LEARNING MADE EASY

Netskope Special Edition

Modern Data Loss Prevention (DLP)

for dummies®
A Wiley Brand

Learn modern DLP techniques

Use zero trust principles to protect data where it moves

Achieve better cloud security

Brought to you by

netskope

Carmine Clementelli

About Netskope

Netskope, a global SASE leader, is redefining cloud, data, and network security to help organizations apply zero trust principles to protect data. Fast and easy to use, the Netskope platform provides optimized access and real-time security for people, devices, and data anywhere they go. Netskope helps customers reduce risk, accelerate performance, and get unrivaled visibility into any cloud, web, and private application activity. Thousands of customers, including more than 25 of the Fortune 100, trust Netskope and its powerful NewEdge network to address evolving threats, new risks, technology shifts, organizational and network changes, and new regulatory requirements. To learn how Netskope helps customers be ready for anything on their SASE journey, visit netskope.com.

We would like to thank a number of individuals who, along with the author, made this book possible:

From Netskope: Amanda Anderson, Chad Berndtson, Jason Clark, Scott Hogrefe, Kathy Jacobsen, Naveen Palavalli, Stephenie Pang, Lauren Polito, Carolyn Robinson, Neil Thacker

From Evolved Media: David Penick, Karen Queen, Evan Sirof, Lauren Wagner, Dan Woods

Modern Data Loss Prevention (DLP)

Netskope Special Edition

by Carmine Clementelli

for dummies®
A Wiley Brand

Modern Data Loss Prevention (DLP) For Dummies®, Netskope Special Edition

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

ISBN 978-1-394-19891-7 (pbk); ISBN 978-1-394-19892-4 (ebk)

For general information on our other products and services, or how to create a custom *For Dummies* book for your business or organization, please contact our Business Development Department in the U.S. at 877-409-4177, contact info@dummies.biz, or visit www.wiley.com/go/custompub. For information about licensing the *For Dummies* brand for products or services, contact BrandedRights&Licenses@Wiley.com.

Publisher's Acknowledgments

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

Project Editor: Elizabeth Kuball
Acquisitions Editor: Traci Martin
Editorial Manager: Rev Mengle
Client Account Manager: Jeremith Coward
Production Editor: Mohammed Zafar Ali
Special Help: Nicole Sholly

Introduction

Data protection as a concept in cybersecurity is not new, but the demands placed on legacy data protection systems have changed drastically in the past decade. Security professionals were once confident that the valuable data they protected was safely tucked away inside heavily fortified data centers. But digital transformation entails that businesses, large and small, move their data to the cloud and across distributed locations. Your data now finds its way everywhere users are, wherever that is. Your business may share digital connections to huge numbers of third- and even fourth-party suppliers, partners, and contractors. These scenarios bring both unprecedented business opportunities (good news) and security challenges, particularly regarding data protection (not good news).

Successful breaches can have devastating consequences for a business. The risks from insiders (whether malicious or negligent) are as dangerous to your business as headline-grabbing attacks by outside actors. All threaten to expose sensitive information. Data protection is now a cornerstone of compliance rules, with industry and data privacy regulations that specifically detail your business's responsibilities and significant penalties for failure.

Companies must take a new approach and apply data protection policies everywhere their data goes — consistently. Ideally, data protection supports business goals while also protecting the business. But managing data protection policies and the tools needed to enforce them can be complex and costly. Organizations need data protection solutions that simplify policy enforcement while ensuring policy effectiveness. A new generation of cloud-delivered data loss prevention (DLP) solutions offers a possible way forward. Organizations must adopt a cloud-delivered solution that is less complex, highly scalable, and more cost-effective while, ideally, protecting data with higher reliability and better accuracy, and minimizing exposure to unauthorized access or misuse. It's a tricky balance to strike, but you can achieve it today with the right guidance.

About This Book

This book can prepare you to make informed decisions about how to evaluate your organization's current approach to data protection and evaluate new data protection solutions to find the best fit for your needs, using zero trust principles to guide how security is contextually and consistently applied. By explaining how modern, cloud-delivered DLP systems work, this book cuts through the marketing clutter to identify the characteristics and capabilities needed to reliably protect your data anywhere it may be used.

Foolish Assumptions

This book assumes you have a baseline knowledge of how businesses have embraced the use of cloud computing to make themselves flexible and better equipped to embrace digital transformation. It also assumes you're here because you want to ensure the right mix of technology and process improvements to protect sensitive data wherever it resides and wherever it moves in your computing environment.

Icons Used in This Book

We use icons to call attention to important information. Here's what you can expect:


TIP

Anything marked with the Tip icon is a shortcut to make a specific task easier.


REMEMBER

The Remember icon flags facts that are especially important to know.

TECHNICAL STUFF

When we offer up highly technical info that you can safely skip, we use the Technical Stuff icon.

WARNING

Heed anything marked with the Warning icon to save yourself some headaches.

Beyond the Book

Although this book is chock-full of information, if you find yourself at the end of it thinking, "Where can I learn more?," just go to `www.netskope.com`.

IN THIS CHAPTER

- Understanding where sensitive data is stored and how it's monitored
- Discovering what data protection is really about
- Learning about data loss prevention (DLP)
- Digging into why legacy DLP is no longer a viable solution
- Switching to a cloud-first strategy with a modern DLP solution
- Dispelling common myths about DLP

Chapter 1
Sensitive Data Is Everywhere and Is Harder to Find

In general, when people talk about sensitive data, they're referring to information that is confidential or personal in nature. What is sensitive depends a lot on whether you look at data from a business perspective or an individual perspective.

A Quick Guide to Sensitive Data

You may notice that most data that's labeled as sensitive has been around in some form for years, decades, or even longer:

- Personal data/information such as Social Security numbers, credit card numbers, driver's license numbers, health details, and home addresses

- Intellectual property (IP) such as product designs, new inventions, patents, and source code
- Confidential information and trade secrets such as financial plans, contracts, tax reporting, mergers and acquisitions (M&A) information, and prerelease documents such as press releases

What's new is that the modern business landscape has completely changed the way data is shared and (yikes!) exposed. Many companies, especially following the onset of the COVID-19 pandemic, now embrace a hybrid work environment.

Nearly every type of sensitive data is created, stored, and moved digitally. Data travels to and from cloud services, corporate networks, and anywhere else users can access it. At the same time, an ever-growing number of applications store and share that data across multiple platforms, making it accessible from virtually any device in remote locations. As the amount, variety, and speed of data increases exponentially, it becomes increasingly hard to identify and protect sensitive information. To make matters worse, the sheer volume of available data makes it difficult for traditional security measures to keep up with constantly new threats.

A Tidal Wave of Data

By 2025, according to IDC, the world will be awash in as much as 181 zettabytes of data! An enormous portion of that will be created and stored directly in the cloud — more each passing year. Among the challenges data protection systems and their operators face as a result are

- **Too many categories of sensitive data:** An increase in data privacy regulations and laws that protect a wider range of individuals and types of information globally is driving massive growth in the categories of sensitive data. This includes information that can identify a person, such as their location, financial and health information, personal preferences, religious beliefs, and sexual orientation. Sensitive data includes things like national ID numbers, credit cards, source code, designs, financial plans, bank accounts, contracts, tax forms, passwords, M&A information, protected health information (PHI), confidential email, gender, and religion.

There are categories of sensitive data that differ from country to country, in localized languages, and that are specific to each country.

- **Too many data formats and types:** PDF, graphic images (like JPG, PNG, and BMP), compressed and encapsulated files (like ZIP, RAR, and ISO), attachments, Slack messages, chats, online forms, screenshots, spreadsheets, computer-aided design (CAD), social posts, text files, presentations, and email.
- **Too much context:** Context must govern a decision on how sensitive data should safely be accessed, used, transferred, and shared. Context helps define what a risky action would be around sensitive data and what should be considered a violation or a breach attempt: who, where, what, how, why, when, to whom, and other factors.

Faced with a surge of inscrutable data, legacy security systems are forced to err on the side of caution, which has increased administrative headaches by orders of magnitude. Why? Incident response security teams face barrages of false positives, most of which must be manually evaluated by already-beleaguered personnel.

Data Protection Is about Much More Than "Just" Data

Businesses need new automated strategies that can effectively identify, monitor, and protect their valuable data. At the same time, the world in which data protection operates continues to introduce new challenges that compound the security predicament. These new challenges include

- **More cyber risks:** Companies face more vulnerabilities to data breaches than ever before. These vulnerabilities can be both intentional and unintentional. Insider behavior, such as employees stealing or mishandling (oops!), is one way a company's sensitive information is at risk for exploitation. Eighty-two percent of data breaches involve the human element, which includes
  - *Malicious insiders:* For example, a disgruntled employee taking screenshots of a critical spreadsheet, sending data to a personal storage software as a service (SaaS) app instance,

or through a personal instance of a corporate email account (that is, personal Gmail versus corporate Gmail).

- *Unintentional exposure:* For example, an employee who inadvertently sends too much information to a vendor or negligently overshares files on a OneDrive folder. These are significant causes of data breaches.

Similarly, external attacks or hacking attempts also put company secrets in danger of being held for ransom or revealed to the public or rival organizations.

- **Cloud including SaaS and public cloud infrastructure as a service (IaaS):** The adoption of SaaS applications, in particular, is increasing at a stunning rate. According to recent studies, the average enterprise uses 2,400+ cloud applications, with 97 percent considered *shadow IT* (unsanctioned by, unknown to, or invisible to the IT department). This presents technical and security challenges because data can be stored and shared across a large number of SaaS apps, moves across corporate networks and managed devices, and can be easily accessed by employees and even by external users connecting from remote locations with unmanaged devices. Cloud apps can quickly become a primary attack vector if not properly monitored and managed. Businesses must take steps to upgrade their data protection solutions to protect against such threats.
- **Hybrid work:** The rise of the hybrid workforce is changing how companies store and access sensitive data. Things are drastically different from the days when companies kept most critical information within a private data center that the company had control over. Hybrid workforce arrangements have brought about a new era in which sensitive data is highly distributed in places beyond corporate borders that the company can't see and doesn't control. Nowadays, data is spread across a variety of environments, both digital and physical, including data centers, corporate headquarter workplace, branch offices, home offices, and remote workers' devices (corporate and personal).
- **New compliance requirements:** Compliance has always been a concern, but as businesses become more heavily regulated and data privacy legislation carries increasingly hefty fines and legal action, companies of all sizes are feeling the pressure to ensure they meet compliance standards and secure their sensitive data. Companies must take measures to meet industry-wide regulations such as the more popular

Payment Card Industry Data Security Standard (PCI-DSS), the Health Insurance Portability and Accountability Act (HIPAA), and the Gramm–Leach–Bliley Act (GLBA), while also ensuring they abide by applicable data privacy laws and regulations, including the General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA), the Colorado Privacy Act, the Connecticut Data Privacy Act, the Virginia Consumer Data Protection Act, and the Utah Consumer Privacy Act to name just a few. Many countries around the globe are regulated by privacy laws, including Brazil, Singapore, Japan, and the United Kingdom. Now more than ever, companies need to show that they're taking the necessary steps to protect their customers' personal information and comply with all relevant legislative policies or risk facing severe penalties.

- **Rare and costly talent:** The specialized, skilled talent needed to run complex data protection programs is in short supply. Data protection technologies require skillful oversight to tackle vast amounts of incidents triggered by the system. That problem compounds when legacy data protection systems monitor cloud services like SaaS applications (something they weren't initially designed for), causing a higher occurrence of false positives and, as a result, much additional work for the team. Based on the required skill sets, these skilled IT personnel command high salaries, which can amount to a hefty cost for businesses — whether they stay and must be paid or whether they get overworked, leave, and must be replaced.

