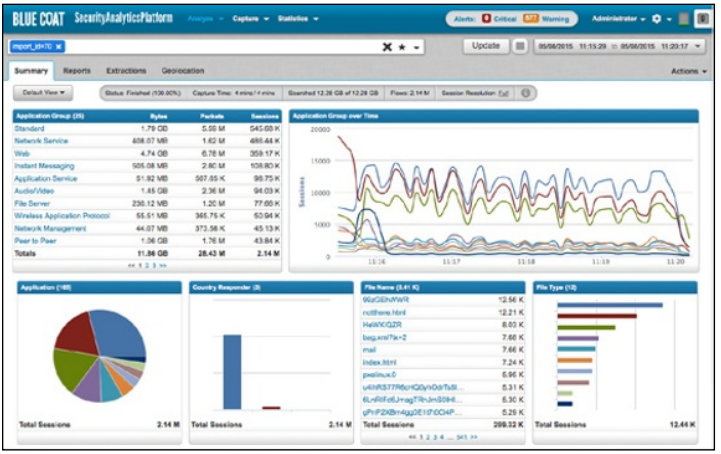

Figure 3-2: Sample security analytics dashboard.

Global threat intelligence

In recent years, security analytics offerings have evolved from solutions designed to investigate threats into solutions that also detect threats that other security defenses miss. To accomplish this objective, modern security analytics platforms incorporate up-to-the-second global threat intelligence from a variety of feeds, including IP, URL, and file reputation feeds and geolocation feeds.

Rules and alerts

Most security analytics offerings enable users to configure rules and alerts. A *rule* is an action that can be taken when recorded traffic meets a certain condition, such as routing captured traffic associated with a known-bad site to an advanced malware protection device for further analysis. An *alert* is simply a notification that one of your rules has been triggered. Alerts are often displayed on the dashboard but may also be sent via email or text message.

Comprehensive reporting

Today's security analytics platforms feature powerful reporting capabilities and provide easy-to-read report templates that IT departments can customize to meet their needs. Reports include the following details (and much more):

- Identification of the network traffic generated by common web-based applications, such as Webmail
- Email sender and receiver addresses, and the subject lines of messages
- Usernames, nicknames, or accounts on instant-messaging, chat, and social media sites
- File characteristics, including names, MIME types, content disposition, and transport method
- IPv4 and IPv6 source/destination addresses and their address space geolocations
- HTTP server names, referrers, web queries, port numbers, Secure Sockets Layer (SSL) common names, and full URLs

Query favorites

Just as web browsers enable you to save favorites, modern security analytics applications enable users to save custom search queries (sometimes called *filters)* for future use. Users can quickly execute favorite search queries to home in on threats, malware, or suspicious traffic.

Metadata retention

In addition to analyzing several days' worth of full packet data, security and incident responders often want to perform long-term analysis of network traffic to evaluate anomalous or suspicious behavior. Unfortunately, storing a full year's worth of packet data is rarely realistic, given the amount of data involved.

Security analytics platforms, however, provide metadata retention, enabling security analysts to devote a portion of storage to full-packet capture and another portion to network metadata. This feature allows analysts to optimize their systems to retain an appropriate amount of full-packet data (perhaps a week's or a month's worth) while allocating enough space for a year's (or more) worth of metadata.

Advanced functions

The preceding features are available in most security analytics offerings, but you'll find the following only in advanced security analytics solutions.

Geolocation

Geolocation is the practice of assessing the real-world location of an Internet-connected computer or device. Preferred security analytics solutions offer geolocation integration, which enables users to view the origin, destination, and flow of network traffic (see Figure 3-3).

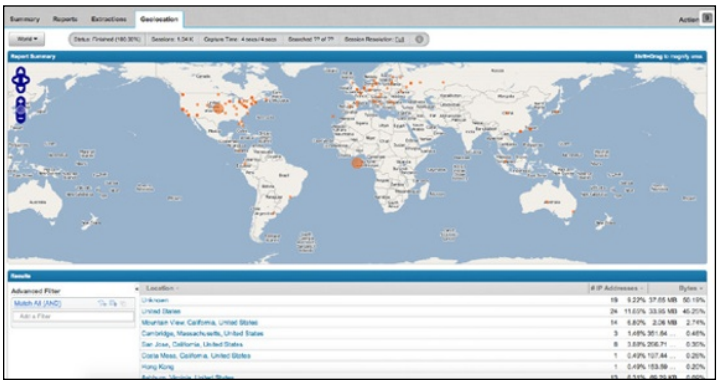

Figure 3-3: Geolocation view of external traffic sources.

Geolocation enables analysts to identify patterns and concentrations of traffic traveling to and from unusual or unexpected locations, such as countries where you have no business dealings. Users can zoom in on specific paths and flag IP addresses, locations, or even countries that appear to be suspicious. Some security analytics solutions even allow users to import data directly into Google Earth.

Root-cause exploration

When a suspected network intrusion occurs, time is of the essence. To accelerate network forensic tasks, preferred security analytics solutions offer a capability commonly referred to as *root-cause exploration.* This feature automates the identification of an actual session or file that caused a security

incident. By enumerating these events, the feature helps analysts quickly trace back the source of an infection or network breach, dramatically reducing time to resolution.

Customizable detection and risk scoring

Leading security analytics vendors allow the customer to determine what, if any, supplemental threat detection will be performed by the security analytics platform. Such solutions typically allow customers to detect threats associated with web traffic, email communications, and stored files. When a threat is detected in one or more of these media, a risk score (rather than a good/bad rating) is assigned to help security analysts prioritize their investigatory efforts.

File reconstruction

Many security analytics solutions allow users to reconstruct original documents, images, or executables that traversed the network. Better solutions have additional reconstruction capabilities, such as email, chat messages, web pages, and file transfers over tunneled protocols. Full-event reconstruction is possible because every packet is recorded, classified, and indexed. All this recording provides a massive amount of data to sift through, however, and only the better security analytics solutions can perform reconstruction at real-time speed. This capability enables incident responders to act as soon as the threat is detected.

