



Data Protection with Secure Internet and SaaS Access (ZIA™)

Reference Architecture — Zscaler for Users

Contents

About Zscaler Reference Architectures Guides	1
Who Is This Guide For?	1
A Note for Federal Cloud Customers	1
Conventions Used in This Guide	1
Finding Out More	1
Terms and Acronyms Used in This Guide	2
Icons Used in This Guide	3
Introduction	4
Zscaler Data Protection Technologies	5
Benefits of Data Protection	8
New to Data Protection?	8
Getting Started with Data Protection	9
Enable TLS/SSL Inspection of All Traffic	9
Managing Logs and DLP Incident Information	10
Cloud Application Security Broker	12
Inline, Out-of-Band, and Multimode CASB	13
CASB Visibility and Reporting	14
Using Inline CASB Identified Applications in Policy Definition	16
Zscaler Cloud DLP	18
Visualize and Report on the State of Your Data	20
Enable Controls Over Data and Refine Your Dictionaries	25
Fine-Tuning Control and Additional Inspection Tools	28
Out-of-Band CASB	33
Incident Receiver	33
SaaS Application Tenants	34
DLP Policy	35
Malware Policy	37
SaaS Security API Scan Schedules	38
SaaS Security Activities and Alerts	39

Cloud Browser Isolation	40
How It Works	40
Use Cases for Remote Browser Isolation	41
Summary	42
About Zscaler	42

About Zscaler Reference Architectures Guides

The Zscaler™ Reference Architecture series delivers best practices based on real-world deployments. The recommendations in this series were developed by Zscaler's transformation experts from across the company.

Each guide steers you through the architecture process and provides technical deep dives into specific platform functionality and integrations.

The Zscaler Reference Architecture series is designed to be modular. Each guide shows you how to configure a different aspect of the platform. You can use only the guides that you need to meet your specific policy goals.

Who Is This Guide For?

The Overview portion of this guide is suitable for all audiences. It provides a brief refresher on the platform features and integrations being covered. A summary of the design follows, along with a consolidated summary of recommendations.

The rest of the document is written with a technical reader in mind, covering detailed information on the recommendations and the architecture process. For configuration steps, we provide links to the appropriate Zscaler Help site articles or configuration steps on integration partner sites.

A Note for Federal Cloud Customers

This series assumes you are a Zscaler public cloud customer. If you are a Federal Cloud user, please check with your Zscaler Account team on feature availability and configuration requirements.

Conventions Used in This Guide

The product name ZIA Service Edge is used as a reference to the following Zscaler products: ZIA Public Service Edge, ZIA Private Service Edge, and ZIA Virtual Service Edge. Any reference to ZIA Service Edge means that the features and functions being discussed are applicable to all three products. Similarly, ZPA Service Edge is used to represent ZPA Public Service Edge and ZPA Private Service Edge where the discussion applies to both products.



Notes call out important information that you need to complete your design and implementation.



Warnings indicate that a configuration could be risky. Read the warnings carefully and exercise caution before making your configuration changes.

Finding Out More

You can find our guides on the [Zscaler website](https://www.zscaler.com/resources/reference-architectures) (<https://www.zscaler.com/resources/reference-architectures>).



You can join our user and partner community and get answers to your questions in the [Zenith Community](https://community.zscaler.com) (<https://community.zscaler.com>).

Terms and Acronyms Used in This Guide

Acronym	Definition
API	Application Programming Interface
AUP	Acceptable Use Policy
BYOD	Bring Your Own Device
CASB	Cloud Access Security Broker
CRM	Customer Relationship Management
CSV	Comma Separated Value
DC	Data Center
DLP	Data Loss Prevention
DMZ	Demilitarized Zone
EDM	Exact Data Match
FTP	File Transfer Protocol
GDPR	General Data Protection Regulation
GRE	Generic Routing Encapsulation
HIPAA	Health Insurance Portability and Accountability Act
HTTP	Hyper-Text Transfer Protocol
HTTPS	Hyper-Text Transfer Protocol – Secure
ICAP	Internet Content Adaptation Protocol
IDM	Indexed Document Match
IoT	Internet of Things
IPSec	Internet Protocol Security
ITSM	IT Service Management
NSS	Nanolog Streaming Service
OCR	Optical Character Recognition
PCI	Payment Card Industry
PHI	Personal Health Information
PII	Personally Identifiable Information
SaaS	Software as a Service
SASE	Secure Access Service Edge
SIEM	Security Information and Event Management
SSE	Security Service Edge
SSL	Secure Socket Layer (superseded by TLS)
SWG	Secure Web Gateway
TLS	Transport Layer Security
VM	Virtual Machine
ZDX	Zscaler Digital Experience
ZIA	Zscaler Internet Access
ZPA	Zscaler Private Access
ZTE	Zero Trust Exchange

Icons Used in This Guide

The following icons are used in the diagrams contained in this guide.

 Zero Trust Exchange	 Laptop with Zscaler Client Connector Installed	 Zscaler Client Connector on Phone	 Zscaler Client Connector on Desktop Workstation	 Zscaler Client Connector on IOT Device	 User Views the Page as an Image	
 Zscaler API	 Zscaler CASB	 Zscaler Log Router	 Zscaler Nanolog	 Zscaler NSS	 Cloud Browser Isolation	
 Data Center	 Branch Office Location	 Generic Cloud Application or Workload	 Headquarters Location	 Legacy Firewall	 Private Residence	
 Generic Application or Workload	 User Accessing an Application	 Log or Data File	 Internet	 SIEM	 Database	
 Data Tunnel	 Positive / True Badge	 Negative / False Badge	 TLS Inspection	 TLS Secured Connection	 Zscaler Issued Certificate	 Public CA Issued Certificate

Introduction

Protecting your organization's data from leaks and exfiltration requires a different approach in a cloud-first world. Before cloud adoption, your data was contained in your data centers, running on your networks, with endpoints issued and managed by your organization. Now your data and applications continue to migrate from your data centers for the public clouds and Software as a Service (SaaS) applications. Your users are shifting to hybrid or fully remote work, and bring your own device (BYOD) is the norm.