What Is DLP and How Is It Meant to Help?

DLP security technologies are systems designed to automatically discover and protect the storage, the flow, and the use of sensitive data anywhere across an organization's networks, users, and services. The technology is implemented to detect a wide variety of sensitive data, such as personal data/information of customers and employees, financial documents, and intellectual property. DLP monitors how such data is accessed and used, preventing leakage, accidental exposure, and theft. DLP helps businesses mitigate their data breach risks and audit their files for accidental publication of confidential information. As the compliance landscape has gotten both stricter and broader, DLP has become an increasingly important security measure for businesses to protect themselves against costly data breaches and meet the demands of compliance legislation.

Why Legacy DLP Is Now Woefully Inadequate

Legacy DLP solutions have been used for data protection for more than ten years. However, over time, legacy DLP has gained a reputation for being complex to implement and manage, costly, limited in scope, less and less accurate, and not providing the comprehensive coverage needed for the current work-from-anywhere world. DLP solutions were designed to protect data within a data center and corporate premises. These solutions have struggled to adapt to the changes brought by the cloud era. Legacy DLP is good at what it was designed for, but it's now being asked to do a job it was never intended for: securing data in the cloud or moving across many clouds. In addition, its perimeter-based model can't keep up with data scattered across multiple locations and applications.

The downside of legacy DLPs

Legacy DLP systems, made of several software and hardware components, can be a pain to implement and sustain. Setup can be complex and expensive, not ideal for any company on a budget or with limited IT resources. Covering highly distributed enterprises is also a major and costly challenge because the on-premises DLP architecture most likely should be replicated at every branch office. And even then, the approach doesn't cover important modern requirements, such as remote employees, cloud, and bring-your-own-device flexibility.

Legacy DLP technologies also need lengthy software upgrades and continuous adjustments that create business disruptions that can't simply be waved away. Because of this disruption, organizations often avoid the upgrades; organizations can find themselves months or years behind on DLP versions, meaning they aren't using current protection against the latest data requirements, compliance, and risks.

Failing to update and patch a DLP system can lead to several issues no organization wants, including security vulnerabilities, data breaches, and inadequate data protection. This can put sensitive data in harm's way and bring an organization out of compliance with data protection regulations. Furthermore, the inherent complexity of legacy DLP often results in inconsistent

and unreasonably specific data protection practices — making inefficient use of resources and time.

WARNING

For some companies, the business disruption caused by their legacy DLP system is considered so severe that they'll shift their DLP systems into "monitor-only" mode, meaning the system watches what's happening but doesn't enforce data protection policy. Running your DLP without enforcing policy is like having a safe but leaving it unlocked and just hoping really hard that no one walks off with your cash, jewelry, and important papers.

The false-positive conundrum

Not only do legacy DLP systems impose complicated deployments and processes, but they also need a lot of resources and human labor to monitor and continuously tune them effectively. Earlier, I mention the pressure that false positives put on security teams, but it's worth looking at the situation in more detail.

The number of incidents to remediate manually has grown to a point where the incident response team isn't able to consider, let alone deal with, all of them. Incident response teams get a lot of alerts that aren't actually problems and lack context to determine their level of risk after the fact (basically, they get these alerts too late after an incident has occurred, so not only do the alerts not have context, but teams are also asked to figure out incidents that happened in the past, and even if they reach out to the employees that caused them, they wouldn't remember what happened). These alerts can number in the thousands or hundreds of thousands daily and arise from many different sources. Because so much is happening, security response teams simply can't look at all these alerts; in fact, they need to ignore many just to keep up.

A significant contributing factor is that data now resides and moves in and among many places outside the managed data center network. Legacy DLP solutions are not equipped to handle the ever-growing variety and amount of data and lack of newer machine learning assisted detection, modern data sharing use cases, and context awareness. Their static policies can't effectively adjust to changing business risks and contexts, such as who's using the data, how, in what environment and application instance, whether they exhibit safe behavior, and the final destination.

Cybersecurity automation and orchestration tools like user and entity behavior analytics (UEBA) have been added to help with some of this by taking in alerts and fixing them more quickly. However, if the DLP system is inaccurate, lacks business context and risk awareness, and has many gaps, then UEBA models won't work well.

To effectively protect sensitive data, a DLP system should be integrated and automated to continuously monitor and verify the identity of authorized individuals and devices, their behavior, their collaboration and external data sharing, the applications that they are using and their risks, and many other contextual factors. This zero-trust approach (see Chapter 3) allows for accurate policy recommendations and incident response rules that adapt to changing risk conditions and the specific business context in which the data is being used. Such an approach doesn't disrupt modern business practices but enables them with safety.

Legacy DLP is missing critical cloud coverage

Legacy DLP systems were designed with a perimeter-based security model that assumes that all data is stored within the corporate network and managed environments. This model is no longer sufficient in the cloud era where data is stored in multiple cloud-based locations and accessed by users and devices outside the corporate network. Additionally, legacy DLP systems may not have been designed to integrate with the wide range of cloud services and infrastructures that are now in use, making it difficult or impossible to provide comprehensive protection for data in the cloud.

Bolting on additional technologies such as cloud access security broker (CASB) and cloud-delivered secure web gateways (SWGs) to a DLP system deployed on-premises may provide some additional coverage for cloud repositories, but it won't address the fundamental limitations of the legacy system. Teams are further challenged having to address disjoined management consoles and uncoordinated data protection policies — two common side effects when CASB and SWG are bolted on to legacy DLP.

In other words, adding extra technologies to an outdated DLP approach doesn't make it cloud-ready, and it would only add complexity. A DLP system must be able to meet the ever-evolving standards of cloud security adaptively with its own

dynamic policies and real-time risk assessment capabilities, so businesses can keep their employees, customers, and data safe. Legacy DLP solutions are on-premises. Period.

REMEMBER

To protect data in the cloud, legacy DLP needs to elegantly integrate with cloud security solutions. Data in the cloud needs security in the cloud.

TECHNICAL STUFF

In most enterprises today, two cloud security solutions are usually combined with legacy DLP: CASB for cloud application traffic and SWG for web traffic from remote workers and branch offices. These solutions are designed for the cloud but usually have limited data protection capabilities. The hope is that integrating these solutions would provide legacy DLP with the "eye in the cloud" needed to extend their existing on-premises capabilities to the cloud and look for sensitive data outside the data center perimeter. Unfortunately, this integration has proven to be very difficult, involving network traffic redirections that rely on the very complicated Internet Content Adaptation Protocol (ICAP), which, thankfully, is beyond the scope of this book.

Even where integration is achieved, the approach proves not to be sustainable. For one thing, CASBs use application programming interfaces (APIs) to connect to corporate cloud applications like Microsoft 365, Salesforce, Slack, Zoom, Teams, Google Workspace, Amazon Web Services (AWS), and Box. These APIs provide the legacy DLP system with the desired window to look inside these cloud applications. So, for example, if there is sensitive data stored on Salesforce, the DLP can scan it and protect it. CASBs also use inline detection to look at data uploads and downloads across thousands of SaaS apps.

It's also tough to consolidate data protection policies between on-premises and cloud systems. For example, CASBs often can't duplicate the same policies as legacy DLPs can. Because these technologies don't have the same abilities, policies and management consoles become fragmented and out of sync.

The problem with this architecture is that integrating an on-premises DLP through CASB with an application in the cloud also creates a delay called *latency.* Latency means that even if your legacy DLP discovers a data violation in the cloud, it may take minutes, hours, or longer to mount a response. Think of this scenario: The violation has happened, it has been detected, but you still didn't stop it on time (meaning your data is compromised!).

Ultimately, combining legacy DLP with cloud technologies is like trying to combine two different animals. One is a cloud service (CASB), and the other is a massive on-premises deployment of hardware and software (legacy DLP). The result is a fragile chimera that is easy to break, causes a lot of latency, and is very hard to optimize and maintain. Ideally, you'd want to get rid of that complexity and make everything streamlined and simple, so there's less chance for problems.

Anchored to on-premises infrastructure and lacking the means to scale quickly and cost-effectively significantly limits the effectiveness of legacy DLP in cloud environments. The approach is simply no longer sustainable.


REMEMBER

For DLP to be effective, the focus must shift from the outer perimeter of your data set to the actual data itself and where and how it moves. Companies can no longer rely on legacy DLP strategies if they hope to protect their information in the cloud effectively.

DLP for the Cloud Era

Digital transformation has revolutionized how organizations deliver customer service and develop products and services. It has also had an immense impact on how data is secured. Big and small companies rely heavily on cloud technology to achieve business growth and enablement, so security strategies must keep up with these changes. The DLP architecture must accommodate the ever-growing hybrid workforce by switching to a cloud-first strategy to bring broader coverage, improved efficiency, scalability, powerful computing abilities, and more effective risk prevention measures. With a reconsidered DLP model in place, modern organizations can be successful in the hybrid work world and future-proof their enterprises. Modernizing your company's DLP is a big undertaking, but with the ever-evolving risks and advancement in cloud-ready DLP solutions, now is the right time to consider it.

With cloud-delivered DLP, you have nothing complicated to deploy, just a cloud service to enable. You don't have to deal with many components and software that you need to update and manually maintain. There are no more DLP databases to maintain or database experts to hire. There are no more DLP servers to

become obsolete and require replacement. And there are no more hardware proxies that need refreshing.

Cloud-delivered data protection platforms are designed to be easily integrated across security, networking, infrastructure, and cloud services while consistently gathering risk and organizational context from other controls. Data surveillance and detection algorithms work better in the cloud, where access to endlessly scalable resources reduces the load on your computing infrastructure while keeping pace with newer use cases and your innumerable and ever-increasing endpoint agents. You're no longer limited by an on-premises infrastructure, so your users are covered wherever they go.

Furthermore, because a cloud-delivered architecture is not tied to your infrastructure and schedule, your DLP remains up to date, with real-time updates available everywhere. This approach makes for a much more efficient tool for protecting your organization's valuable data.

Myth Busting

When it comes to cloud-delivered DLP, it's no secret that the market is saturated with buzzwords, inflated promises, and tech jargon — leading people to feel overwhelmed and confused by their options. But the truth is, not all DLP solutions are created equal. In this book, I help you distinguish between fact and marketing hype when weighing your choices, with a guide to important features and functionalities in each one.

So, let's take a step back and start by debunking some common myths around data protection delivered from the cloud so you can cut through the noise and make an informed decision tailored perfectly for your business.

Myth: New DLP is the best DLP

Reality: When it comes to data protection programs, you don't want to leave anything to chance. Not only do you need enough features within the program to ensure safety, but you also need a dedicated and knowledgeable vendor with established experience in DLP. Legacy solutions may not have been built with cloud

technology in mind, but they have lessons to teach about maturity to most cloud-delivered DLP solutions.

The most reliable data protection solution has undergone an extended maturation period and developed new features along the way. If you're considering investing in a comprehensive data protection program, make sure your vendor can meet all your needs — from cloud support to feature maturity — for maximum data security. The newest vendor solution shouldn't be confused with the best.

Myth: Legacy DLP was inaccurate

Reality: Legacy DLPs were built by vendors who invested a decade or more into developing accurate algorithms and policies to identify and prevent the unauthorized transfer of sensitive information.

Accuracy is not the real issue. The real issue, as I mention earlier in this chapter, is false positives. False positives can lead to a dangerous situation where real threats go unnoticed and sensitive data is accidentally leaked. It also leads to skilled (that is, expensive) incident response teams getting bigger and bigger to deal with a larger volume of incidents. In Chapter 2, I explore why DLP systems must be precise and accurate to maintain trust.