File reconstruction is also useful for determining what, if any, confidential information has been exfiltrated from the network in a security breach. When an organization must publicly report a breach, having this information can save it millions of dollars and immeasurable bad press. What if only 10 database records were stolen and not all 10,000,000? Would that be worth hearing about on the evening news? Security analytics helps organizations right-size their incident response efforts so they don't overreact — or underreact — to successful cyberattacks.

Web, email, and chat reconstruction

Some sophisticated security analytics solutions enable analysts to view web pages exactly as users originally saw them. Analysts can also review instant-message (IM) and email conversations in their original forms for clues to the source of a security event.

Many cyberthreats begin with a compromised website, a malicious file, or a link in an email or IM. When you evaluate security analytics solutions, consider reconstruction (of web pages, emails, and chat messages) to be a must-have feature.

Third-party integration

No IT security solution should work in a vacuum. Rather, security solutions should work in concert to simplify monitoring, increase effectiveness, and reduce total cost of ownership.

Better security analytics solutions offer application programming interfaces (APIs) to integrate with popular SIEM, IPS, next-generation firewall (NGFW), malware analysis appliances, and unified threat management (UTM) solutions, as well as popular endpoint forensics platforms. (For details on these information security solutions, see Chapter 2.)

Some security analytics providers also offer a universal-connector client application that integrates with popular web browsers. If you see something suspicious while you're using a traditional IT security product, you can click the universal-connector sidebar to view the network activity you want to investigate within your security analytics application. Universal connectors save analysts time and effort by making it easy to conduct quick analysis of suspicious indicators.

Deploying Security Analytics

After you choose the best possible security analytics solution for your organization, with just the right combination of standard and advanced features, you face a new task: deployment. This section gives you some pointers.

Why size really matters

Whether you're deploying your security analytics solution on purpose-built appliances (as most organizations do) or on your own hardware platform, make sure that you've got enough CPU, memory, and storage to accommodate your organization's requirements for sustained throughput and archiving.

If your hardware falls short, some element of operation could be compromised, and you may not have all the data you need for real-time analysis.

Leveraging SPAN ports and TAPs

A common way to give your security analytics appliance full network visibility is to connect it to the SPAN ports on your network switches. *SPAN ports* replicate all traffic flowing through a switch, typically for the benefit of network security and performance tools. Enabling SPAN ports can negatively affect the performance of switches, however.

If your SPAN ports have already been allocated to other tools — a problem commonly referred to as *SPAN-port contention* — or if your switches are operating at capacity, you can use a network TAP to tap your network (typically, between a router and a switch or between two switches) and replicate traffic to your security analytics appliance.

New high-end TAPs, called *network packet brokers,* enable you to aggregate traffic from multiple SPAN ports or TAPs for the benefit of a single security analytics appliance. Figure 3-4 shows a typical security analytics deployment, featuring a security analytics appliance connected to the network through a network TAP. This diagram also features an additional security analytics storage array — an option commonly available from leading security analytics vendors.

Supporting a distributed architecture

If your organization is large and/or geographically dispersed, you may require a distributed security analytics architecture. In this case, you can use a central console to manage multiple security analytics appliances (plus virtual appliances and/or servers) to aggregate views and reports across many security analytics systems. Figure 3-5 depicts such an environment.

Removing SSL blind spots

As I mention in Chapters 1 and 2, threat actors often cloak their activities within encrypted sessions. Be sure to use an

SSL visibility appliance (see Figure 3-6) to effectively inspect and decrypt SSL traffic before capturing to maximize success in detecting and investigating advanced threats and potential data exfiltration.

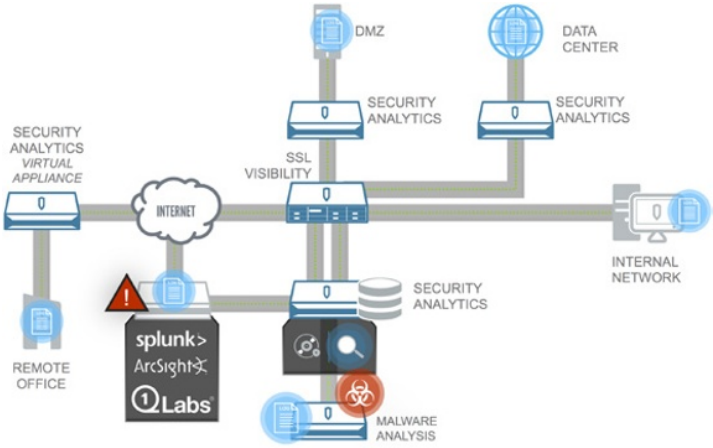

Figure 3-4: Typical security analytics deployment architecture.

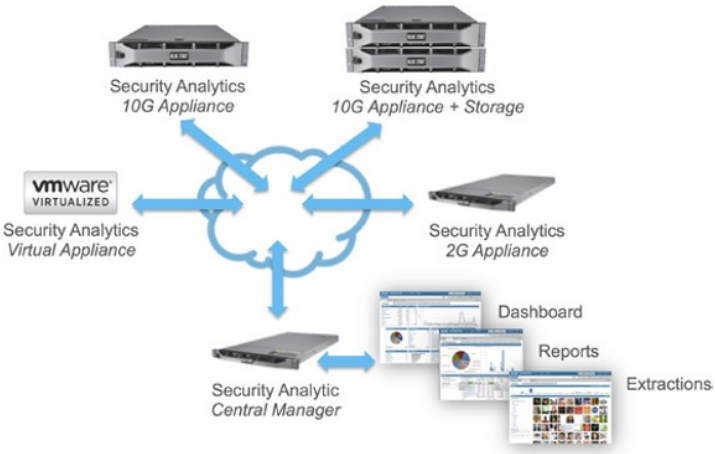

Figure 3-5: Distributed security analytics architecture.

Figure 3-6: Sample SSL visibility appliance from Blue Coat.

A security analyst's worst nightmare

Imagine that you're director of network security at a Fortune 500 financial-services company. Your home phone rings at 3 a.m., rousing you from the calm of a deep, warm sleep. As you fumble for the phone, you look at Caller ID and realize that your boss — the company's chief information security officer — is calling. Your heart begins to race.