The legacy centralized security solutions that many organizations rely on now operate in a world they weren't designed for. Backhauling traffic to these solutions, only to send the traffic back out to the internet to reach cloud applications, is inefficient. Latency and lack of localization can significantly reduce your user's experience. This can drive shadow IT, where users avoid sanctioned apps for personal alternatives that are faster and easier to use but might not have proper security configured.

Secure Internet and SaaS Access (ZIA) data protection gives you the tools to identify and control cloud applications and limit sharing of sensitive data delivered in the cloud. All your traffic passes through the Zscaler Zero Trust Exchange where ZIA Public Service Edges inspect traffic in real time, applying multiple technologies to inspect and enforce policy. Zscaler inspects all traffic, including all TLS/SSL encrypted sessions.

SaaS scanning for data at rest happens via API inspection of your applications. The scanning occurs regularly and looks for sensitive information being stored incorrectly. This could be collaboration tools where the information shared is not appropriate for the channel or folder where it's residing. When the system finds a violation, it alerts you and takes action to remediate the issue automatically.

With over 150 data centers around the world, your traffic passes through to the geographically closest data center for inspection. At each data center sit several ZIA Public Service Edge instances that inspect your traffic in real time. The same policy is applied no matter where you are in the world, providing a consistent and predictable access to cloud applications and websites.

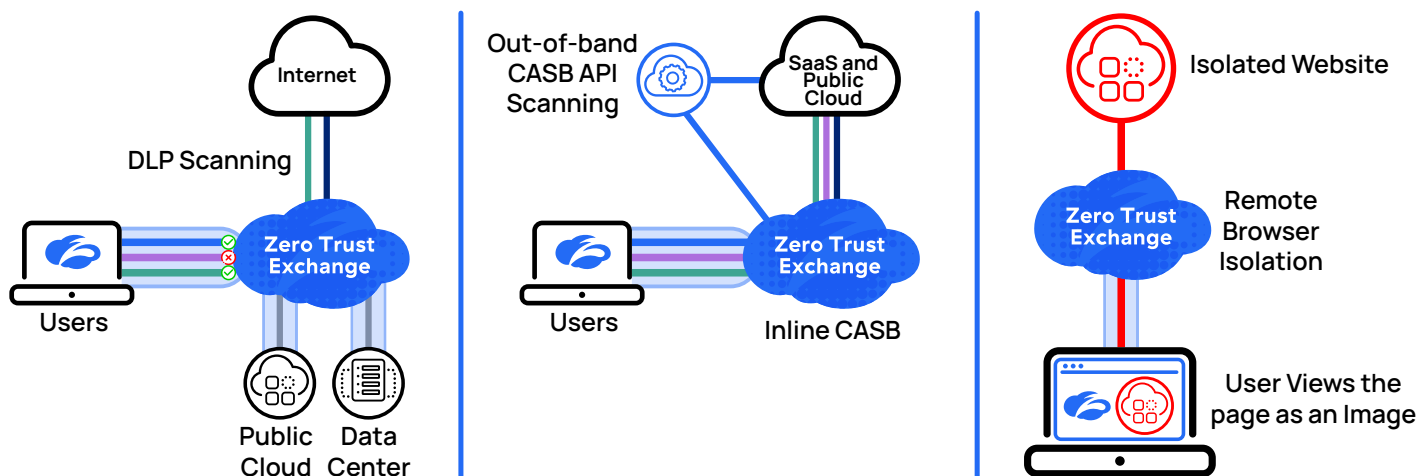


Figure 1. The Zscaler Zero Trust Exchange enforces policy and applies CASB, DLP, and remote browser isolation technologies

Technologies for data protection include cloud access security brokers (CASB) both inline and out-of-band, data loss prevention (DLP), and remote cloud browser isolation. Together these tools give you the visibility and control to protect your data at rest or in motion through your organization. This includes identifying and controlling access to cloud applications, preventing data leakage through file sharing sites and social media, and discovering shadow IT applications. A strategic partnership with Microsoft enables MIP integration with dynamic labels for easy identification and actions based on document type. Auditing your organization for compliance and reporting activities are easily achieved through the same interface.

Zscaler Data Protection Technologies

Zscaler Data Protection technologies are cloud services for identifying and protecting your sensitive data. This can include items such as personally identifiable information (PII), personal health information (PHI), payment card industry (PCI), financial records, and more. Sophisticated application controls let you allow or deny entire applications, or a subset of actions a user might take. Documents can be indexed and identified in file upload attempts, and you can index your data to prevent exfiltration.