Myth: "Good enough" is enough when it comes to DLP

Reality: When it comes to making sure your company's data is safe and secure, don't cut corners. Are you thinking of using a cloud-delivered solution that promises "good enough" security? Think twice. You may end up with a reduced feature set or limited focus on only the most surface-level attack vectors and data types, which leaves you at risk of malicious activities, false positives, and inaccurate detection.

Invest instead in a modern, cloud-delivered DLP system that delivers high data detection accuracy, provides additional security layers, and ensures complete protection against possible threats to your business data or other confidential material. Don't play fast and loose regarding your company's data; make sure you invest in the right DLP system for maximum security and performance.

Myth: Cloud-delivered DLP is less capable than legacy DLP

Reality: Currently, many cloud DLP systems use fewer than 100 data identifiers (see Chapter 2) and scan only a few file types, which means they're barely detecting anything. The reason for this is a lack of maturity in the technology. Unlike DLP systems that have been developed and in play for a decade, these systems have been designed to focus on solving specific new use cases, such as particular cloud applications, and protect only a few popular types of files. This lack of broad focus means they still lack the accuracy needed to effectively balance data protection and business needs, leading to continued friction between the two. Cloud-delivered DLP technology should be superior to legacy DLP due to its ability to offer massive scale. You would think that with greater scale, you would be able to solve for false positives and improve accuracy.

WARNING

When it comes to data protection, the old saying is true: "Experience counts!" Although tempting new options may look great on paper or at first glance, mature DLP solutions can offer a deeper level of security and insight because they've been grown and refined over time. Go with a proven provider and test several systems yourself for total peace of mind when safeguarding your essential data.

Myth: A bundle of data protection systems is just as good as a complete, integrated data protection solution

Reality: When it comes to data protection, security initiatives and programs that attempt to bundle a variety of separate DLP products and services from different vendors may seem a logical progress. After all, DLP services may come already integrated with certain SaaS applications, public cloud services, firewalls and SWG solutions. But soon or later, these multi-service data protection programs will certainly fall short. In bringing together discrete systems that weren't developed together, the solution may offer little in the way of awareness regarding business context and risks. In addition, data protection practitioners will end up dealing with disjointed data protection policies and multiple consoles. In fact, the scope of each integrated DLP service is

often limited to specific environments and channels, covering for example only web traffic or specific control points such as one or a few SaaS apps. This will leave your data vulnerable after it's out in the open.

To protect yourself and your organization, seek fully integrated solutions that offer comprehensive data protection to cover all potential areas of risk across cloud services, on-premises locations, email services and end points, and get full coverage across multiple types of data and controls.

IN THIS CHAPTER

- Learning the challenges that legacy data loss prevention (DLP) faces
- Preparing to scale for possible future changes and growth
- Knowing the realities and limitations of cloud-based DLP
- Understanding how DLP makes other security tools more effective

Chapter 2
Protecting the Entire Cloud-Centric Enterprise

Why is it important for an organization's data protection systems to protect the entire enterprise, including its cloud applications? Because the loss or unauthorized access of data can have serious consequences for the organization and its stakeholders. This approach may seem like a no-brainer, but in practice, many forces work against this goal. In this chapter, I explain why achieving complete data protection throughout your entire enterprise is a process that delivers quick wins and long-term strategic benefits.

Enterprises without Borders

A decade ago, the concept of an enterprise was largely defined by the physical boundaries of a building or location. This typically included the employees, equipment, and resources contained within those walls. However, the definition of the enterprise has evolved over time to reflect the changing nature of business and technology, and the enterprise is no longer limited to a physical location.

With the rise in remote work, valuable data likely crosses the devices and the home networks of your employees. With the rise of cloud services, your data may be scattered across a variety of cloud locations, including software as a service (SaaS) apps like Microsoft 365 and Salesforce, as well as in online conversations on collaboration apps like Slack and Microsoft Teams (see Figure 2-1). The scope of the enterprise now includes the numerous endpoints that employees use to connect to corporate resources, as well as the thousands of approved and (ahem) unapproved cloud-delivered applications that may be used within companies.

WARNING

If you don't know that or where sensitive data exists, you can't protect it. If you know sensitive data exists but you don't know where it lives and travels, you still can't protect it.

REMEMBER

To ensure that all sensitive data is detected and protected no matter where it lives or travels, you need to take a comprehensive approach to detect and protect sensitive data. That means no gaps in coverage or blind spots where data could be exfiltrated or accidentally exposed without your knowledge.

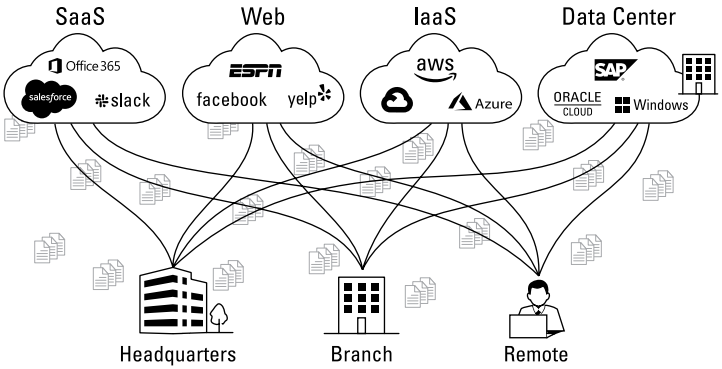

FIGURE 2-1: In the modern highly distributed enterprise, data resides and flows across many new environments.

The Challenge DLP Faces as It Evolves

As I discuss in Chapter 1, DLP systems have been primarily focused on protecting data that was stored *within* a corporate data center. Today, it's important to protect data everywhere it may go, whether that be in the cloud, on remote devices, through the

corporate network, or across external locations. This means that legacy DLP systems, which were designed to protect data within the enterprise, are no longer sufficient.

REMEMBER

Although you do need to identify all the locations where data lives and moves, by shifting your data protection's focus to the data itself rather than the locations where data is originated and held, you can gain huge advantages in flexibility and effectiveness. Think of it like a basketball team shifting from a zone defense to a man-to-man defense. As I explain later, by taking this comprehensive approach, you can safeguard your sensitive information and keep it out of the wrong hands.

Any replacement for a legacy DLP system must provide the enterprise with complete coverage for both cloud channels and traditional on-premises channels. Even most modern, cloud-delivered DLP solutions have been designed to cover only specific on-premises channels. They may address a network or a given set of endpoints or specific applications, but they don't address the full variety of modern use cases.

To provide complete enterprise coverage, a DLP solution must protect all transmissions of data to and from any location and device. That includes managed and unmanaged devices across every place where users are, both inside and outside the corporate network, as well as in SaaS applications, infrastructure as a service (IaaS), email, private apps, and endpoints. This requires a comprehensive and flexible DLP solution that can adapt to the constantly evolving needs of a highly distributed enterprise.

In the following sections, I take a look at key considerations in designing a DLP solution for the new, borderless enterprise.

Scalability and Future-Proofing

Not too long ago, the use of SaaS applications was fairly limited in the enterprise, but over time, there has been a significant increase in the number of SaaS applications being used by employees within enterprises. Now, it isn't uncommon for companies to use hundreds of approved SaaS applications, and (scary thought) employees may be using thousands of additional apps that the company isn't even aware of.

TIP

Scalability means not only meeting the present needs but also preparing for possible future changes and growth. A forward-looking approach is essential for creating flexible, agile solutions capable of dealing with an ever-increasing workload or ongoing expansion without compromising performance or functionality. Scalability helps you make sure your systems remain effective and efficient in the face of unpredictable change.

But scalability isn't just about addressing new environments and covering new places where data goes. Scalability is also about handling the increasing speed, variety, and volume of data. The amount of data being generated and collected today is unprecedented. With the rise of collaboration apps and online tools, data can now come in the form of conversations on apps like Slack, Teams, and Zoom, as well as cloud-delivered email apps like Gmail. It can also be in the form of images, such as photos and screenshots. People are as likely to take screenshots of important information as they are to paste it into a document. Scalability means protecting all different data formats and all these use cases, including use cases not yet developed.

REMEMBER

The "Modern DLP in Action" section, later in this chapter, gets into the details of how DLP systems work. For now, keep in mind that the basic functionality of a DLP system is to detect sensitive data and protect it.

How DLP evolved from hero to problem child

As organizations adopted cloud applications and expanded into new locations, the deployment of legacy DLP systems became increasingly unmanageable. These systems were designed to be installed and maintained on-premises, which meant that they had to be duplicated and installed in each new location and branch office. This added a significant amount of complexity and required a lot of resources, including hardware, maintenance, and personnel. Furthermore, the increasing trend of remote work added yet another layer of complexity as employees began to access sensitive data from a variety of different devices and locations. All of this made it difficult for organizations to effectively manage their DLP systems, leading to increased costs and potential security risks.

Have you ever shied away from updating your phone or laptop out of fear that it would disrupt a favorite app or cause some

annoying problem? Multiply that by thousands, and imagine having to upgrade legacy DLP software across numerous servers and branch offices, as well as thousands of employee devices. No wonder some customers hang on to old versions of their DLP software — it's a lot less effort than attempting an update.

WARNING

Not performing regular upgrades leaves data exposed and increases the risks of compliance violations and data breaches.

DLP needs to work smarter, not harder

Legacy DLP systems scan all data forms and identify sensitive information to protect. The idea is that only sensitive data needs to be protected, as protecting nonsensitive data can negatively impact productivity. For example, although it may be important to prevent certain sensitive data from being shared via email with third parties, protecting and possibly delaying any email communication with third parties isn't necessary because doing so can hinder communication and collaboration and produce too many alerts for the incident response team. Additionally, employees may be allowed to use company resources for non-work-related activities, such as uploading personal pictures on social media, as long as the content isn't sensitive and doesn't contain company secrets. Because legacy DLP systems are made of software and hardware components, having it scan all web traffic and all file repositories and having it look for all types of sensitive data, requires the deployment of additional detection servers, modules, and larger databases.

Due to their nature of being deployed on-premises, legacy DLP systems rely on hardware computing resources that are necessarily limited. For example endpoint DLP software installed on employees' computers are forcefully designed with limitations in their data detection capabilities, such as relying on basic detection engines that are less resource intensive. This means that although they can detect some sensitive data on endpoints, they're unable to use advanced detection methods, which may result in considerable quantities of sensitive data going undetected. For example, legacy DLP can't use advanced technologies that require substantial processing resources like machine learning (ML) and exact data matching (see the next section). Cloud-delivered DLP offloads resource-heavy activities to the cloud while still enforcing them on the endpoint. The scalability of this approach is a dramatic improvement, allowing DLP to fingerprint data such as specific names, Social Security numbers, and other sensitive details associated with individuals.

REMEMBER

The cloud can provide the effectively infinite scale that's needed to power these detection capabilities, allowing DLP systems to focus on the most important data and protect it from unauthorized access.

The Need for Precision

A prevailing myth that I discuss in Chapter 1 is that legacy DLP was inaccurate. But accuracy is not the real problem, or at least not the main one. The main problem is false positives (also discussed in Chapter 1) mainly due to lack of context. Sure, with data expanding like crazy across multiple devices and applications outside an organization's perimeter walls, and sensitive data becoming harder to detect as a result of the explosion of data types, legacy DLPs couldn't keep up and accuracy levels have dwindled. But the main issue is that legacy DLP solutions tended to be too restrictive — flagging beneficial actions as violations, and even blocking them, without understanding business context or risk level. In a world where collaboration is fundamental for the new way of conducting business, these false violations have become too many.

It's important that DLP does not cause friction for the business and disrupt the flow of data needed for beneficial business practices. For instance, if an employee wants to send a file to a trusted contractor who's engaged in a project, you don't want DLP to stop that transmission. Ideally, DLP should empower response teams to be more effective by amplifying legitimate incidents of potential data loss and filtering out the noise of false positives.