It turns out that your boss just received a call from the company's vice president of public relations. Overnight, a prominent hacking group infiltrated your company's network and exfiltrated data pertaining to more than 100,000 credit-card accounts. Anonymous alerted the *Washington Post* and *The Wall Street Journal*, which could make this incident tomorrow's front-page news.

The security executive gives you until 7 a.m. to answer the following questions:

- Did a network breach truly occur?
- If so, how did the hackers get in?
- What systems were affected?
- What, if any, data was stolen?

Oh, yes — he also wants to know how you're going to make sure that this never happens again.

You hang up the phone; throw on some clothes; grab your car keys and your computer; and head off to work, leaving tire tracks in your driveway. When you get to the office, however, your security products show no signs of intrusion. How can you be 100 percent sure that your network wasn't compromised?

This scenario may seem to be extreme, but situations like it have taken place countless times in recent years. Just ask the folks at Sony Pictures, Anthem, JPMorgan Chase, and Target.

Security analytics is the ultimate Big Data Security solution. It's always on, it's always watching, it knows everything, and it can save your hide in a scenario like this one.

Chapter 4

Use Cases for Big Data Security

In This Chapter

- Identifying the most common use cases for Big Data Security
- Increasing your network security posture
- Seeing how real-world organizations benefit

If you've read the first three chapters of this book, you probably have a better sense of why traditional security defenses fail — or at least aren't foolproof — and of how Big Data Security can help. In this chapter, I peel back the layers of the proverbial onion a little further to show you various use cases for Big Data Security.

Enterprises and government agencies rely on Big Data Security to achieve some or all of the following objectives:

- Incident response and forensics
- Situational awareness
- Cyberthreat detection
- Data loss monitoring and analysis
- Policy compliance verification
- Security assurance

I explore each use case in its own section, discussing tangible benefits that you can expect to achieve when rolling out a Big Data Security solution in your own organization.

Incident Response and Forensics

Incident response and forensics is arguably the most popular (and perhaps most obvious) use case for Big Data Security. When your traditional security defenses trigger an alert or you discover a cyberthreat through your Big Data Security solution, you need to respond — *fast.* Every second that ticks away could mean the loss of more data or the compromise of another host.

When a security event is identified by an intrusion prevention system (IPS), next-generation firewall (NGFW), or malware analysis appliance, the event that triggered the alert usually doesn't represent your network's first contact with this threat. To ensure that the incident doesn't continue and can't happen again, you must determine the root cause of a security event.

The effects of an incident are often related to the scope of the systems and data accessed in an attack. Here are a few ways that Big Data Security can help you recover from these effects:

- **Tracking the source:** Most traditional network security tools can't determine the scope of an incident on your network; they only alert you at the moment when the threat was detected. By contrast, Big Data Security solutions allow you to track the pathway of a threat by fingerprinting files and searching all network data.
- **Monitoring all network connections:** After an incident occurs on your network, how can you be sure that it's over? Many organizations that suffer high-profile breaches deal with the lingering effects of those breaches for months or years.

  With Big Data Security, however, IT security personnel have real-time situational awareness, so they know whether an incident is over, whether attackers are still present on the network, and whether any machines are still compromised. By monitoring the connections among machines and out to the Internet, Big Data Security makes it possible to prove that your network is secure.
- **Collecting forensic evidence:** If you're lucky enough to identify the culprit by name, your Big Data Security solution can assist law-enforcement computer forensics officers by collecting digital evidence that may be used to prosecute the alleged perpetrator in a court of law.

Loading the ultimate incident-response weapon: CSIRT

Virtually every Global 2000 enterprise and large government agency has a team of cybersecurity professionals called CSIRT (pronounced "see-sirt"), which stands for Computer Security Incident Response Team. Although the complete list of CSIRT responsibilities varies by organization, one task common to all CSIRT teams is incident response.

When a member of the organization's IT department (often, someone from the help desk) calls the CSIRT hotline, a CSIRT incident handler is dispatched within minutes to investigate the situation. When that person arrives and confirms the cyberthreat, he attempts to answer the five dreaded questions of incident response:

- Do we know who did this to us?
- How did they do it?
- What systems were compromised?
- Can we be sure that the attack is over?
- Can we be sure that it won't happen again?

If you use traditional security tools, answering these questions is challenging, because you have access to only a subset of the information you need. Sure, log data aggregated by a security information and event management (SIEM) system is helpful for piecing the puzzle together, but you can't derive packet payload from log files.

Think of reviewing a SIEM report as being like reviewing your phone bill. You can certainly see when (long-distance) calls were made, including the duration of each call and the number that was dialed. But there's no way to confirm who was actually speaking and what the parties were discussing. Implementing Big Data Security is like (legally) tapping the phone line to record everything that was said — 24 hours per day, 7 days per week.

A Big Data Security solution built on security analytics (see Chapter 3) records every packet that traverses the network. CSIRT incident handlers have access to every packet captured, which can provide definitive evidence of an attack; this evidence allows members of the CSIRT team to make accurate conclusions rather than educated guesses.

Today, a CSIRT team without Big Data Security is like a team of air-traffic controllers equipped with just binoculars. If you're not using Big Data Security to its fullest potential, you're asking for trouble.

Situational Awareness

As security analytics platforms capture every packet that traverses the network (or every packet that it's configured to see), they provide a treasure trove of host and network intelligence that security analysts can use to achieve unprecedented situational awareness.

Suppose that an intrusion detection system (IDS) located on the network's perimeter detects a Windows-based exploit targeting an order entry system located in the DMZ (demilitarized zone). Should the security analyst who witnessed this alert respond? The answer: "It depends."

If that order entry system is running on a Linux server, the answer is no, because that Windows-based exploit can do no harm to a non-Windows host. If that system is running on Windows Server 2012 and is vulnerable to the exploit, however, the answer is a resounding yes.

Big Data Security enables security analysts to make more-informed decisions so they can better prioritize their incident response efforts.

Cyberthreat Detection

In Chapter 2, I discuss the merits of a defense-in-depth strategy. Big Data Security plays an important role in that strategy, especially in detecting advanced cyberthreats.

Zero-day malware, stealth botnets, and advanced persistent threats (APTs) have dominated headlines recently. The trend is clear: Hackers are no longer motivated just by the prospect of fame or the thrill of vandalism. They're focusing on economic benefit and even information warfare. Targeted attacks, engineered and carried out by sophisticated threat actors, jeopardize the economic welfare of virtually every organization. These types of attacks are designed to be difficult for traditional network security tools to detect.