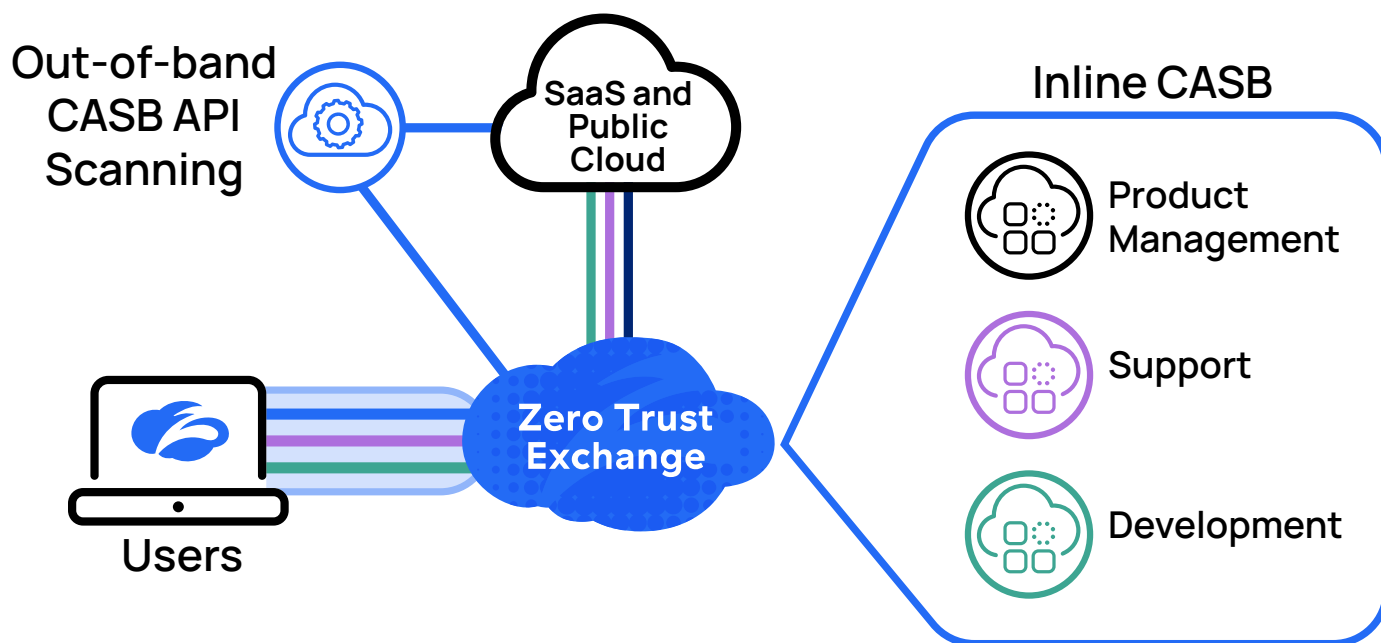


Figure 2. CASB identifies your applications in use and allows you to write policy specific to an application

CASB is a set of technologies that gives your organization the ability to understand and control cloud-based applications. With CASB, you gain insight into what applications are in use, where your data is stored, and how your data moves through the cloud. CASB is also a reporting and discovery mechanism that helps you ensure compliance with regulations and reporting requirements.

CASB has evolved from a tool for identifying shadow IT applications to identifying and controlling access to cloud applications. CASB itself has grown to include both inline (proxy) inspection and out-of-band scanning via APIs. With inline scanning, CASB acts as a proxy for connections, ensuring that data is protected as it moves between your users and the cloud. With out-of-band scanning, the CASB system can signal back enforcement actions via APIs to the user's security gateway. The two CASB types can be used together for more comprehensive scanning, called multimode CASB. Out-of-band CASB leverages Zscaler Cloud DLP and Zscaler malware detection.

Users leveraging bring your own device (BYOD) policies can be both a benefit and a risk to the organization. Zscaler Data Protection can enforce Zscaler inspection for corporate applications. Zscaler identity proxy requires users to go through the Zscaler Zero Trust Exchange to access the applications. Users attempting to access the application directly will be denied access. This prevents BYOD users from attempting to go around security controls. To access these applications, users need to install Zscaler Client Connector, be at a trusted site with a direct link to ZTE, or leverage browser access.