Accuracy and precision were not the main problems of legacy DLP systems, but they are for the less mature new cloud-delivered DLP solutions. There are two aspects:

- Inaccuracy of data detection may detect and unnecessarily protect too much data that is not sensitive — identify too much data as sensitive when it's not! — and possibly stop legitimate business communications.
- There may be a lack of detection methods to identify data that is actually sensitive — basically, missing sensitive data — such as missing certain file types like images or compressed formats, or missing passport numbers, health

information, international routing numbers, and country-specific national IDs because the system doesn't have the ability to identify those data formats and file types.

REMEMBER

To maintain trust and confidence, DLP systems must be precise and accurate, flagging and blocking only truly malicious data transfers and not generating too many false positives.

Key ingredient #1: Data identifiers

Data identifiers are used to find sensitive information like Social Security numbers or credit card numbers based on generically described content such as regular expressions (known as *regex*), a powerful tool that helps DLP automatically identify specific data types using natural, everyday terms, expressions, and patterns ("search for a nine-digit number"). One possible answer is that the number is a Social Security number, but how can you know for sure?

Data identifiers seek the answer by using special rules based on the number of numeric digits, text patterns, sequences, separations, and proximity keywords (like Social Security number [SSN], password [pwd], credit card number [CCN], and so on) to recognize these numbers and keep them safe. Here are some important points to keep in mind regarding data identifiers:

- Thousands of predefined data identifiers and the ability to customize them to suit your business are necessary to keep your information safe and to meet governance rules. Additionally, the ability to edit or create custom data identifiers is critical — every organization may have different sensitive information that needs to be protected.
- Data identifiers must support thousands of file types (Word, XLS, JPG, PNG, PDF, CSV, ZIP, RAR, and so on), formats, and categories (Image, Analytics, Archive and Compressed, Spreadsheet, Audio, Video, Database, and so on) (see Chapter 1).
- You must have support for a wide range of country-specific identification numbers (such as international banking information, addresses, postal codes, national IDs, passport numbers, and phone area codes) and regulatory and privacy compliance profiles to ensure that the DLP solution can keep up with the latest governance requirements.

TIP

For your DLP system to be effective, it needs thousands of data identifiers. This enables it to accurately identify and flag potentially sensitive information, across states, regions, and countries, no matter where it's located.

Key ingredient #2: Exact data matching (EDM)

EDM is a way to find specific structured information from sources like spreadsheets and databases. EDM allows a DLP solution to fingerprint and index confidential customer and employee records, which can be used to identify an individual using their full name, Social Security number, address, and other identification numbers. EDM can also be used to find financial records that identify an individual's assets, such as credit card numbers or bank account numbers. It can even be used for health-care information and product identification and pricing databases. With EDM, a DLP solution can index this information and then find it anywhere it's supposed to be. For EDM to be effective and accurate, it must match various pieces of indexed data and combine data fields from a particular record. It must also be able to index billions of records in order to support growing organizations, their expanding databases, and today's ever-increasing amount of information. Therefore, the scale of processing is important to EDM.

Key ingredient #3: Advanced data-detection capabilities

With more data types and ways of transferring them than ever before, organizations need their DLP system to be able to detect sensitive information with precision. *Advanced detection capabilities* is a bit of a catchall term that refers to things such as:

- **Optical character recognition (OCR) and artificial intelligence (AI)–based image recognition:** These features are becoming more and more important for data protection. Today, people take pictures of documents, forms, ID cards, whiteboards, and pictures of other pictures very easily. For example, people often take screenshots or snap photos to quickly capture information and share it with a colleague. Using OCR, a DLP solution can extract text from an image and then apply data classification based on the detection policies that are in place.

- **AI and ML:** AI and ML image classification, using sophisticated detection methods, can recognize common file and document types — like credit cards, tax forms, nondisclosure agreements (NDAs), mergers-and-acquisitions (M&A) forms, and patents — without necessarily extracting the content they contain. These methods can detect blurry, crumpled, and damaged pieces of content, even when the information is hard to read clearly. This is because the algorithms are trained to identify patterns and features specific to each type of document, such as the layout, fonts, and colors used. Additionally, they can also consider the context in which the document is being used. This allows the AI to accurately classify the document even in challenging conditions, such as low-quality images or damaged documents.
- **File and document fingerprinting:** This is an essential technique for organizations to ensure the security and confidentiality of their mission-critical documents and highly sensitive files. By indexing the entire document and detecting exact or partial copies of its content, organizations can prevent unauthorized exfiltration and duplication of their confidential information (such as M&A documents, prerelease information, engineering designs, or investor-related data). This technique is especially useful in detecting copies of sensitive files across risky environments and transmission channels, such as outbound emails and emails uploads to personal application instances.

Legacy DLP solutions actually provided some answers in the on-premises-only past, but they can't keep up anymore. They simply don't have enough computing power or the scalability.

Key ingredient #4: A lot of context and a zero trust data protection model

Just as the waves of the ocean are constantly changing and moving, so too are the people, networks, applications, data, and governance rules within a company. To stay ahead of potential risks, a DLP system and related strategy must be able to adapt and respond quickly and effectively to the constantly shifting data landscape, also known as *understanding context.* This agility allows the DLP system to effectively protect sensitive data, minimize data breach risks, and ensure compliance with relevant regulations without impacting user productivity and causing friction to the business continuity.

To achieve such nuance and flexibility, a cloud-delivered data protection platform should integrate with an organization's broader security and networking infrastructure. That DLP platform also should constantly gather information from various sources such as identity management, behavioral analytics, network logs, cloud security tools, threat analysis, network security, SaaS and cloud security postures, cloud access security broker (CASB)–native cloud confidence indexes, and endpoint security postures. This information can be used to accurately identify the specific circumstances of a user's access to sensitive data, the business context, and the potential risks involved in such an action and, therefore, determine the appropriate level of access and the right data protection response, all based on factors like a person's identity, location, and behavior; the safety of their device; the trustworthiness of the network; the reputation of the application being used; the final destination of a data transfer; and so on.


TIP

By being aware of risks and context, a data protection platform can continuously adapt and provide high efficacy and precise incident response.

Chapter 3 covers the concept of zero trust and its central role in effective DLP. For now, keep in mind that zero trust is an essential security strategy that assumes all users, devices, and networks within an organization's environment are potentially malicious and should be treated with suspicion at any time.

This means that all access to resources and systems is strictly controlled and verified, regardless of whether the user or device is inside or outside the network perimeter. Context is the engine that powers a zero trust strategy because it makes it possible for the DLP system to make informed choices about when to allow or disallow data-related activities to occur.


REMEMBER

Working with integrated security solutions and adjacent data protection technologies is what differentiates a data protection *tool* from a true data protection *platform.*

Modern DLP in Action

DLP sits at the heart of a company's information security framework and helps make other security tools more effective. It performs several critical functions, including the following:

- **DLP identifies sensitive data wherever it resides and moves, such as:**
  - *Data in motion,* which is data crossing the Internet, networks, applications, and devices (like uploads and downloads).
  - *Data at rest,* which is data being stored. This can be anything from storage in your private applications to a company-adopted SaaS application, such as when customer data is put into Salesforce, or internal-only documents stored and shared on Microsoft OneDrive or Microsoft SharePoint.
  - *Data in use,* which is data that's actively in use and collaborated on, such as transfer to USB, w activities, printing, or data that's faxed. (Do people still send faxes?!)
- **DLP monitors the data environment** to detect who's accessing data and what they're doing with that data. By monitoring actions, DLP can detect incidents, such as unauthorized sharing of confidential information, that may be in violation of corporate policy and take action to address them. This helps to ensure that sensitive data is not accessed or used without the right privileges (employees versus an outsider, or corporate versus personal device) or authorization or moderation (like suspicious bulk downloads of large amounts of files) and that any potential security breaches are quickly identified and addressed.
- **DLP automatically takes action to enforce policies** by, for example, stopping the data flow, encrypting the data, quarantining the confidential information, or unsharing the data on a SaaS application. For example, if an employee uses OneDrive to intentionally or accidentally share a file containing confidential information with external users, DLP can automatically unshare the file to prevent the unauthorized disclosure of the information.
- **DLP provides user coaching** by automatically notifying users of violations and the reasons behind them, while educating them on safe data-handling practices. Notification also helps to instantly educate users on security policies, reducing the need for incident response teams to manually triage issues. A good DLP should also be able to notify users instantly, without delay, and escalate notifications to managers, the response team, or HR as necessary.

Now Is the Time to Change Your DLP

Legacy DLP had been a dependable security solution for years, and it's no wonder that so many practitioners are still fans. After all, as I note earlier, those systems have undergone intense development over the past decade to protect on-premises networks from threats in the pre-cloud era.

Legacy DLP vendors have tried to bridge the gap between their systems and modern, cloud-first business requirements using technologies such as cloud-delivered secure web gateways (SWGs) and CASB solutions, using Internet Content Adaptation Protocol (ICAP) integration.

WARNING

Unfortunately, most legacy DLP systems are not designed to handle cloud and hybrid work use cases, which require integrations and capabilities with cloud services that legacy DLP systems don't readily support. This can result in compatibility issues and poor performance.

All these limitations and many more discussed in previous chapters have made legacy DLP unpopular, leading many organizations to simply turn off these tools altogether. As organizations increasingly move their data to the cloud, there is a growing need for cloud-delivered DLP systems that can recognize the changing contexts and risks associated with data management. These systems should be easy to deploy, expand, and scale, while covering both legacy and modern use cases. Because they're cloud-delivered, they're also always up to date, providing improved protection as business context and risks change.

IN THIS CHAPTER

- Learning how outdated data security can harm your business
- Discovering data context types and keeping business activities rolling
- Adapting to changing risk conditions to protect your data
- Ensuring modern business use cases can happen safely
- Assessing business context, risk, and user behavior to keep your data safe into the future

Chapter 3
The Role of Zero Trust in Modern DLP

A very important concept in security today — DLP and otherwise — is zero trust. A zero trust strategy assumes all users and devices, even those inside the organization's network, may be harmful and can't be trusted. That means access to sensitive data and systems isn't automatically granted based on personal identification and organizational affiliation. Access is given after careful authentication, check of security postures, and consideration of risk context, which is continuously reassessed. Zero Trust should not hinder productivity, it should instead enable a safe use of sensitive data and support modern business practices with security in mind, adapting automatically to changing risk conditions.

Zero trust continuously reevaluates the trustworthiness of each individual or device and operating environment before granting them access to sensitive data or to certain use of that sensitive data. Even if someone is an employee and has been granted

access before, they still need to be carefully assessed, such as by having their identity verified, their device and network connection checked, the risks of the applications that they're accessing gauged, and their behavior monitored to ensure they *remain* trusted. If they start behaving suspiciously or show signs of negligence, such as oversharing data, the system addresses their actions by, for example, reducing their privileges. This helps protect sensitive data from potential data loss risks and ensures that only trusted individuals can access and share it with other trusted individuals.

Zero trust aims to create a secure and controlled environment for data access and transfer, reducing the risk of data breaches and protecting sensitive data from unauthorized access. It does this by implementing strict access controls and continuously monitoring and verifying user actions, contextual risks, and behavior. In data loss prevention (DLP), a zero trust security model helps to minimize data breach risks, produce more accurate data protection results, and optimize incident response cycles by taking into account organizational context and risks. By allowing only safe access and use of sensitive data by authorized users and preventing any malicious, suspicious, negligent, or risky attempts to access or transfer that data, organizations can better protect their assets.

The Risks of Outdated Security

DLP systems were created to help prevent sensitive information from leaving a company. Legacy versions address a limited number of common data loss scenarios; their main purpose is to identify sensitive data and keep it within the organization, using a perimeter-based approach that is focused on controlling the flow of data in and out of the organization's network.