Big Data Security can mitigate advanced cyberthreats before, during, and after attacks.

Before an attack

Before an attack, Big Data Security can help you do the following:

- **Gain situational awareness.** The solution familiarizes you with the types of traffic on your network so that you can recognize out-of-the-ordinary communications.
- **Reduce your network's attack surface.** The solution identifies applications, communications, and operating systems that pose a security risk and/or aren't approved for use in your organization.

During an attack

During an attack, Big Data Security can help you do the following:

- **Detect the threat.** Identify anomalous communications, such as an internal host connecting to an outside host for unusually long periods, an internal host transmitting an abnormally large amount of data, or an end-user host (desktop or laptop) communicating with other end-user hosts rather than servers.
- **Identify rogue hosts.** *Rogue hosts* (computers planted inside the organization for nefarious reasons) are clearly outside the operating system and/or application parameters set by your IT department.
- **Quarantine the threat.** The solution identifies other hosts that may have been compromised so you can quarantine them for remediation.

After an attack

After the attack, Big Data Security can help you do the following:

- **Verify attack termination.** Verify the attack has ended and confirm whether any lingering threats need to be remediated.
- **Confirm exfiltrated data.** Determine the scope and extent of the data breach.

- **Identify the root cause.** Understand exactly how the breach happened so you can ensure that it doesn't happen again.

As I discuss in Chapter 2, traditional security defenses simply aren't equipped to mitigate advanced cyberthreats. Big Data Security is your secret weapon for detecting these advanced threats, so use it — and use it often.

Data Loss Monitoring and Analysis

Data breaches are occurring at an alarming rate. As I mention in Chapter 1, the average cost of a data breach in 2014 was $3.8 million, according to Ponemon Institute's 2015 report on the topic. Cleaning up after a breach is often the most expensive part of the process, and it's directly related to the type of information and number of records exposed.

If you don't know what was exposed, however, the cleanup job gets even more difficult and even more expensive. Unfortunately, most organizations find themselves in this predicament. Log tools alone can't explain exactly what happened. In cases related to credit-card theft, organizations may be financially exposed to the maximum possible scope. Knowing what happened can save millions of dollars.

With Big Data Security at your disposal, you can protect yourself in the following ways:

- Monitor and record all files leaving your organization through web, email, and instant messaging traffic
- Monitor queries to your SQL database and relate them to the source endpoint
- Construct policies that extract, re-create, and take policy-based action on potentially sensitive files leaving the organization, such as forwarding such files to a data loss prevention (DLP) system for analysis
- Record all traffic to and from your critical data stores (like a video camera in a bank) and replay events if a breach occurs

Policy Compliance Verification

To IT security professionals, the word *compliance* is most often associated with some industry or government regulation. But the term also relates to organizations' internal policies that regulate the use of computers and the Internet. These policies are commonly referred to as *acceptable-use policies* (AUPs).

Some organizations combine multiple specific policies into one large AUP; others maintain a separate AUP for each topic. Regardless, AUPs often cover the following topics related to computers and the use of computer networks:

- Internet and social media
- Email and instant messaging
- Operating systems and applications
- Laptops and mobile devices
- Personal computing devices

Employees may wonder why their organization's IT department is so restrictive about computer use and Internet access. Is the chief information officer simply mean? Does the chief executive officer want to make sure that her employees are working every moment? Although I can't comment on the personalities or work habits of your organization's senior executives, I *can* tell you that the primary reason for implementing AUPs is to achieve one objective: reducing the network's surface area of attack.

Every computing device is a target for hackers. Virtually every operating system and application has at least one exploitable vulnerability. Vendors know about many of these vulnerabilities and may already have patched them, but some of them may be zero-day vulnerabilities, which I discuss in Chapter 1. The more freedom organizations grant their employees to select and customize operating systems, applications, and computing devices and to use the Internet, the less secure the organization's IT infrastructure will be.

How can Big Data Security help you monitor and enforce AUPs? Easy. Your Big Data Security solution sees everything,

so you have a real-time catalog of communications and applications in use on the network. You also can see when users violate AUPs by visiting unauthorized websites, downloading unauthorized content (such as pirated music and movies), and posting inappropriate content on social media websites during work hours from work computers.

Sure, you can realize definite employee productivity gains by enforcing AUPs. At the end of the day, however, by monitoring and enforcing your AUPs, you're reducing your network's attack surface, which results in reduced risk and fewer successful cyberattacks.

Security Assurance

In the days following a zero-day attack, vendors of signature-based network security tools (see Chapter 2) scramble to publish new signatures that defend against the new threat. But a dangerous window of vulnerability is wide open between the time when the zero-day attack is discovered and the time when your network security devices are updated with the new signatures. How can you be sure that your organization wasn't victimized by this zero-day attack?

Big Data Security gives you *security assurance* by enabling you to replay captured traffic during this dangerous period through security devices equipped with updated signatures. Replaying traffic helps you determine whether your network was previously victimized by the zero-day attack. If a security event related to the new signatures is triggered, you can leverage your Big Data Security solution to determine the scope of the attack and quarantine its effects.

When you're replaying captured traffic for inspection by your network security devices (IDS, IPS, or NGFW) after they're updated with new signatures, you'll want to work in a nonproduction environment, such as a lab, to prevent contamination.

Purchasing redundant network security devices for the purpose of security assurance is highly recommended. If budget is a concern, using freely available open-source solutions may be a viable alternative.

Big Data Security pays big dividends

Recently, the vice president of information technology of a multibillion-dollar Wall Street investment firm sought a solution that would better equip his incident response team to investigate and remediate security incidents before they could evolve into data breaches. Keenly aware that other large financial services firms — including JP Morgan Chase, Citigroup, and Heartland Payment Systems — have fallen victim to large-scale data breaches, this IT security executive knew that his status-quo security defenses simply weren't enough.

Having just found out about security analytics technology and the benefits of Big Data Security, the IT executive asked his staff to evaluate a leading solution from the company's existing secure web gateway (SWG) vendor, Blue Coat (`www.bluecoat.com`). Soon after, a Blue Coat Security Analytics Platform appliance was delivered for on-site evaluation.