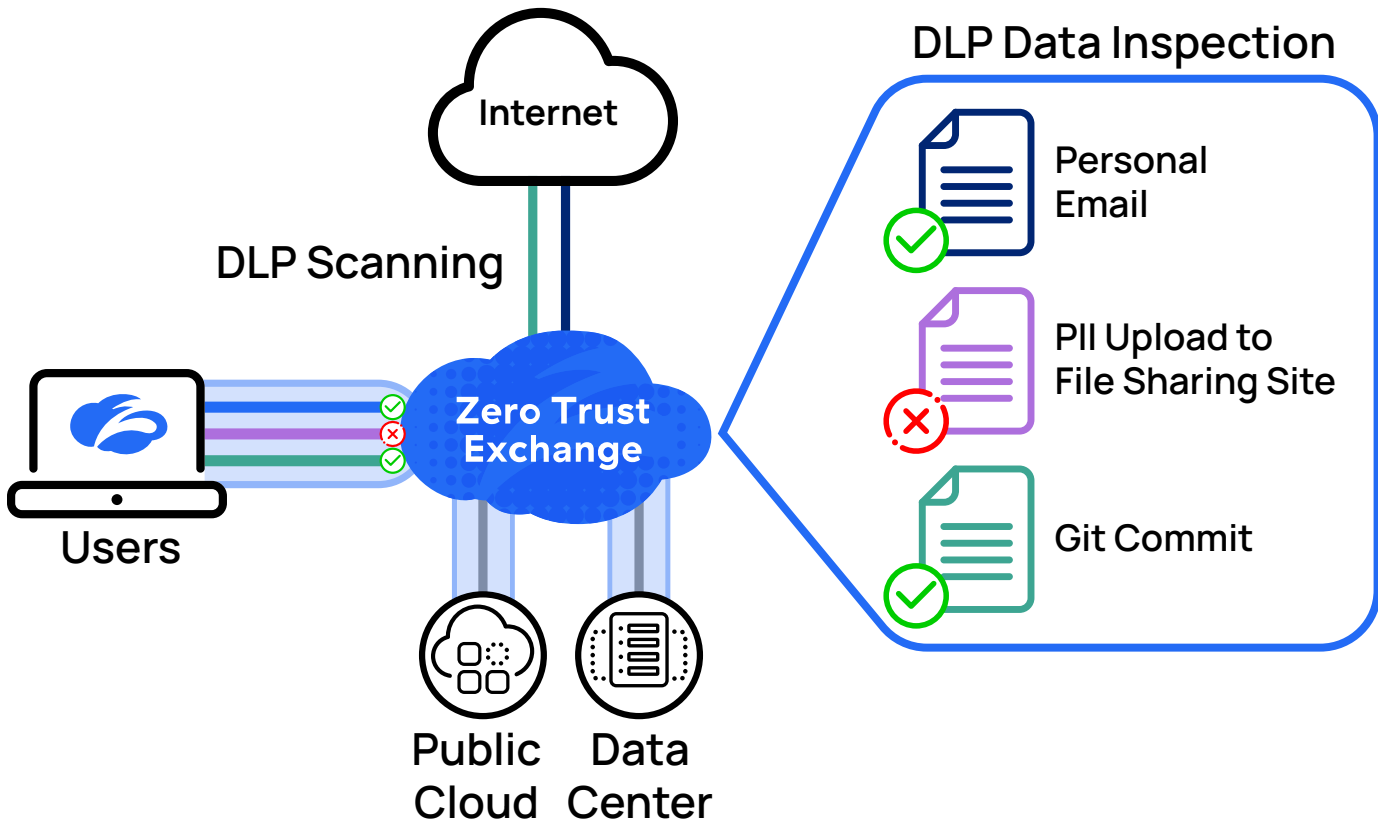


Figure 3. Data loss prevention for your data being transmitted or stored in the cloud

With Zscaler Cloud DLP, you monitor and inspect data on your network to prevent exfiltration of critical information and intellectual property. Many common dictionaries are available to match items such as credit card or social services numbers in various countries. As you analyze your data, you'll start to develop your own dictionaries specific to your data. Using the Zscaler Index Tool, you can fingerprint documents in your existing document stores. Advanced features such as exact data match allow you to describe and fingerprint your data, and then look within documents for exact matches. These same dictionaries and engines are used to scan documents at rest as a part of the out-of-band CASB scanning.

Zscaler DLP supports many predefined dictionaries to identify PCI, PII, PHI, and other standard data classifiers. These dictionaries were built based on standard regex and hyperscan-based libraries. Customers can also build custom dictionaries based on keywords, phrases, regex, and proximity keywords.

Zscaler DLP supports out-of-the-box DLP engines that detect PCI, PII, Health Insurance Portability and Accountability Act (HIPAA), and General Data Protection Regulation (GDPR) data. Zscaler Advanced Data Classification supports exact data match (EDM), index document match (IDM), and optical character recognition (OCR) to protect data on image files, embedded images, screenshots, and handwritten texts.

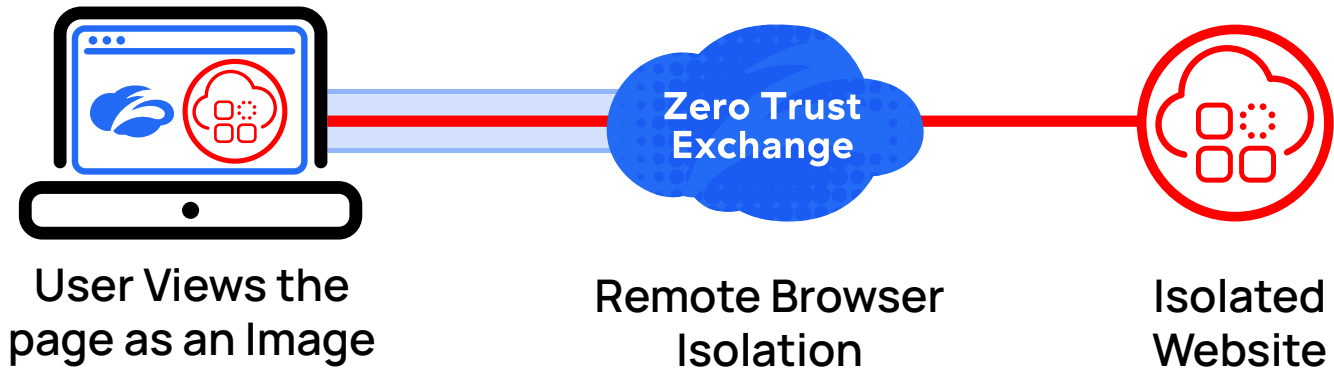


Figure 4. Cloud browser isolation separates the user from the website and its content

Some users require an extra layer of protection between themselves and their web destinations. These are typically your users who have access to confidential information or are subject to phishing-type attacks. They could also be devices that can't install a software agent, such as internet of things (IoT) devices like printers and cameras.

For these users, remote browser isolation provides a separation between the website and the user's browser. With remote browser isolation, a user's web session takes place on a remote server and a rendered version of the page is sent to the user. By completely separating the user from the browser, attacks against the browser or OS are not possible. You can also prevent users from saving files locally.

Deploying Zscaler Data Protection is performed in stages:

1. Enable TLS/SSL inspection for full visibility of all traffic.
2. Leverage CASB policy development and detection of unsanctioned applications.
3. Deploy Zscaler Cloud DLP and fine-tuning of dictionaries and engines to reduce false positives.
4. Deploy Zscaler out-of-band-CASB to scan your SaaS application tenants for data leaks and malware detection.
5. Remote browser isolation deployment for critical users, services, and untrusted devices from SaaS applications.

ZIA data protection provides robust security for your organization's cloud applications, data, and users. You can regain visibility into your data and user's actions, as well as fine-grained controls to enforce policy.

Benefits of Data Protection

- Unified protection – Zscaler Data Protection provides consistent unified security for data in motion and data at rest across SaaS and public cloud applications.
- Full TLS/SSL inspection of all traffic – Around 95% of outbound traffic is encrypted. Unlike static appliance-based tools, Zscaler has the capacity needed to inspect all TLS/SSL traffic.
- Compliance reporting and remediation – Zscaler enables unified compliance across SaaS and public cloud apps, measuring configurations against 17 frameworks and automating remediation.
- Elastic scale with inline enforcement – Zscaler secures data in real time, not after the fact. Its services are user-based, not capacity-based, so security scales with performance guaranteed by SLAs.