Using an approach called *implicit trust*, legacy DLP focuses on detecting and responding to predefined data violations. But this approach lacks context about users and their business reasons and the associated risks of a specific action.

For example, a legacy DLP system may search for Social Security numbers and block any attempt to send a Social Security number outside the enterprise's perimeter. In another case, it may stop sensitive data from being uploaded to a SaaS application

unequivocally, without any discern between a corporate instance of an approved SaaS app like Microsoft Team and a personal instance of that same app. This approach may seem secure, but it's actually quite rigid and lacks insight into users, devices, networks, applications and destinations that may reveal sanctioned activity. Implicit trust is a business inhibitor that prevents the effortless communication and flow of data necessary to grow a modern business.

WARNING

Because it doesn't continuously reconsider business context and risk, a legacy DLP system can't make informed decisions about data protection and can cause unnecessary disruptions to business operations.

With loose policies, implicit trust grants access to sensitive data without continuously verifying the identity and trustworthiness of the user or device. This is problematic because it leaves the organization vulnerable to potential mishandling of its sensitive data. After sensitive data leaves the perimeter, it's beyond the control of the organization's security.

This situation is a big problem in the cloud era. Sensitive data is used and shared outside the company's borders for even the most routine business functions. For example, common cloud applications and services such as Dropbox and Google Drive allow employees to access, share, and collaborate using sensitive data inside and outside the corporate environments. But legacy DLP systems that use implicit trust would either disrupt a legitimate collaboration or carelessly let data leak to the outside world, making it vulnerable to potential threats.

Zero trust data protection allows for the use and sharing of sensitive data as long as security conditions are continuously verified. It enables sensitive data to flow and be shared across users and devices and stored in different cloud services because it continuously verifies security conditions, such as user identity, device, network and application safety, and user behavior over time. Zero trust data protection applies specifically to sensitive data and ensures all security conditions are met at all times, enabling hybrid work, cloud, and modern business use cases.

REMEMBER

A modern, cloud-delivered DLP system that uses zero trust principles monitors and controls data anywhere corporate users want to connect and access data from, and anywhere data can be stored

and transferred across both cloud application repositories and on-premises environments.

Another issue with traditional security approaches based on multiple products and implicit trust is that they are very siloed, applying only one security control at a time without integrating all the security controls and without sharing risk intelligence. This means that different security controls are isolated and not integrated into a cohesive security platform, leaving gaps in your overall security strategy. To fully protect your data, you need multiple security controls working together and sharing intelligence.

The zero trust approach takes a more holistic and dynamic approach to data protection. It considers the context of the user, device, network, and other factors to make more informed decisions about data protection. This approach support the integration of DLP with other security controls and productivity tools and can continuously monitor and adapt to changing threats, risks and business conditions.

Overall, organizations that use DLP based on implicit trust must rely on the false assumption that users within an organization are trustworthy, are careful about security, and will never compromise sensitive data. In fact, because of their lack of security context, a restrictive enforcement of DLP policies would often cause disruption of legitimate business processes. In contrast, DLP based on zero trust closely monitors and controls how data is used at all times to adaptively prevent data policy violations.

A DLP system based on implicit trust would protect a credit card number by allowing authorized users access to the sensitive data while denying access to unauthorized users. This assumes that authorized users can be trusted to handle the data securely and not misuse it.

In contrast to legacy DLP systems based on implicit trust, a DLP system based on zero trust principles does not rely on the assumption of trust among users. Instead, it protects sensitive data, such as a credit card numbers, by requiring all users to undergo authentication process before accessing that data, regardless of their authorization level. This could include multifactor authentication, such as a password and a one-time code sent to a mobile device.

The system also continually assesses potential risks from devices, users, data, and apps. It verifies that the devices are trustworthy and secure, that the applications and their instances (i.e. corporate vs. personal) used are safe and compliant, that the network is safe and trusted, that data is shared with trustworthy destinations and recipients, and that the user's behavior is compliant. These conditions are continually verified, and the system adapts its protection response accordingly. Additionally, the system monitors and tracks user access to sensitive data, alerting administrators to any suspicious behavior or potential breaches and coaching the users on safe data use practices in case of violations of corporate policies. This approach reduces the risk of unauthorized access to sensitive data because the system verifies all users before granting them access and also minimizes risks to sensitive data over time be educating users in real time.

Context Lets Your DLP Say Yes to Important Business Activity

Zero trust helps data protection systems make informed decisions about allowing or restricting certain activities. It does this by considering multiple factors or contexts, such as the user's identity, the device used, the trustworthiness of the application, and the context of the data involved. (Zero trust gathers the context with the help of other solutions, which I discuss in the "DLP Shouldn't Stand Alone" section.) By taking all these contexts into account, applying zero trust principles can more accurately determine whether a particular activity is beneficial and necessary for the business and can say yes to it. This helps ensure that data is protected and the risk of security breaches or other threats are minimized while enabling business operations to continue running smoothly.

The following list defines the types of context used in zero trust:

- **User context:** Who's doing an action or who is the recipient of an action. This information helps determine if a user's behavior is good or if something is off. For example, suppose a user is suddenly moving a lot more data than usual, logging in from weird places, or acting strangely compared to before. That could be a sign of risky or malicious behavior.

The same applies if a user is accessing or using sensitive data and/or sending it to personal apps. Based on their identity and behavior, you may change a user's privileges to ensure sensitive data is protected and only authorized users can access that data share it with authorized recipients and transfer it to safe destinations.

- **Device context:** The device trying to access your data. You need to consider whether the device is personal or corporate, its security posture, and whether it's patched and up to date. You may also look at factors near the device, like the trustworthiness of the location it's connecting from. Considering all these things, you can determine the right level of privileges for the device based on how trustworthy and risky it is. Even if a user is usually reliable, their device may still be compromised or pose a security risk, so device context is critical in determining the privileges you should grant.
- **Application context:** The reputation and trustworthiness of the app used to access or handle data. This is important because if an app has a bad reputation or is untrustworthy, it could pose a risk to the security of the data being accessed or handled. Data protection systems may rely on other systems like a cloud access security broker (CASB) to gather information about the app's compliance-related and risk-related attributes. This can help the system determine if the app poses a risk, such as violating the General Data Protection Regulation (GDPR) by potentially overexposing sensitive data.

  A user may have access to multiple instances of a cloud app, which requires more granular control over sensitive data to prevent accidental sharing with personal accounts. Collaborative communication apps like Slack and Microsoft Teams may also pose a risk if channels within those apps have both corporate and external users, so the system must be able to differentiate between them to prevent data leaks. Keep all this in mind to ensure that the apps you're using are reputable and trustworthy and to protect your data from potential risks.
- **Data context:** How sensitive a specific piece of data is, its format, size and other factors. Where your data is being used and whether that use is legitimate. It helps to know what type of data is being accessed or moved and whether it belongs where it's being used. Sensitive data being accessed or transferred to an unauthorized location requires action to

prevent a data leak or breach. Data context is crucial for ensuring that data is appropriately handled and only accessed by authorized users in authorized locations based on its level of criticality. It helps determine whether an activity is necessary for the business and whether it's worth the risk.

WARNING

Most DLP solutions, not just legacy DLP, cause problems with how your business runs because typically they don't collect enough information about the business and the risks involved. Most DLP solutions force your organization to rely on incident response teams to make manual decisions about what to do. This is frustrating, inefficient, and expensive!

With zero trust, these problems are certainly minimized. A modern DLP system based on zero trust principles considers all the risks from things like users, devices, data, networks, and applications. This way, the system has a much better understanding of the risks involved and can automatically make the right decisions about protecting your data based on dynamic data protection policies adapted to your specific business needs. Zero trust helps you keep your data safe and your business running smoothly.

DLP Shouldn't Stand Alone

Data controls are used in legacy and new DLP systems. DLP is in fact designed to identify sensitive data and protect it. The problem with most of these data controls is that they lack context. DLP needs to be part of a larger platform, based on zero trust principles, that uses all the available context to make informed decisions. DLP needs help and intelligence from other solutions to gather all the necessary context, such as user context, device context, application context, and data context. That's why a zero trust–based system is integrated and focuses on contextual data controls instead of just blindly trusting everything. It's a way to adapt to changing risk conditions and automatically protect your data every time with the most appropriate response.

TIP

In zero trust data protection, look for consolidated controls where each control shares information and works together seamlessly to protect your data. For example, Netskope Intelligent Security Service Edge (SSE) directly enables zero trust and allows for sharing

and context between controls — including DLP at its core — making it super easy and efficient to protect your data.

Netskope Intelligent SSE supports its comprehensive DLP platform with several other security solutions. Some of the most important are:

- **Secure web gateway (SWG):** A SWG is a security solution that sits between users and the Internet, ensuring safe web connections and protecting against web-based threats. Netskope DLP via SWG ensures that sensitive data doesn't get leaked over untrusted and risky web traffic, including encrypted traffic. It detects, monitors and protects sensitive corporate data from being leaked and exposed over every web connection, including home offices, branches and public wifi locations.
- **CASB:** Netskope DLP through CASB discovers, monitors, and protects sensitive data across software as a service (SaaS) applications, infrastructure as a service (IaaS), corporate networks and branch offices, the mobile workforce, email services, and employees' endpoints. This centralized cloud-delivered service enforces unified data protection policies everywhere sensitive data is stored, used, or transferred and covers sensitive data in motion and at rest. It covers thousands of SaaS apps, and uniquely has awareness of data transmitted to personal app instances (i.e. corporate OneDrive to personal OneDrive) and risky apps. It scans thousands of different file types, as well as posts and asynchronous communications through collaboration apps and email services. Data protection, compliance, and data privacy policies are consistently enforced across public cloud services and automatically synchronized across the entire DLP platform.
- **SaaS security posture management (SSPM) and cloud security posture management (CSPM):** These technologies provide posture management for SaaS and public cloud environments to ensure security and compliance. They continuously monitor and assess security posture, identifying potential risks and misconfigurations and providing actionable insights and recommendations. Automated remediation capabilities address identified issues in real time.
- **Endpoint protection software:** Netskope Endpoint DLP is a solution that detects, monitors, and protects sensitive data on employee endpoints. Because the solution is integrated

into the single Netskope client, there is no need for deploying a separate agent. Netskope Endpoint DLP minimizes resource utilization while featuring a full suite of capabilities, including ML-based classifiers, optical character recognition (OCR), file fingerprinting, exact data matching (EDM), and more. Leveraging the cloud-delivered DLP service and intelligence sourced across the entire DLP platform helps avoid duplicate scanning of data originated in the cloud, resulting in a frictionless user experience and stronger protection outcomes.

- **User and entity behavior analytics (UEBA):** This security control continually assesses user behavior to identify any unusual or potentially risky activity. In the past, UEBA was often a siloed security control, but it needs to be integrated with DLP to be effective. By ingesting DLP violation logs and flagging risky behavior for further evaluation, UEBA can inform subsequent changes to policy enforcement and help keep your data safe.
- **Identity and access management (IAM):** IAM is the practice of managing and controlling access to resources based on user identity. It includes technologies such as multifactor authentication, single sign-on, and access control lists. Netskope integrates with many IAM vendors to ensure that only authorized users can access specific resources and protect against unauthorized access. IAM is an essential part of any organization's zero trust security strategy, helping to protect resources and ensure compliance with security policies and regulations.
- **Email protection.** Netskope provides a very extensive DLP solution for email like Microsoft 365 and Gmail, and for both data in-motion and at-rest. The solution protects outbound sensitive emails in real-time through SMTP proxy and webmail, and can discern sensitive data leaving via a personal email account from data sent via a corporate email account or via private email services.
- **Zero Trust Network Access (ZTNA):** Netskope DLP delivered via Netskope Private Access (NPA), remote access solution, prevents data loss and exfiltration across private resources in the data center and in public cloud environments ensuring data protection for browser-based access to private applications anywhere the users are connecting from.