Within days, the organization's IT security staff experienced the time-saving benefits this technology has to offer. Responding to a high-priority security incident used to take security analysts hours to validate. And after an incident was validated, it could take an incident responder days to terminate the threat and assess its impact. These durations were reduced to minutes of effort as everything security analysts and incident responders need is right at their fingertips.

After a few months, the organization expanded its use of the security analytics platform to preventive activities. Administrators constructed rules to monitor the network for potential violations of the company's internal AUP. By detecting unauthorized use of cloud-based file sharing applications, for example (a challenge commonly referred to as *shadow IT)*, the IT organization ensured the confidentiality and integrity of its data.

Chapter 5

Integrating Big Data Security

In This Chapter

- Integrating Big Data Security into your infrastructure
- Seeing real-world examples of integration
- Exploring universal connections

No IT security solution should operate in a vacuum. This rule applies to every endpoint and network security product on the market — and especially to Big Data Security. The era of static security products is over. Long live context-aware security and situational awareness!

You have many reasons to integrate Big Data Security into your existing security fabric. Here are a few of the most important things Big Data Security does:

- Accelerates incident response when time is of the essence
- Delivers context-aware security, providing deeper identity, application, content, reputation, vulnerability, and threat context to IT security teams at the point when a security and/or policy enforcement decision is made
- Provides situational awareness to offer unprecedented visibility into files, applications, and flows
- Reduces risk by detecting advanced threats that might otherwise go unnoticed
- Mitigates security threats by verifying that a cyberattack is truly over

In this chapter, I provide some real-world examples to get you started.

SIEM Integration

In Chapter 2, I briefly discuss the roles that a security information and event management (SIEM) system plays in larger companies and government agencies. Specifically, I'm referring to two roles:

- **Log aggregation,** which allows you to query all your log data in one place
- **Correlation,** which allows you to leverage prebuilt and custom rules to correlate security events

SIEMs are particularly useful for gluing together disparate pieces of information so you can understand security-related events. SIEMs work kind of like a jury in a criminal trial:

1. The jury listens to testimony by individual witnesses.

   In other words, the SIEM collects data from various security systems, such as intrusion prevention systems (IPSs), antivirus software, and data loss prevention (DLP) systems.

2. The jury weighs all the evidence.

   That is, the SIEM correlates the collected data.

3. The jury determines innocence or guilt.

   The SIEM reports on the presence or absence of a cyberattack.

For a SIEM user to weigh all the evidence about a potential cyberthreat, he must consult his Big Data Security solution by investigating host traffic related to the suspected threat (source and destination IP addresses) at the exact moment of the alleged attack. He could certainly log into his Big Data Security console and construct a new query, but by leveraging the integration between the Big Data Security solution and his SIEM, he can initiate that query directly from the SIEM console (see Figure 5-1), thereby saving valuable time and effort.

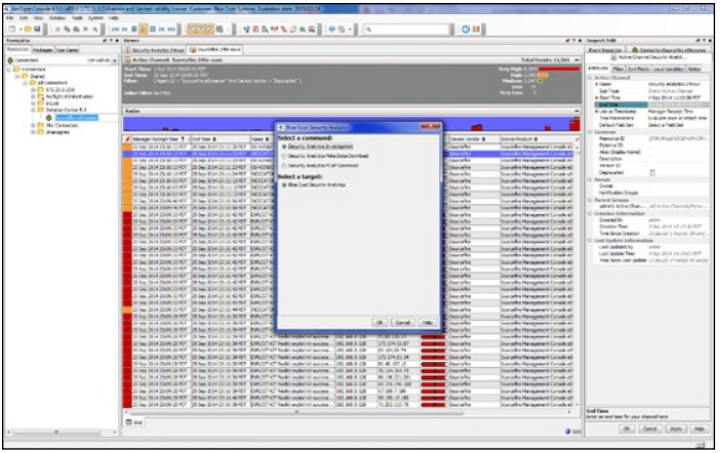

Figure 5-1: Big Data Security integration with ArcSight.

Integrating your Big Data Security solution into the consoles of your mission-critical security products isn't just a matter of convenience. When you're investigating a critical, high-impact security event, every second matters.

IPS Integration

As I discuss in Chapter 2, an IPS can be placed in *inline IPS mode,* meaning that it can block cyberthreats, and/or in *passive intrusion detection system (IDS) mode,* meaning that it can only provide alerts about detected threats. Over the past decade, more organizations have become comfortable placing IPS appliances inline, because improved detection technology yields far fewer false positives.

Even when an IPS is placed inline, not all signatures in a typical IPS detection policy are configured to block threats (as opposed to merely detecting them). Certain vulnerability-based signatures may be more likely to yield false positives than are more clear-cut exploit-based signatures, which trigger only on detection of known exploits.

When an IPS signature fires, triggering an intrusion event, security analysts must determine whether the cyberthreat is relevant to the operating system or application that it's

intended to attack. Then, if the threat is relevant, analysts must determine whether the attack was successful.

To accomplish these objectives, a security analyst must query the Big Data Security solution so that she can analyze traffic before, during, and after the suspected attack was detected. Specifically, she needs to analyze packet captures associated with the source and destination IP addresses involved in the intrusion event precisely when the event was registered.

This user could connect to the Big Data Security solution's console manually and then configure a new query, but she'd be losing precious time when every second counts. A better approach would be to connect to the Big Data Security solution directly from the IPS management console and review the details of the intrusion event.

By leveraging application programming interfaces (APIs) from both your Big Data Security provider and your IPS vendor, you can streamline security analysts' workflow, which saves time, saves effort, and reduces the effect of successful cyberattacks.

NGFW Integration

A next-generation firewall (NGFW; see Chapter 2) is a multifunction security appliance equipped with firewall, IPS, and application control technologies. When it comes to detecting cyberthreats, the IPS component of an NGFW operates quite similarly to a stand-alone IPS. Users can launch preconfigured queries directly from the NGFW console through corresponding NGFW and Big Data Security integrations.