New to Data Protection?

If you are new to CASB or Zscaler and the Security Service Edge, see the links below to find out more.

- For a brief introduction to CASB, see [What Is A Cloud Access Security Broker \(CASB\)?](https://www.zscaler.com/resources/security-terms-glossary/what-is-cloud-access-security-broker) (<https://www.zscaler.com/resources/security-terms-glossary/what-is-cloud-access-security-broker>).
- For a brief introduction to Zscaler Cloud DLP, see [What Is DLP?](https://www.zscaler.com/resources/security-terms-glossary/what-is-dlp) (<https://www.zscaler.com/resources/security-terms-glossary/what-is-dlp>).
- For a brief introduction to remote browser isolation, see [What Is Remote Browser Isolation?](https://www.zscaler.com/resources/security-terms-glossary/what-is-remote-browser-isolation) (<https://www.zscaler.com/resources/security-terms-glossary/what-is-remote-browser-isolation>).

Getting Started with Data Protection

The widespread adoption of SaaS and public cloud apps has rendered data widely distributed and difficult, if not impossible, to secure with legacy on-premises appliances. As a result, it is easy for both careless users and malicious actors to expose enterprise cloud data. Zscaler Data Protection follows users and the apps they access, continuously protecting against data loss. The Zero Trust Exchange (ZTE) inspects traffic inline, encrypted or not, and ensures your SaaS and public cloud apps are secure, delivering protection, visibility, and regulatory compliance.

Data protection focuses on protecting your organization's information from intentional or accidental release. To do that, Zscaler needs to look inside your traffic. Today, encrypted traffic makes up the 95+% of internet traffic. These technologies open the encrypted streams and application flows for inspection, giving you visibility into how data is used and moves through your clouds. They also give you the ability to restrict access to applications down to granular controls over users and actions within approved applications.

Zscaler achieves this by first decrypting all your TLS/SSL encrypted traffic. For many organizations this has not previously been possible, as hardware appliances are often rated at only a fraction of the traffic when decryption is enabled. Zscaler assumes all traffic will be decrypted, and that you might occasionally opt traffic out of inspection. When you gain visibility into your traffic, you can then start to see what your users are doing and where your data is moving.

Enable TLS/SSL Inspection of All Traffic

Getting data protection tools in place requires that you inspect all your traffic, including traffic that is encrypted with TLS/SSL. For many organizations, this is a large change as previous security appliances typically can't handle decryption at scale. With Zscaler and the ZTE, this scalability is not an issue. ZIA Public Service Edge data centers are designed to handle full decryption and inspection of all your traffic.

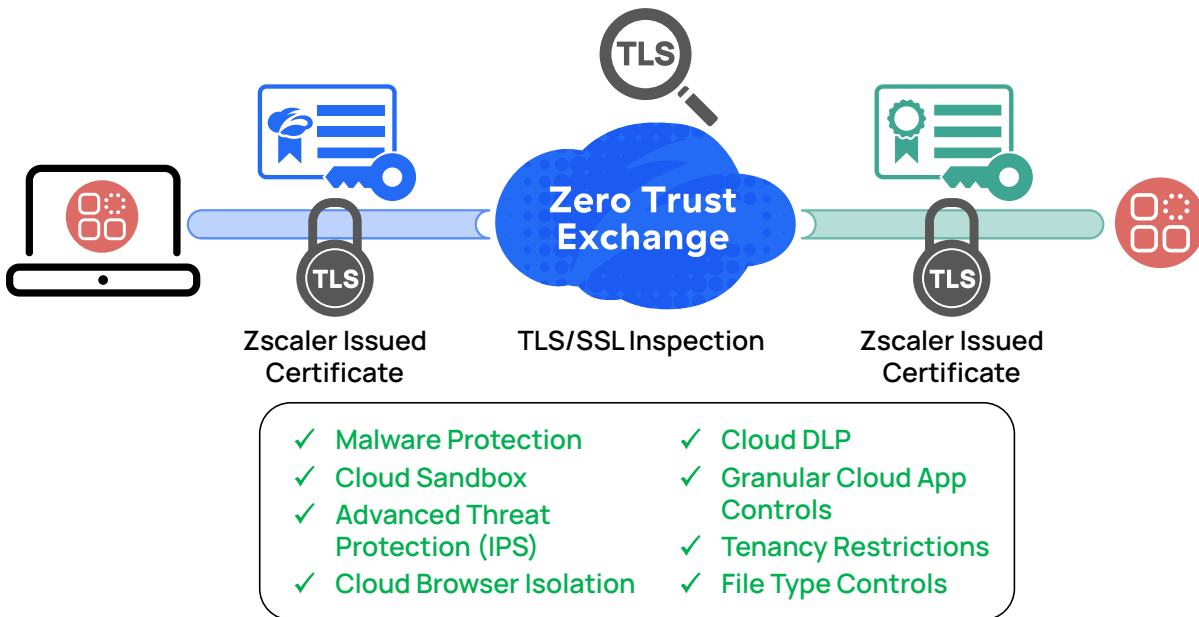


Figure 5. Zscaler TLS/SSL Inspection allows all your security subscriptions to be used on all traffic

Enabling TLS/SSL inspection is beyond the scope of this guide. If you have not already enabled TLS/SSL inspection in your organization, we recommend the reference architecture [TLS/SSL Inspection with Zscaler Internet Access](https://help.zscaler.com/zia/tls-ssl-inspection-zscaler-internet-access) (<https://help.zscaler.com/zia/tls-ssl-inspection-zscaler-internet-access>). This critical first step ensures you have visibility into all your data.