By combining these core components into a single, integrated platform, Netskope's SSE platform provides a comprehensive security solution that can protect your organization from a wide range of threats.

Putting Zero Trust Principles to Work with DLP

REMEMBER

The purpose of zero trust data protection isn't to stop sensitive data from leaving the enterprise. It's also to allow modern business use cases to happen while always keeping safety and risk in mind.

This means supporting users in different locations and encouraging collaboration, all while keeping your data secure. Zero trust data protection is about being able to work from anywhere and still have access to all the resources you need and being able to collaborate with team members and external partners without worrying about data leaks. Using a unified solution like the Netskope SSE, you can protect your data and take advantage of all the benefits of modern business data workflows. Here are a couple examples of how this works in practice:

- Imagine you're working on your laptop, logged in to your company's network using the Netskope SSE. You access some important sales documents and start working on them. But then you accidentally try to save a copy of the documents to your personal cloud storage account instead of the corporate instance of that same cloud storage application.

  With DLP based on zero trust principles, the system recognizes that you're trying to send sensitive company data to a personal app instance and prevents the data from being saved. Instead, the system displays a user coaching notification, a pop-up that immediately informs you of the violation and reminds you of the correct location to save the documents. This way, you can work from anywhere and still have access to all the resources you need without worrying about accidentally sending sensitive data somewhere it doesn't belong. Coaching notifications educate users about safe practices and company policies, minimizing the risk of data

loss over time and reducing the need for time-consuming training throughout the year.

- Let's say you're collaborating with external partners on a project and you want to share some documents with them. With DLP based on zero trust principles, the system will check the reputation and trustworthiness of the app you use to share documents, your identity and behavior, the device used, and the transmission destination.

  If you're using a personal cloud storage app that has a different level of security than your company's corporate app, the system may prevent you from sharing the data through that app. Instead, it may suggest you use a different app or send the documents through a secure channel. The DLP will also check the destination of the transmission, such as if the recipient is an external user or an employee and whether the destination is safe. The DLP may send a notification to you that asks if you're sure about sharing sensitive data with the external recipient and may even ask you to justify your action. This way, you can collaborate confidently, knowing your data is protected and only authorized users can access it.

Adaptive Zero Trust

Adaptive zero trust is all about recognizing that things change over time. This means that zero trust data protection needs to continually assess business context, risk, and user behavior to keep your data safe.

To visualize this point, imagine a bouncer at a nightclub. One night, they're standing at the door when a group of people approaches. The bouncer checks their IDs, and everything looks good, so they let the group in. But as the night goes on, the bouncer starts to notice some strange behavior from one of the people in the group. Maybe they're acting aggressively or trying to access areas of the club they're not supposed to. With adaptive zero trust, our bouncer would recognize this change in behavior and take action to protect other people and the club. They may keep a closer eye on them to ensure they don't cause any problems or even ask the person to leave. This way, you can keep other people and your club safe and secure, even if someone's behavior changes.

Consider these common scenarios your business is likely to face:

- **Someone's behavior changes.** You have a trusted employee who has always had access to certain sensitive company data. One day, maybe after a performance review, they start behaving differently. They begin accessing and downloading more sensitive data than usual or logging in from unusual locations. With adaptive zero trust, the system will recognize this behavior change and adjust the employee's privileges accordingly. For example, the system may restrict their access to specific data or notify the security team for further evaluation. This way, you can protect data even if a trusted employee's behavior changes.
- **Applications' reputation and trustworthiness change.** Applications change over time; not only their functionality, but also their reputation, their security postures and trustworthiness can change. For example, a cloud storage app that was once considered secure may have a new vulnerability exposure or misconfiguration that affects its trustworthiness. With adaptive zero trust, the solution will continually assess the app's risk level and adjust privileges as needed. This way, you can protect your data even if an app's trustworthiness changes.
- **Devices become compromised.** Devices can become more vulnerable or even compromised without the user even realizing it. For example, a laptop that was once considered safe may get infected with malware or have its security settings changed without the user's knowledge. With adaptive zero trust, the system will continually assess the device's security posture and adjust privileges as needed. This way, you can protect your data even if a device becomes compromised.
- **Data flow changes.** Data flow can change because of changes to compliance rules at different levels. For example, a data flow may be considered acceptable, but if the destination becomes noncompliant or unsafe, regulations may still require that the organization protect the flow of data. This is the case with GDPR, which says that certain private data may not be allowed to exit the EU unless adequacy or a valid transfer agreement is in place. With adaptive zero trust, the system will continually assess the risks and adjust privileges as needed. This way, you can protect your data even if the rules change.

- **A user's role or status changes.** Users who give their two-week notice may still have access to sensitive data during that time. With adaptive zero trust, the system will continually assess the risks involved and adjust privileges as needed. For example, the system may restrict the user's access to specific data or notify the security team of an action that needs further evaluation.

TIP

Adaptive zero trust evaluates data usage from as many perspectives as possible in order to adjust privileges to protect the sensitive data, and the company reputation, and support business activity.

Adaptive zero trust unlocks increased protection while making data and people more productive. It brings dynamic, adaptive data protection policy to life by continually assessing risks and adjusting privileges as needed. This is a significant shift from the typical approach of DLP systems, both legacy and new, which rely on a one-time approach based on implicit trust, leading to many false positives and incident triage fatigue. With such a cumbersome approach, the incident response team is forced to manually evaluate each incident to determine if it was an actual violation and then contact the user responsible (often after they had forgotten their previous action). The team then has to decipher the entire data flow — a long, resource-intensive process. Adaptive zero trust provides a model for continuous protection, making it much easier to keep your data safe and your business running smoothly.

Netskope Adaptive Zero Trust Data Protection

Netskope's implementation of adaptive zero trust data protection is all about context. By monitoring traffic among users, devices, applications, networks, and destinations, Netskope builds a deep understanding of what's happening in your organization. This enables the system to exert granular control over data access, allowing you to protect your sensitive data without hindering business operations.

For example, imagine a user trying to access sensitive company data from a personal device. With Netskope, the process begins

with the accurate detection of sensitive data. Additionally, by considering various contextual factors, the incident response becomes more precise and effective, reducing the need for manual triage and minimizing the burden on security teams. The system would evaluate the security posture of the device, the user's identity, and the user's behavior to determine whether access should be granted.

Other factors considered include the network connection and location, potential vulnerabilities, available threat intelligence, and more. The application's associated risks and reputation will be considered by the Netskope Cloud Confidence Index (CCI), a database of nearly 60,000 cloud apps (and growing!) that Netskope has evaluated based on about 50 risk-based criteria. These criteria measure an app's enterprise-readiness, taking into consideration an app's security, auditability, and business continuity.

If the device is deemed risky or the user's behavior is deemed unusual, access may be restricted, or the security team may be notified for further evaluation. If the device is secure and the user's behavior is normal, access may be granted.


TIP

The foundation of Netskope data protection is its SSE, part of the broader Netskope Secure Access Service Edge (SASE) platform. This converged, cloud-native security solution consolidates the vital security technologies I define earlier into a single, integrated platform. By combining these technologies into a single platform, Netskope makes it easy to manage your security from one central location. Netskope SSE is cloud-native, which means it can scale quickly and efficiently to meet your organization's needs. It's also designed to be highly flexible, so you can customize it to meet your specific security needs.

Netskope SSE was designed with the understanding that security is about more than policy enforcement. It's also important to coach employees and encourage safe data handling behavior. That's why the solution preserves the user's ability to make business decisions while keeping your data safe. For example, when a violation occurs, Netskope SSE may direct employees to a training on how to handle sensitive data, ask questions to evaluate context further, or provide guidance about tips and best practices for working securely from home. By taking a holistic approach to data protection, Netskope helps you create a culture of security in your organization.

IN THIS CHAPTER

- Comparing modern and legacy DLP solutions
- Staying safe anywhere you access data
- Using unified policies and access controls
- Sizing up Netskope DLP's benefits and differentiators

Chapter 4
Why Netskope for Modern DLP

Chief information security officers (CISOs) and information security teams often face a difficult decision: Should you stick with mature, but complex and costly, legacy data loss prevention (DLP) solutions, or go with easy-to-deploy cloud options that likely lack the depth and breadth you fully need? You'll be prepared to answer that question after reading this chapter and learning about the main benefits of all cloud-based DLP solutions:

- **They can provide comprehensive coverage.** No matter where your data is stored, where it's transferred, or how it's accessed, a cloud-delivered DLP can protect it.
- **They can provide coverage for cloud environments.** SaaS applications, IaaS public cloud services and web access no matter where your users are connecting from across the modern hybrid-work enabled enterprise.
- **They eliminate the need to set up additional infrastructure because they can be deployed quickly and easily as cloud services.**
- **They protect your sensitive data without putting a strain on your network and endpoint resources.** A cloud-delivered

DLP system can handle all the data scanning and detection algorithms you need at maximum capacity.

- **They're easier to integrate with a wide range of other security tools.**
- **They deliver increased visibility of data that is transferred and stored outside of your corporate premises.**
- **They're easier to maintain and update in real time, and offer the ability to scale faster and more easily than with older models deployed on premises.**

After reading this chapter, you'll have a good understanding of how these benefits may apply to your organization and be well equipped to make an informed decision about which cloud-delivered DLP is the right fit for your company. Along the way, we give you specific information regarding the differentiators of the Netskope platform.

Differentiating Between Cloud-Delivered DLP Choices

Modern DLP needs to be cloud-delivered. Two types are available. Cloud-native DLP is typically embedded in infrastructure as a service (IaaS) platforms and software as a service (SaaS) apps from cloud service providers. Integrated cloud-delivered DLP solutions are usually part of a security service or product such as a secure web gateway (SWG), next-generation firewall (NGFW), or cloud-access security broker (CASB).

Type 1: Netskope DLP versus cloud-native point solutions

Netskope DLP offers a number of advantages over more limited cloud-native point solutions. One key advantage is its broader coverage using a single enterprise-grade DLP policy engine, which ensures that data sensitive data is protected across a wider range of formats, communication channels and environments, including SaaS applications, IaaS services, private applications, email services, file sharing, and web transactions anywhere your users are. Netskope DLP also includes endpoint DLP protection,

which is important because it helps ensure all of your sensitive data is protected, even across endpoints in remote locations that may or may not be connected to the cloud through any specific network. The single DLP policy engine also significantly reduces complexity compared to having to manage different DLP policy rules for different channels and different cloud services.

Another advantage of Netskope DLP is its superior detection accuracy. By scanning the entire spectrum of file types and data formats, using a broad range of data detection algorithms and ML to understand a wide variety of information and documents, and its specific context, it's able to accurately identify and classify sensitive data, even if that data is stored and transferred in different structures, formats, languages, or embedded in images. This is important because it helps to ensure that any type sensitive data is not accidentally leaked or exposed, which could have serious consequences for an organization, and that the system produces true data security events rather than false positives.

Finally, Netskope DLP has zero trust context baked in, which means it's designed to work within a rich zero trust security framework. This is important because it helps to ensure that all access to sensitive data is carefully controlled and monitored, in the right risk context, reducing the risk of unauthorized access, overexposure or data leaks.

Today, many cloud service providers (CSPs) and SaaS vendors offer native DLP capabilities within their platforms. These readily available cloud-focused solutions are often chosen by organizations that are pursuing a cloud-first strategy or those that are just starting their data protection journey. Although these solutions may address the specific cloud data protection use cases for which they were designed, they may lack broad coverage and may not be as comprehensive as legacy DLP solutions.