Endpoint Forensics Integration

Cautious IT security organizations are augmenting their traditional host-based security defenses with endpoint forensics software that detects advanced threats and helps security analysts remediate them. To accelerate the remediation process, leading endpoint forensics solutions integrate with Big Data Security consoles through the same methods used by

IPS and NGFW solutions. Automated date, time, and IP address queries can be triggered in the security analytics platform directly from the endpoint forensics console, saving time and effort. Additionally, endpoint forensics tools can be triggered when certain network activity is identified that affects an endpoint.

Malware Analysis Appliance Integration

Enterprise acquisition of malware analysis appliances has risen dramatically in recent years. The capability to analyze files within the safety of virtual sandboxes makes this technology particularly useful for detecting unknown, targeted, and zero-day threats.

Analyzing files takes anywhere from 30 seconds to a few minutes, depending on the malware, its intended effects, and the complexity of the analysis approach. During this process, the file or URL in question reaches its final destination, as when a user downloads a malware-infected file from an untrusted website. When the malware has been identified, the console of the malware analysis appliance (and even the security analytics appliance) can issue an alert with a risk score indicating severity and prioritization.

By integrating your Big Data Security solution with your malware analysis appliance, you can launch preconfigured network forensics queries straight from the appliance's console (see Figure 5-2). Just like IPS and NGFW console integration, malware analysis appliance console integration saves security analysts considerable time and effort.

In addition, better Big Data Security offerings allow you to configure threat detection policies. You can automatically extract, reconstruct, and redirect suspicious files stored within the security analytics platform to your malware analysis appliance for analysis, thereby providing a last line of defense against advanced threats.

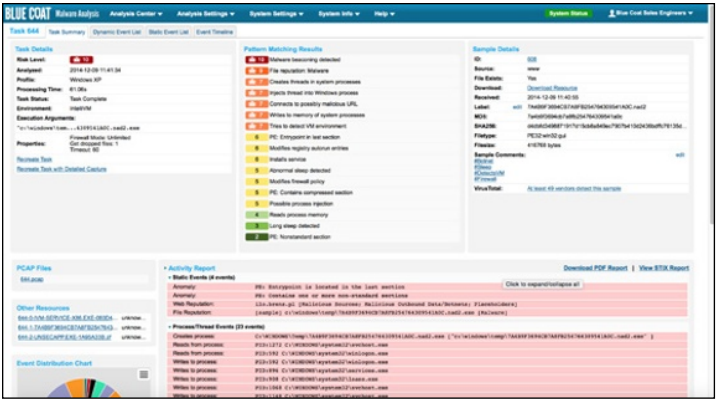

Figure 5-2: Integration with Blue Coat Malware Analysis Appliance.

Universal Connectors

Each of the Big Data Security integration examples mentioned in the preceding sections involves APIs from Big Data Security providers and network security vendors. If your network security vendor doesn't offer such an API, are you out of luck?

Not necessarily. Some leading Big Data Security vendors provide a *universal connector* — a small piece of software that installs directly in your web browser as a plug-in. When you install this plug-in and launch the browser, a small sidebar like the one shown in Figure 5-3 appears. While you stay on the web page that's hosting your network security tool's console, you can send queries about any event to your Big Data Security solution for further analysis.

Although this technique is by no means as seamless as integrating access to your Big Data Security solution into your network security solution's console (by using corresponding APIs), a universal connector is the next-best thing, because it simplifies the process of configuring search queries in your Big Data Security solution.

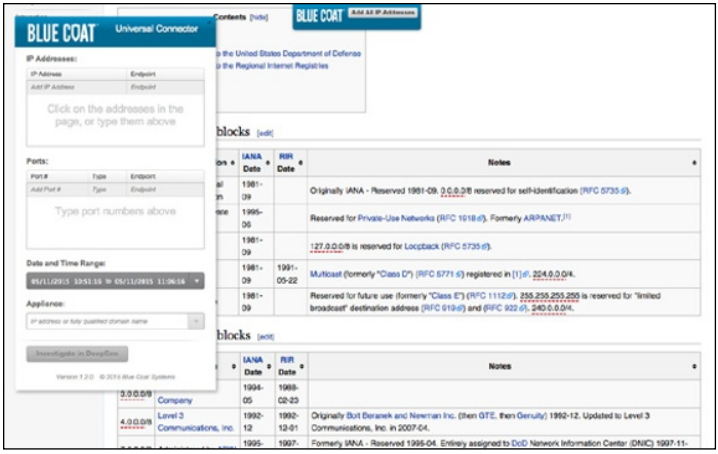

Figure 5-3: Big Data Security universal connector from Blue Coat.

Lottery commission hits security jackpot

The chief information officer of a large U.S. state lottery commission recently challenged his staff to find new ways to strengthen the organization's network defenses. He wanted new ways to detect the presence of advanced threats and to assess the causes and effects of successful attacks.

At the same time, the commission wanted to refresh its web security strategy by adding a more modern secure web gateway. After evaluating three leading web security vendors, the commission chose Blue Coat (`www.bluecoat.com`).

The commission selected Blue Coat's ProxySG for web security, its Content Analysis System and Malware Analysis Appliance for mitigating advanced threats, and its Security Analytics Platform for network forensics. The latter component not only reduced the time it took security analysts to validate and investigate security incidents, but also helped the organization monitor compliance with internal acceptable-use policies.

After a few months, the IT security architect integrated the company's Security Analytics Platform with its NGFW and SIEM platforms. Integration allowed security analysts to jump to forensic data (such as packet captures and session information) relevant to security incidents at the precise moment when the incidents were triggered. This technique dramatically simplified the process of investigating security alerts and reduced time to resolution.

Chapter 6

Ten Buying Criteria for Big Data Security

In This Chapter

- Avoiding inferior Big Data Security offerings
- Creating a checklist of buying criteria
- Understanding what to look for in a solution

Not all Big Data Security solutions are created equal. The performance, scalability, and capabilities of solutions vary considerably, as do the quality of the vendors.