Learn more at [About TLS/SSL Inspection](https://help.zscaler.com/zia/about-ssl-inspection) (<https://help.zscaler.com/zia/about-ssl-inspection>).

Managing Logs and DLP Incident Information

Zscaler Data Protection services generate multiple logs that provide insights into how your organization moves and stores information. Web security logs and SaaS security insight logs are available for review in the ZIA Admin Portal for 180 days. The ZIA Admin Portal provides robust filtering capabilities to isolate issues.

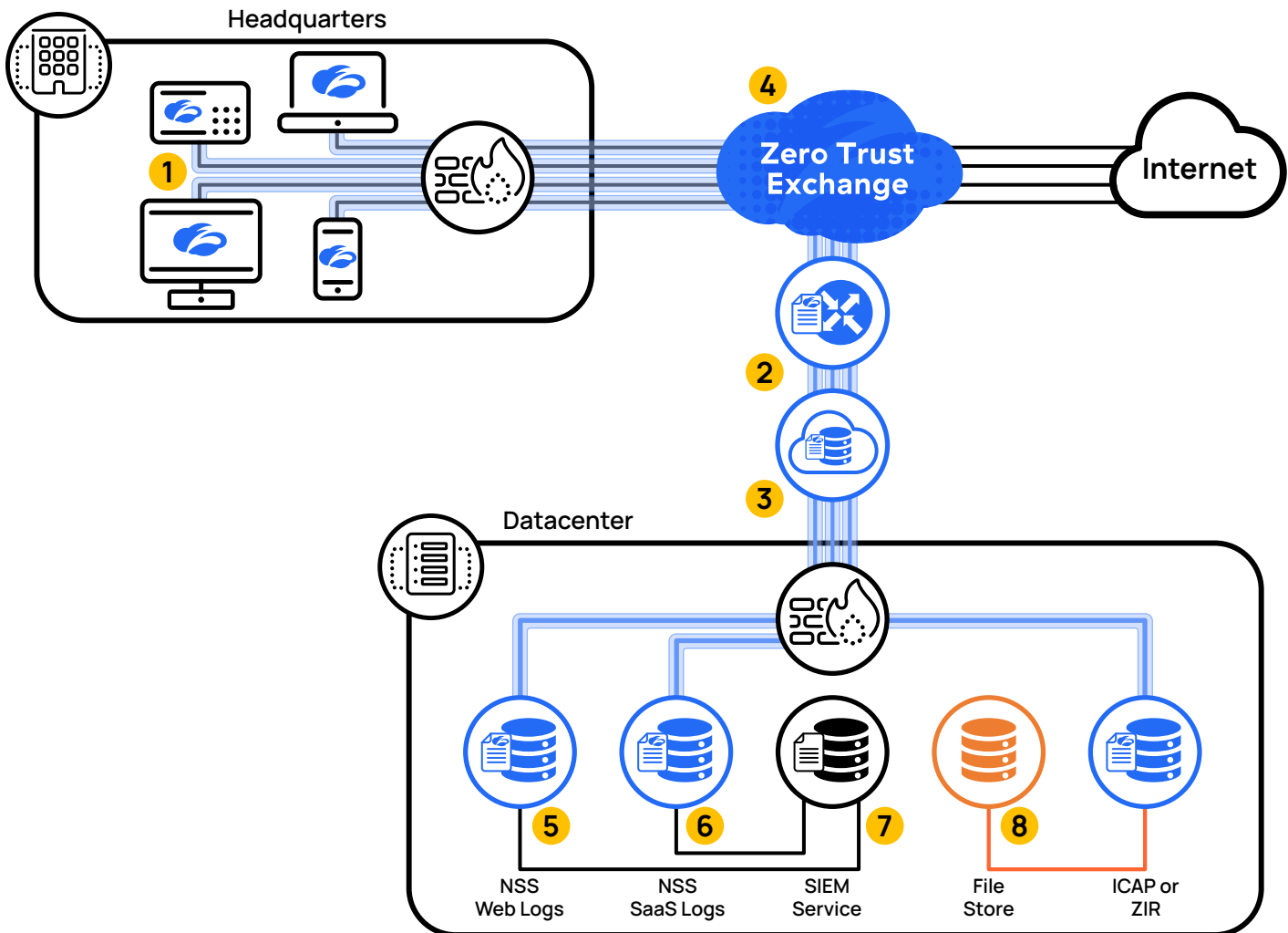


Figure 6. The Nanolog Streaming Service (NSS) receives logs from ZIA, decompresses and detokenizes the logs, then streams them to your SIEM servers

1. Logs are generated as your users interact with the internet via the Zscaler Zero Trust Exchange. Every transaction is logged using a compressed format leveraging WAN compression techniques.
2. These logs are sent via log routers to the Zscaler Nanolog server. No logs are stored on the ZIA Service Edge itself.
3. The Zscaler Nanolog servers sit in different regions around the world. When you set up your Zscaler service, you select the region where your organization's logs are stored.

4. Your logs are viewable via the Zscaler Central Authority interface in near real time. You can search and filter logs from the last 180 days.
5. If you require longer retention of logs, a log stream can be sent to your SIEM event manager via an NSS virtual appliance. Because web logs and SaaS incident report logs are different streams, they require two separate NSS servers, one per log stream.
6. The NSS decompresses your logs and filters the output if needed to meet your policy. From there, the logs are streamed to your SIEM server.
7. Workflow automation is configured to direct DLP incident data and metadata files.
8. A Zscaler Incident Receiver is deployed in the data center to receive DLP incident data and metadata files for inspection. The Zscaler Incident Receiver does not store files itself, but acts as a destination for the log stream and then writes them to local storage you provide.