WARNING

Some enterprises start with these cloud-native DLP solutions because they can be quick and easy to implement. However, it's important to approach these solutions with your eyes wide open, understanding that they may not be sufficient to meet all your data protection needs. In some cases, organizations may find themselves forced to adopt multiple, disconnected, siloed DLP options for subsequent use cases, leading to a fragmented and potentially less effective data protection strategy.

Type 2: Not all integrated cloud-delivered DLP solutions are created equal

When it comes to choosing a cloud-delivered DLP solution, keep in mind that many newer solutions on the market have significant shortcomings:

- They may offer breadth of coverage but lack the technology depth and features necessary to protect your organization's sensitive data effectively and accurately across all modern use cases.
- They may offer some of the latest methodologies and features for a few specific use cases and data formats, but lack the breadth of coverage necessary to protect your organization's sensitive data comprehensively.

WARNING

Some newer cloud-delivered DLP solutions may be well-marketed, but are far from being as sophisticated and mature as the legacy DLP solutions they're meant to replace.

It's important to thoroughly research and compare DLP solutions to ensure that you choose one that will effectively meet your organization's needs. Look at factors such as the maturity and the sophistication of its data detection capabilities (for example, how many types of files it can scan and how many data identifiers it uses, including localized data types specific to different countries), the variety of channels it covers, its ability to adapt to changing risks and environments, and the level of integration and customization it offers.

If you're considering using a cloud-delivered DLP solution, you may be wondering which type is best for you. Let's look in more detail at what to consider:

- **Breadth of coverage:** Integrated DLP solutions are typically included as part of a SWG, CASB, or NGFW, and are often part of a zero trust network access (ZTNA) service. These solutions are delivered from the cloud and integrated typically within a network security service. They are limited in scope, lacking, for example, data protection for outbound emails, endpoints, a larger spectrum of SaaS applications and their specific instances (i.e. corporate vs. personal accounts).

- **Limitations of the solutions:** Be aware that these solutions may not cover all modern and traditional use cases like cloud collaboration with external users, data transfers via personal email or email drafts, USB file transfers, screenshots and pictures of sensitive documents, new compliance templates, data in foreign languages and formats etc. Most important, they may have weaker detection capabilities. Additionally, their ML and AI capabilities may be underwhelming.
- **Accuracy of sensitive data detection:** Many newer cloud-delivered DLP solutions fall short in their ability to accurately and granularly detect sensitive data. They often scan only a limited number of file types and lack the breadth of data identifiers that more mature solutions possess. These solutions may make a big splash by focusing on one or two flashy features but ultimately fall short in their ability to provide comprehensive data protection.

  A mature solution will offer thousands of predefined data identifiers, including a wide spectrum of personally identifiable information (PII), passports, bank accounts, International banking information, national IDs, financial data, medical data, biodata, and industry-specific information, as well as localized languages and customizable identifiers. It would also provide a wide range of predefined policy profiles to support use cases and compliance requirements such as the General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA), the Payment Card Industry Data Security Standard (PCI-DSS), the Health Insurance Portability and Accountability Act (HIPAA), and the Gramm–Leach–Bliley Act (GLBA) just to mention a few.
- **Integration into a platform:** A cloud-delivered DLP must tightly integrate with a broader security platform to effectively protect sensitive data within the entire risk context available of users, devices, networks, applications, behaviors and destinations. A well-integrated DLP solution will use intelligence from other control points, such as user behavior analytics, next-generation security web gateways, CASB, ZTNA, and security posture management, to comprehensively understand an organization's security posture and the risks associated with every single interaction with sensitive data. This includes being aware of the specific instances of

SaaS applications and devices in use, distinguishing between personal and corporate email accounts user identities, the recipients of a data sharing and much more. This level of integration allows for a more accurate and granular approach to detecting and protecting sensitive data.

TIP

Not all data protection solutions are created equal, and many lack the maturity and sophistication needed to effectively replace legacy solutions. Some vendors may offer DLP as an add-on to their core products, but without the necessary breadth and depth, these solutions may not provide the level of protection that organizations require. Any solution you consider should be tested to ensure it supports all the data types and volumes needed today, and covers any data egress point both on-premises and in the cloud without compromises.

Carefully evaluate the capabilities of different DLP solutions, and choose one that will meet your organization's needs, both now and in the future. Mature feature sets and a dedicated vendor are essential for success. Relying on just the basics can lead to inaccuracies, partial detection, and tons of false positives.

With a decade's worth of continuous innovation and full dedication to data protection, Netskope has been recognized as an industry standard bearer compared to other SASE and security service edge (SSE) vendors. In the following sections, we delve into the features and capabilities that set Netskope DLP apart.

How Netskope DLP Keeps You Safe

Netskope DLP is a comprehensive, cloud-delivered integrated solution that helps protect your data across all fundamental channels, including clouds, networks, emails, endpoints, and users from any location. It's designed to be risk-aware and context-aware, so you can trust that your data will always be safe wherever it moves.

Netskope DLP is *fully integrated* into the comprehensive Netskope SSE described in Chapter 3 and delivered as part of a full SASE platform. This means you get a converged, cloud-native security platform that helps eliminate blind spots, provides consistency, enhances performance, and reduces costs and complexity.

Netskope DLP covers all channels and data transfer, as shown in Figure 4-1, so you can be sure that your sensitive information is always protected. It covers

- Nearly 60,000 SaaS applications, with new apps dynamically classified, and every instance of these applications
- Every major IaaS provider, including Amazon Web Services (AWS), Google Cloud, and Microsoft Azure
- Private applications in the data center or hosted in the public cloud
- Your corporate networks and branch offices
- Your mobile workforce
- Every email service, on-premises and in the cloud, including webmail
- All your employee endpoints, on- and off-premises

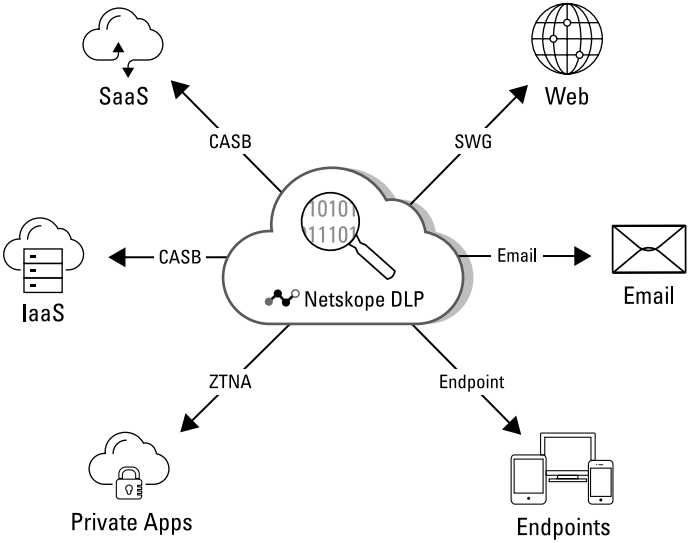

FIGURE 4-1: Netskope DLP has your data covered, no matter where that data is.

Key Differentiators

A myth of legacy DLP solutions is that they're inaccurate; in reality, false positives are a bigger problem, which requires greater precision to solve. We explain this in Chapter 2, where we also introduce and explain key ingredients that can help DLP systems achieve precision. Here, we explain how Netskope has turned these key ingredients into key differentiators and provided a modern DLP solution that can be customized and automated to meet your business needs.

Full coverage of all critical channels with unified policies

Sensitive data moving everywhere outside the traditional corporate premises becomes harder to track and protect, and is more prone to both intentional and unintentional exposure. Netskope cloud DLP comprehensively discovers, monitors, and protects sensitive data in-motion, at-rest and in-use across the entire enterprise ecosystem including SaaS applications, IaaS public clouds, corporate networks and branch offices, the mobile workforce, email services and through employees' endpoints.

It provides unified data protection policies for every location where data is stored, used or transferred and delivered from a centralized cloud service.

Single console, with role based access control, ensures that policy configurations, monitoring, reporting and incident response for all channels are all managed through a single pane of glass by the practitioners.

Superior detection and protection of sensitive data

Data identifiers are fundamental to help a DLP solution identify sensitive data based on certain characteristics like descriptive keywords, regular expressions, number of digits, special characters, patterns, proximity analysis etc. When shopping for a DLP solution, ensure it has the identification capabilities to cover all

your current and future use cases. A good DLP solution should be able to provide several thousands of predefined identifiers to accurately search for and identify the broadest variety and the slight variation of sensitive data. This is particularly important for global enterprises that need to have identifiers for multiple countries. Netskope provides all those features with ML and the ability to granularly customize identifiers and policy templates, and makes sure that it covers all your data protection needs.


TIP

Don't focus only on the data identifiers you need now. You need a future-ready solution that can address types of data, applications, and regulations yet to be invented. Look for a solution with thousands of predefined data identifiers and policy templates for compliance regulations like GDPR and CCPA. And don't forget the ability to create and edit custom data identifiers to fit your specific needs.

There are thousands of file types that may have sensitive info such as compressed files (ZIP, RAR, ISO, and so on), presentations, emails, images (BMP, JPG, PNG, and so on), spreadsheets, computer-aided design (CAD) files, social media posts, online forms, chat messages, various attachments, and graphics. That's a lot of different types of data to keep track of, so you want a DLP solution that can handle all of them.

Exact data matching (EDM) scale is a fundamental aspect to consider, especially if you have a large enterprise — or plan to one day. The DLP solution should be able to process millions or even billions of records with ease, modern cloud-based DLP solutions like Netskope can leverage cloud computing to perform high-scale data fingerprinting analysis even on endpoints, without slowing down other essential processes. This way, the entire collection of personal data of employees, customers and partners, and much more, will be fully protected.

To ensure your sensitive data is further protected, look for advanced detection capabilities — optical character recognition (OCR), AI, ML, file fingerprinting, and zero trust strategies — in your DLP solution, all of which are included in Netskope DLP (and which we cover further in Chapter 2).

TIP

HONEY, LET'S SHRINK THE ATTACK SURFACE

Organizations that want to protect their sensitive data from cyber threats must eliminate any gaps in protection. The *attack surface* is the total amount of potential vulnerabilities or entry points that attackers could leverage and that insiders could user intentionally or maliciously. Limiting the attack surface can make it harder to find and exploit weaknesses; closing any existing gaps in protection can also significantly reduce an organization's risk of a successful attack and accidental exposure. Ensuring that all devices, applications, and networks are properly secured is essential for eliminating any loopholes that may facilitate risky exposure.

Netskope DLP can accurately identify sensitive data, even if it's stored in modern unstructured formats like images (screenshots and pictures), or different languages. Thanks to its sophisticated ML-classifiers the solution is able to discern sensitive images such as driver licenses, credit cards, IDs, contracts, patents, M&A documents and checks even if such images are not clear, blurry, distorted and damaged. It actively protects sensitive information, so you can trust it to keep your data safe in the constantly evolving world of the cloud. This also reduces the workload on your security teams by automatically identifying and protecting sensitive data.

Netskope DLP has a variety of advanced ML-based classification tools, including thousands of data identifiers. It scans more than 1,600 different file types with contextual detection policies, highly scalable exact data matching, structured and unstructured document fingerprinting, precise ML-based image classification, advanced OCR, and AI/ML data classifiers for data discovery and identification.

Context and risk-aware data protection

Effective data protection is all about context. By monitoring traffic among users and apps, you can exert granular control and allow or prevent risky use of sensitive data based on many factors, like who the user is, what they're trying to do, and why they're doing it. This data-centric approach is the best way to manage risk in modern, hybrid enterprises.