In this chapter, I provide ten buying criteria to consider in evaluating Big Data Security solutions. First, though, here are a few pitfalls to avoid:

- Solutions that sample packets rather than capture every packet that traverses your network
- Solutions that can't keep up with the speed of your network
- Solutions that don't provide the flexibility of software-based and virtual appliance options (thereby tying you to a single vendor for hardware, software, and storage)
- Solutions that don't integrate with your existing security infrastructure
- Solutions that take a rocket scientist to configure and a packet scientist to use
- Vendors that lack a track record of success in enterprise, defense, agency, and public-sector market segments

Now that you know what not to look for, read on to review the attributes of a Big Data Security solution that you *should* look for.

Deployment Flexibility

Although most organizations prefer to acquire Big Data Security solutions that incorporate purpose-built hardware appliances, some prefer the flexibility of selecting their own hardware platform, perhaps to accommodate larger storage arrays than the vendor can provide or simply to reuse existing hardware. Also, some government agencies (especially military) require all software to be installed on hardware from an approved-products list.

By selecting a Big Data Security vendor that offers hardware, software, and virtual appliance versions, you avoid being tied down to proprietary hardware and maintain deployment flexibility based on your business requirements.

24/7 Full-Packet Capture

Some aspiring security analytics offerings capture only sample traffic, because the hardware doesn't have the horsepower to perform full-packet capture at speeds up to 10Gbps. Other rudimentary offerings may have enough computing power, but they perform only statistical sampling of data to speed report generation. And others only capture traffic after an alert is triggered, eliminating evidence of what happened before the alert.

Although packet sampling and selective data capture can improve performance, the risk of missing important packets for accurate, reliable threat detection is too substantial. You wouldn't buy a video surveillance system that does nothing but take still photos every 10 seconds. Be sure to limit consideration of security analytics solutions to those that capture and analyze packets in their entirety and can do so around the clock.

Deep Packet Inspection

A high-quality Big Data Security solution should be capable of performing deep packet inspection at Layers 2-7 of the Open Systems Interconnection (OSI) model. Most important, users should be able to view and analyze their data with the full context afforded by such metadata attributes as username and application name.

Accurate data categorization is of utmost importance when investigating security incidents, enabling users to find data quickly instead of having to do manual packet analysis to identify attributes, such as usernames and applications. It's also important to the cause of enforcing compliance with an organization's acceptable use policies (AUPs). Many organizations, for example, have policies that permit Facebook browsing from the office but not posting messages or sending messages within that application. Only the best Big Data Security solutions can differentiate these activities within a single application like Facebook.

Enterprise Performance and Scalability

A Big Data Security solution could have every possible feature under the sun, but if it can't collect and analyze data at the speed of your network, you're out of luck.

In addition, your solution must scale with your organization as it grows (or decides to monitor more network segments). To achieve the enterprise-class scalability that you require, you may need to acquire multiple security analytics appliances. If so, be sure to select a solution with a central management console that can aggregate data from your underlying appliances, providing aggregated views that make it easy to construct dashboards, reports, and alerts.

Virtual Platform Visibility

Physical security analytics appliances can't capture and analyze VM-to-VM (virtual machine to virtual machine)

traffic, so select a security analytics platform with virtual appliance software that integrates directly into your virtualization platform's virtual switch. Data captured by the virtual appliance should be accessible from the central management console for centralized aggregation, monitoring, reporting, and alerting.

Content Reconstruction and Replay

In Chapter 3, I describe how security analytics platforms reconstruct content for viewing in its original form — anything from documents and images to chat messages and emails. Select a security analytics solution that offers full-featured content reconstruction to maximize the effectiveness of your staff's investigation and visibility efforts.

Better security analytics offerings allow you to configure policies that automatically replay reconstructed content to a third-party network security device — such as a next-generation intrusion prevention system (IPS), data loss prevention (DLP) system, or a malware analysis appliance — for further analysis.

Global Threat Intelligence

Security analytics solutions have evolved from tools that investigate threats to tools that also detect threats. Modern security analytics platforms offer continuously updated global threat intelligence that feature IP, URL, file reputation feeds, and geolocation intelligence.

If a security analytics platform that you're evaluating doesn't offer global threat intelligence — backed by large-scale threat intelligence communities — for the purpose of identifying threats, move on. Otherwise, you're investing in a solution that only pays partial dividends.

Extensive Third-Party Integration

You should integrate your Big Data Security solution into your existing security infrastructure whenever possible to accelerate investigations and detect cyberthreats. Common Big Data Security integrations embraced by IT security organizations include the following:

- Security information and event management (SIEM)
- Intrusion prevention system (IPS)
- Next-generation firewall (NGFW)
- Endpoint forensics
- Malware analysis appliances

For more on this topic, see Chapter 5.

Select a Big Data Security provider that can integrate its solution into popular network security platforms. At the very least, select a vendor that offers a universal connector (see Chapter 3), which makes it easy to submit queries to your Big Data Security solution through your network security product's web-based console.

Ease of Use

No matter how powerful, scalable, and feature-rich a Big Data Security solution is, it's practically worthless if users can't figure out how to use it. Be sure to consider only Big Data Security solutions that feature easy-to-use, customizable dashboards, reports (see Figure 6-1), and alerts. Such solutions should also make it simple to locate and analyze traffic flows of interest and piece together evidence to help incident responders determine the root cause and material impact of successful attacks.

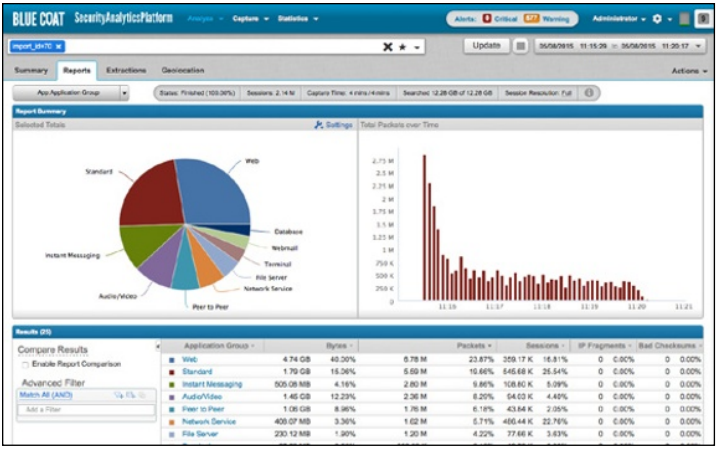

Figure 6-1: Sample Big Data Security report.