Zscaler web logs and SaaS security insight logs are stored for 180 days in the Zscaler cloud. If your organization requires additional storage, the Zscaler NSS provides log integration to your SIEM. The NSS virtual machines receive logs from the Zscaler Nanolog cluster that holds your log files in near real time. Based on your configuration, the NSS reformats, filters, and delivers the logs streams to your SIEM service. Your SIEM service can be local or in the cloud. Zscaler requires two instances of NSS to receive both web security logs and SaaS security insight logs.

For DLP violations, a different service is required. When a DLP violation occurs, the Zscaler service forwards the incident data and metadata files for you to investigate. If you already have an existing internet content adaptation protocol (ICAP) server deployed for DLP reporting, Zscaler can leverage that system. If you don't have an existing ICAP server, or your ICAP server cannot be made accessible to the internet, you can deploy a Zscaler Incident Receiver virtual machine. The Zscaler Incident Receiver instance will accept the incoming log stream and write it as files to a storage location you provide.

NSS and SIEM Integration

The NSS is a virtual appliance that provides the ability to save logs locally. The NSS streams log files from the Zscaler cloud to your SIEM server for retention and analysis. The most common use cases are:

- Your organization is required to store logs for greater than the 180 days ZIA provides in the cloud.
- Your organization wants to correlate Zscaler logs with other systems' logs in a single application.
- Your organization is subject to regulatory mandates requiring local storage of logs.



The NSS requires an additional subscription from Zscaler for the virtual appliance. You will require two subscriptions and instances of NSS to receive both web logs and SaaS security insight logs. Please contact your Zscaler Account team for more information.

- To see a list of Zscaler partners for SIEM integration, see [Zscaler + SecOps \(https://www.zscaler.com/partners/technology/operations\)](https://www.zscaler.com/partners/technology/operations).
- To learn about deploying NSS and SIEM integrations, see [About Nanolog Streaming Service \(https://help.zscaler.com/zia/about-nanolog-streaming-service\)](https://help.zscaler.com/zia/about-nanolog-streaming-service).

Zscaler Incident Receiver

The Zscaler Incident Receiver is a virtual machine that allows you to securely receive information about DLP policy violations. The Zscaler service sends information about policy violations via ICAP to the incident receiver. This tool sends the policy-violating content and a JSON file containing the DLP policy scan metadata (e.g., the URL, DLP dictionaries, DLP engines, etc.) to a local file server you provide for this purpose.

The incident receiver is only available for organizations that do not already have an ICAP service available, or cannot make one available to external services. If your organization has an existing ICAP server, the Zscaler service can leverage that server directly. If you need to allow the ICAP stream through your firewall, Zscaler provides a list of IP addresses for you to configure the necessary firewall rules. You can find the complete list of Zscaler IP addresses at [Zscaler Config \(https://config.zscaler.com/\)](https://config.zscaler.com/).

- To learn more about Zscaler incident receivers and workflow automation, see [Managing DLP Application Integrations in Workflow Automation \(https://help.zscaler.com/zia/managing-dlp-application-integrations-workflow-automation\)](https://help.zscaler.com/zia/managing-dlp-application-integrations-workflow-automation).
- Information on the various Zscaler logs are available on the [help site \(https://help.zscaler.com/zia/documentation-knowledgebase/analytics/dashboards-reports-and-logs/logs\)](https://help.zscaler.com/zia/documentation-knowledgebase/analytics/dashboards-reports-and-logs/logs).

Cloud Application Security Broker

Cloud-based applications and infrastructure changed how organizations looked at access to the web. Shadow IT, the use of unapproved applications and network equipment in the organization, has been a problem for IT managers for decades. With applications now easily procured with a credit card, this led to many organizations seeing the use of unapproved applications drastically increase. The tools that would become grouped under the term cloud access security broker (CASB) started out combating this new wave of shadow IT. These tools brought visibility to the organization by identifying and categorizing the applications in use. Traffic was analyzed and policy actions were sent to on-site enforcement devices via API calls.

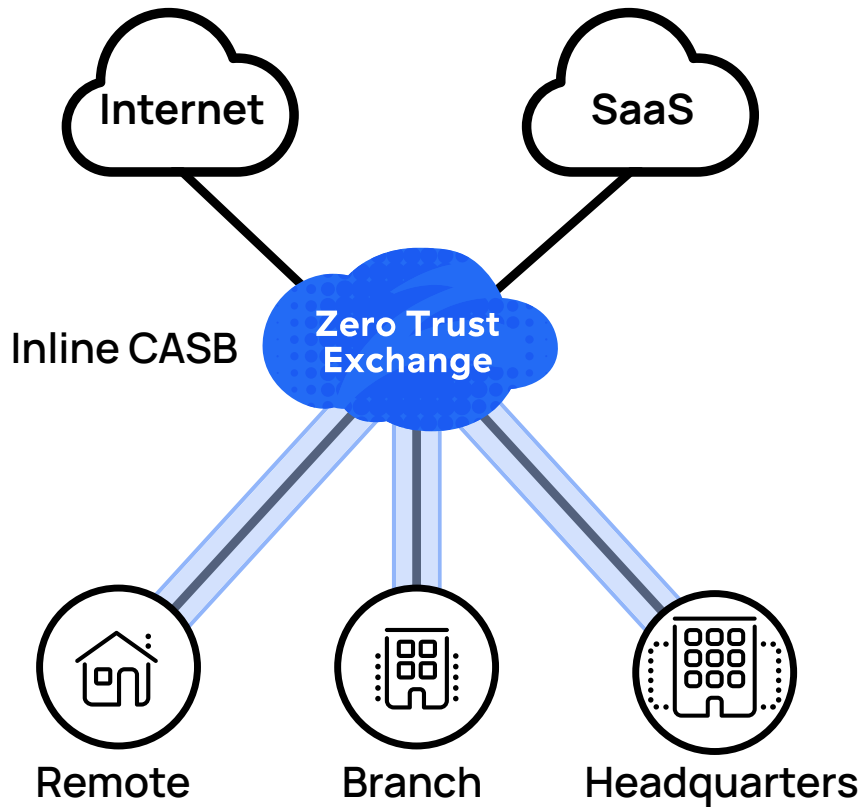


Figure 7. Inline CASB scans traffic to identify and apply policy to applications

With inspection into the data, policy can be applied to specific applications. As your users access cloud applications, you can allow and deny, but you can also get more granular control. CASB inspects and tracks what your users are doing and gives you sophisticated controls over their actions.