Incident response fatigue and business disruption are problems of the past with Netskope DLP. In fact the the Netskope DLP solution goes beyond the static approach of discovering sensitive information and responding to a predefined violation policies, and factors in organizational context and security risks in order to dynamically enable the proper protection based on changing conditions.

Netskope DLP is natively integrated to the comprehensive Netskope Security Service Edge (SSE) solution, a fully converged cloud-native security platform that consolidates security technologies, like SWG, CASB and UEBA, onto a unified, integrated cloud-native platform. This approach eliminates security blind spots, provides policy consistency, and dramatically reduces costs and complexity. The platform is continually aware of users behavior, geolocation, security postures, device risks, application risks and reputations, personal application instances etc. and allows DLP to tailor incident response to true data security incidents, minimizing false positives, incident triage and business disruption.

You can increase visibility and risk mitigation across all key vectors with a single, converged SASE data protection solution based on zero trust principles and advanced data protection controls. Plus, you can simplify data classification, policy definition, and incident management with a converged platform that uses ML, rich reporting, and advanced analytics. And with flexible, context-driven policies and a lightweight agent, you can improve end-user agility and reduce friction.

REMEMBER

To ensure your data protection program is a success, you must train your employees and encourage safe data handling practices. Netskope DLP offers real-time user coaching and awareness programs to do just that. It also integrates with leading learning management systems and has a customizable end-user portal for self-service data protection education.

Work Smarter with DLP

Netskope DLP is delivered from the cloud, so it doesn't rely on on-premises components. It also offers always-on, up-to-date protection, eliminating the need for manual software updates like legacy DLP solutions.

With unified data protection policies and single-console and role-based access control (RBAC), managing policy configurations, monitoring, reporting, and incident response is a breeze.

In the past, companies had to build separate policies for separate channels (for example, web, email, and each individual app), which was resource-intensive and time-consuming. Netskope DLP is a unified, centralized cloud service where you can define a single policy for your company and have it automatically synchronized across all channels. This way, you can build your policy once, and you don't have to refine it constantly to replicate it and refine it constantly in different places.


TIP

Legacy DLP solutions needed a lot of system administrators to build and manage policies. Today's talent shortage makes it important to choose a solution that's easier to manage.

A centralized user interface (UI) and a unified management console are also crucial for effective and efficient incident response. You may have had separate consoles for on-premises and cloud-delivered tools, which can be confusing and time-consuming to manage. Even today, some newer DLP vendors still use a multiple-console approach, which can further complicate things. With Netskope DLP, you receive all violations in one place, sensitive data detection and incident response is delivered consistently, and in real time, so you can respond quickly and effectively to potential threats.


TIP

A centralized UI and unified management console make it easier to keep track of everything and streamline the incident response process.

Chapter 5
Ten Keys to a Successful Transition to Modern, Cloud-Delivered DLP

Replacing long-established, legacy security implementations like data loss prevention (DLP) can seem intimidating. Your current iteration is the result of years of complicated, interlocking processes. Like a house of cards, each element touches the other, and removing one threatens to bring the whole structure tumbling down.

Don't be intimidated! Innovative digital transformation is worth aiming for. And change doesn't have to happen overnight. Take small steps, use your current investments wisely, and you'll be on your way to a comprehensive data protection solution that protects sensitive information across all platforms — whether on-premises or in the cloud.

- **Assess your data protection needs**. Take the time to assess your organization's current technology environment thoroughly. Identify and understand what data must be protected, which services and repositories are being used to store and process sensitive information, and how these

services are being used by departments and individuals. Have your security team specifically identify and assess all corporate applications, email services, collaboration tools, network locations, users' hybrid work practices, connecting devices, and business processes to map data flows and determine how data is shared among employees or with external parties.


TIP

Don't limit yourself to the security team. Your company's chief data officer, legal staff, and HR personnel are among other stakeholders that can provide insight into how your company uses data.

Examine all categories of stored data and any transactions involving data moving across networks. Find out how much priority needs to be given to protecting various types of data within your organization. This stage can represent a quick win for organizations that need regulatory compliance support or require new DLP deployments due to ineffective legacy systems.

- **Identify and mitigate your highest risks.** When looking to transition to a cloud-delivered data protection solution, determine which areas of your current technology environment pose the highest risks. Think about unintentional data sharing, malicious exfiltration, and other cloud-based cyber threats that are associated with corporate software as a service (SaaS) applications, cloud email, and infrastructure as a service (IaaS). Netskope's market-leading cloud access security broker (CASB) solution embeds DLP as its core component to protect data security for both corporate-sanctioned cloud applications and (you're fooling yourself if you think you don't have them) unsanctioned apps.
- **Choose your data protection vendor wisely.** Make sure you choose a vendor that meets your company's needs in every environment today and into the foreseeable future. Netskope DLP is the only vendor that provides comprehensive coverage for all cloud needs and beyond. This includes protection of data at-rest, in-transit and in-use across clouds and on-premises locations, endpoint DLP, email DLP, network DLP for web and for email, DLP for SaaS and IaaS, and DLP for private apps. This comprehensive coverage across all modern data movements ensures corporations have maximum visibility across their whole system and untrusted locations as well. Carefully evaluate the depth of

capabilities of each solution, for example how many and which file types the solution is able to scan, the ability to understand image formats, and the coverage of the widest variety of sensitive data including international and country specific identifiers. Consider the ability of the system to leverage as more risk and business context as possible, and therefore to make automated and informed incident response decisions with every use of sensitive data adaptively. Basically make sure you don't adopt a superficial approach to data protection that will create more problems than providing solutions.

- **Protect your email services and your collaboration apps.** Discover the power of cloud-based email and SaaS protection with Netskope DLP. This comprehensive DLP solution is designed to secure all your company's sensitive information including outbound sensitive emails and asynchronous communications via SaaS-based collaboration apps like Slack and Teams. With application programming interfaces (APIs), real-time protection inline, protection for external collaborations, and even instance awareness, like personal email and SaaS instances vs. corporate instances of the same services, you can be sure that your corporate data is secure no matter what. With the help of Netskope, you'll have peace of mind when it comes to collaboration and communications.
- **Protect your cloud-based email.** Discover the power of cloudbased email protection with Netskope DLP. This comprehensive DLP solution is designed to secure all your company's sensitive information by protecting against malicious attacks and unintentional data sharing. With application programming interfaces (APIs), real-time protection inline, and even data protection through personal email instances, you can be sure that your corporate data is secure no matter what. With the help of Netskope, you'll have peace of mind when it comes to migrating your email service to the cloud.
- **Secure data in motion.** Data that is transferred across different locations, connection, services and devices, such as home networks, corporate offices, branch offices, corporate devices and personal devices, can be difficult to manage and secure. Traditional proxy-connected DLP solutions are not always enough protection when it comes to data in motion. Netskope's provides a unified DLP service that is delivered

through the entire Netskope intelligent security service edge (SSE) platform, and is designed to secure sensitive data from anywhere people work. This way, you'll have the ultimate security for your data transactions, including all the benefits of zero trust principles and all the risk context available and none of the hassles of obscure hardware configurations. With Netskope's innovative DLP solution, you can ensure that your data is safe at all times in all places.

- **Protect data on employees' endpoint devices.** Even though more and more data is stored in the cloud, it's still important to be sure that sensitive files don't get lost or stolen on endpoints that may or may not be connected to a corporate network, or may not be connected at all. Whether sensitive data is created on the endpoint or downloaded from the cloud, Netskope DLP can help with this. This lightweight endpoint solution offers all the advanced DLP capabilities — such as machine learning (ML)–based classifiers, optical character recognition (OCR), file fingerprinting, exact data match (EDM), and more — with minimal resource utilization because it leverages the cloud. It enables a variety of use cases including detection of data transferred via USB and provides USB device protection and other device control policies to make sure your sensitive data stays safe no matter where your people are connected from.
- **Stick with what works as you plan for the future.** If you've recently invested in DLP capabilities from a cloud service provider or SaaS vendor, it may make sense to stick with them for the short term. For example, if a SaaS vendor is already doing a good job of protecting your office suite applications, you don't have to change immediately. But keep a watchful eye for the point where you're managing too many discrete, disconnected policies. If you're looking to expand data protection across multiple clouds and SaaS apps, you could end up dealing with too many consoles and different policies. Netskope DLP offers a simpler solution: one console with consistent policies that can protect your data no matter where it's stored or accessed.
- **Unlock comprehensive data protection.** Netskope DLP offers a modern approach to data protection that's more efficient and effective than ever before. Advanced detection technologies like ML, data fingerprinting, and image recognition are use at full potential and unprecedented

scale, even on the endpoints, as the computing capacity is delivered from the cloud. The single console with unified policies makes it simple to manage the entire organization's data protection needs. Gathering and analyzing risk intelligence and contextual information about users, devices, data, networks, clouds, and behaviors uniquely enables Netskope DLP to evaluate every interaction with sensitive data and dynamically adapt response to each specific policy violation. This new approach supports safe collaboration and modern data sharing practices and doesn't hinder productivity, minimizes false positives and produces more accurate data protection outcomes. Netskope DLP is natively integrated into the overall Netskope SSE platform, and therefore always aware of business risks, behaviors, and security vulnerabilities. Netskope DLP is fully integrated with Netskope SSE, so organizations are always aware of business risks, behaviors, and security vulnerabilities.

- **Preserve institutional knowledge.** Transitioning to a new cloud-based DLP can seem overwhelming, but it doesn't have to be. Leverage the experience and knowledge of the people who've maintained your legacy DLP system, including your policy administrators and incident response team. Their expertise can help ensure that best practices are replicated when transitioning to a cloud-based system. That expertise also can help your organization meet technological expectations by producing compliance policy profiles and developing new incident remediation workflows. Netskope DLP helps reduce the demands on your DLP team, so your security teams will spend less time managing frustrating incidents and more time focusing on proactive initiatives that keep your company safe.
- **Value maturity over hype.** Success will require more than technical know-how. From developing metrics for top-level management down to guidance and action items for staff, you have a lot to consider. Make sure you lean on your vendor's support teams to help structure your journey and ultimately help you to unlock the value of the enterprise's innovation and make the journey worthwhile!

Security that's ready for anything

Data Protection

Netskope, a global SASE leader, is redefining cloud, data, and network security to help organizations apply zero trust principles to protect data. Fast and easy to use, the Netskope platform provides optimized access and real-time security for people, devices, and data anywhere they go. Netskope helps customers reduce risk, accelerate performance, and get unrivaled visibility into any cloud, web, and private application activity. Thousands of customers, including more than 25 of the Fortune 100, trust Netskope and its powerful NewEdge network to address evolving threats, new risks, technology shifts, organizational and network changes, and new regulatory requirements. Learn how Netskope helps customers be ready for anything on their SASE journey, **visit netskope.com**.

netskope

Build for a cloud-first future with modern DLP technology

Rapid adoption of the cloud and the trend toward work-from-anywhere make once-cutting-edge data protection techniques woefully inadequate. Data security efforts must provide consistent data protection everywhere data and people move. The ideal solution for modern data loss prevention (DLP) must be built for the cloud — not retrofitted for cloud use cases. It must apply zero trust techniques, reduce complexity, and provide consistent policy enforcement — everywhere.

Inside...

- Evaluate your approach to data protection
- Protect data and support business goals
- Learn how modern DLP works
- Minimize unauthorized access to data
- Simplify security policies while ensuring their effectiveness
- Safely move data to the cloud and among cloud applications

Carmine Clementelli is a cybersecurity expert and technology leader for data security, cloud security, zero trust and security service edge (SSE) at Netskope. He brings decades of experience as an author, speaker, and advisor, previously at Palo Alto Networks, Symantec, and other global organizations.

ISBN: 978-1-394-19891-7
Not For Resale

WILEY END USER LICENSE AGREEMENT

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