Responsive Customer Support

Selecting a Big Data Security vendor is just as important as selecting Big Data Security products, if not more so. Be sure to select a vendor that focuses on security instead of offering it as a byproduct of network forensics or network performance monitoring.

Find an excuse to contact the customer-support department on multiple occasions during the evaluation phase. Consider how quickly the vendor responds to each phone and/or email inquiry and whether the issue was resolved to your satisfaction.

Glossary

acceptable-use policy (AUP): A set of internal rules that restricts the way end users can use company-owned computer, network, and Internet resources.

advanced persistent threat (APT): A sophisticated cyberattack that employs advanced stealth techniques to remain undetected for extended periods, usually targeting a government or commercial entity for espionage or long-term reconnaissance.

baiting: A social-engineering attack in which physical media (such as CD-ROMs or USB flash drives) containing malware are deliberately left near a targeted organization's facilities, where they may be found and later installed by curious victims.

Big Data: A collection of data sets so large and complex that they're awkward to work with in traditional database management and analysis tools.

Big Data Security: A computer-based solution that captures and stores some or all of an organization's Big Data sources to uncover and mitigate cyberthreats.

cyberwarfare: Politically motivated hacking to conduct sabotage and/or espionage against another nation.

data loss prevention (DLP): A solution that detects and in some cases prevents potential data breaches by monitoring data in use, in transit, and at rest.

defense in depth: A strategy that involves installing a series of cybersecurity defenses so that a threat missed by one layer of security may be caught by another.

drive-by download: A form of cyberattack that occurs when an unsuspecting user visits a compromised website. The website downloads malware to the victim's computer without the victim's ever noticing. See also *malware*.

false negative: In the context of information security, malicious content that has been misclassified as benign.

false positive: In the context of information security, benign content that has been misclassified as malicious.

hacktivism: The use of computers and computer networks to protest and/or promote political ends.

intrusion detection system (IDS): A passive device or software application that monitors network traffic and provides alerts when it detects cyberthreats.

intrusion prevention system (IPS): An active (inline) device or software application that monitors network traffic and blocks cyberthreats upon detection.

malicious insider: A current or former employee, contractor, or other person with authorized access to an organization's computing resources who intentionally uses that access to harm the organization.

malware: Malicious software created to disrupt computer operation, gather sensitive information, or gain access to private computer systems. See also *Trojan*.

malware analysis appliance: High-performance appliance designed to automatically analyze suspicious files in an effort to uncover malware by executing those files within the safety of a sandbox virtual machine environment.

network behavior analysis (NBA): A cybersecurity solution that continuously monitors flow data from routers and switches to detect anomalous network behavior.

next-generation firewall (NGFW): A multifunction security device, typically marketed to enterprises, that bundles firewall, IPS, application control, and URL filtering technologies into one platform.

phishing: An attempt to acquire personal information (such as usernames, passwords, and credit-card details) by masquerading as a trustworthy entity.

ransomware: A type of malware designed to restrict access to the victim's computer system and demand that ransom be paid to the perpetrator. See also *scareware.*

sandbox: One of many virtual machines in a malware analysis appliance used to detect the presence of malware and other cyberthreats.

scareware: A simpler, less severe type of ransomware that displays bogus antivirus or clean-up tools that claim to have detected high-risk threats and demand that the user pay to remove them.

security analytics: Security software installed on high-performance, purpose-built appliances designed to aggregate and analyze every packet that traverses the network to uncover advanced threats and streamline incident response.

security information and event management (SIEM): A solution that aggregates and correlates log data from network security and network infrastructure devices to provide analysis of security events.

spearphishing: An attack on a specific organization or a specific person within that organization.

SSL visibility appliance: High-performance appliance designed to inspect SSL (and TLS) encrypted traffic, enforce SSL usage policies, decrypt SSL traffic for inspection and/or data capture by one or more network security devices, and reencrypt (benign) traffic post-inspection and/or data capture.

Trojan: A type of malware that masquerades as a legitimate file or application while granting a hacker unauthorized access to a computer.

unified threat management (UTM): A multifunction security device, typically marketed to smaller organizations, that bundles firewall, IPS, antimalware, URL filtering, and other security technologies into one platform.

watering-hole attack: A cyberattack that occurs when a hacker infects a website with malware known to be frequented by his targets. Visitors become infected by downloading content or simply by connecting to the website.

zero-day attack: An attack on an unknown operating system or application vulnerability. The attack occurs on "day zero" of awareness of the vulnerability, when neither a patch nor a threat-detection signature exists.

Protect the Web. Secure the Cloud. Manage Encrypted Traffic. Stop Advanced Threats.

Leverage Big Data Security to uncover advanced threats and streamline incident response!

If you're charged with securing your organization's network or responding to security incidents, this book is for you. Enterprises are turning to Big Data Security as the newest weapon to fight cybercrime, collect digital evidence, and uncover advanced threats that traditional security defenses miss.

- ***Defining Big Data Security* — explore Big Data Security form factors, features, and deployment strategies**
- ***Reviewing use cases* — look at common use cases for improving network visibility and strengthening your security posture**
- ***Integrating Big Data Security* — understand how to integrate Big Data Security into your existing security fabric**
- ***Establishing buying criteria* — know what to look for and what to avoid when evaluating Big Data Security solutions**

Open the book and find:

- **Lists of common internal and external Big Data sources**
- **Network diagrams depicting typical deployment strategies**
- **Strategies for detecting advanced threats and targeted attacks**
- **Real-world examples of Big Data Security implementations**
- **Ten buying criteria for evaluating Big Data Security solutions**

Go to Dummies.com®
for videos, step-by-step examples, how-to articles, or to shop!

Steve Piper is a high-tech veteran with more than 20 years of experience. An award-winning writer and consultant, Piper has written more than a dozen books on IT security, networking, and Big Data. He holds a CISSP security certification from ISC² and Bachelor of Science and MBA degrees from George Mason University. Follow him on Twitter at @StevePiper or find out more at www.stevepiper.com.

FOR DUMMIES® A Wiley Brand

Also available as an e-book

ISBN: 978-1-119-13154-0
Book not for resale

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.