CASB inspects data with pattern-matching tools to find sensitive data. Based on the data match and the distribution scope of the channel or file share, the CASB system can take remedial action. This remedial action is dependent on the specific integration and can include disabling collaboration scope through links or explicit collaborative calls, alerting, and notification.

With visibility, these applications and actions become inputs to policy. When combined with a secure web gateway (SWG), CASB application identification provides policy inputs for control of sanctioned and unsanctioned applications in real time.

Inline, Out-of-Band, and Multimode CASB

CASB operates in one of two modes: inline and out-of-band. When operating inline, applications are identified by the inspection engines. This allows you to apply policy controls to specific applications in real time. ZIA provides out-of-band CASB scanning using APIs to connect to your applications and storage providers. Files and data are scanned by Zscaler Cloud DLP, and anti-malware engines scan data at rest.

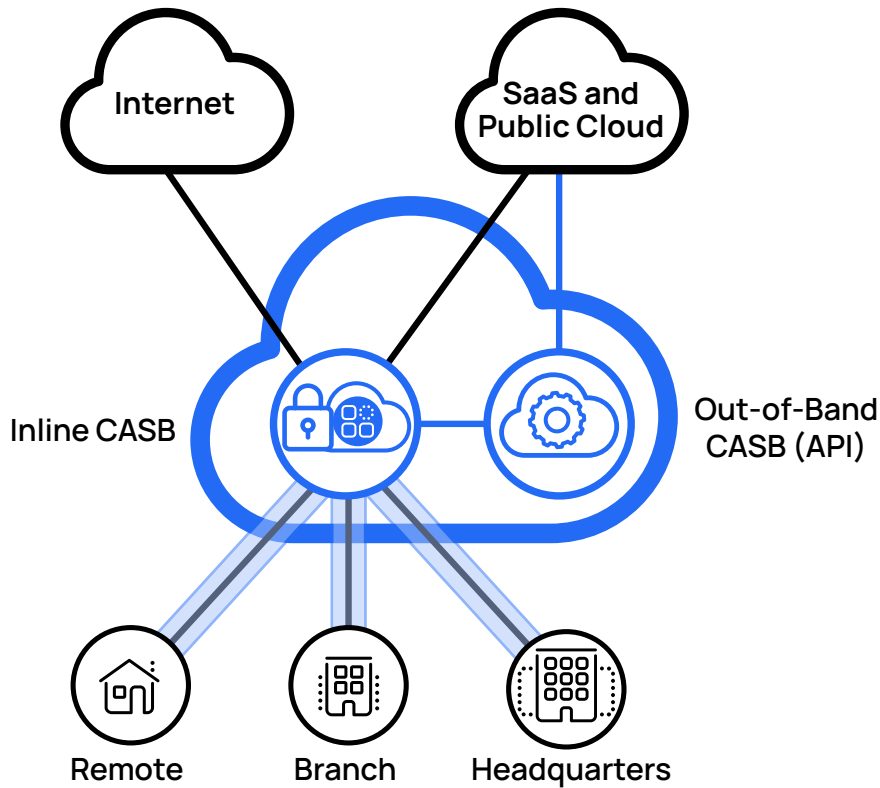


Figure 8. Inline and out-of-band CASB work together to provide robust coverage of data in motion and at rest

These two approaches provide more coverage when used together than you can achieve individually. When combined, these tools are referred to as multimode CASB. For example, you might use inline CASB to prevent posting of sensitive data, or to control the use of cloud applications by your users. Users are prevented from posting confidential content or using unapproved applications in real time.

You might also leverage out-of-band CASB scans for data leaks, as well as looking for malicious files in cloud applications that might have been uploaded by partners or customers. Leveraging API calls for inspection and reporting, out-of-band CASB can also be a part of your compliance strategy. Both inspection and reporting are achieved for your SaaS and cloud applications. Together, the two CASB solutions keep your users safe while accessing cloud applications and keep the data they are accessing secure. This chapter primarily focuses on inline CASB, with [Zscaler Cloud DLP](#) covered in the next chapter.

CASB Visibility and Reporting

Reporting requirements for regulatory and compliance are greatly simplified by leveraging CASB identification and risk classification. By inspecting all your data and matching it to applications and application classes, you gain a big picture view of your data, cloud applications, and users. This allows you to produce reports to meet regulatory, industry, or internal reporting requirements.

ZIA reporting allows you to also view your application usage through the lens of compliance standards. This lets you quickly identify applications or behaviors that would violate certification requirements so that your organization can take corrective action before the issue becomes a larger problem.

Privacy controls in both reporting and logging can be used to obfuscate usernames and device information. This prevents admins from tying an action back to a particular user where privacy policy or regulations protect such information. In this case, an auditor role should be established for when this information is required to be disclosed. That individual can be either a Zscaler Admin or someone separate from the admin team.

This visibility comes in multiple forms, from logging to interactive reporting and application discovery. This section covers the options for viewing your data as logs in real time. Zscaler offers the Nanolog Streaming Service (NSS) that enables integration with SIEM servers.



While SIEM integration for reporting does not require an additional license, it does require feeds to be enabled for the SIEM server. Contact Zscaler Support or your account team for more information.

- To see a list of cloud applications by category, see [About Cloud App Categories](https://help.zscaler.com/zia/about-cloud-app-categories) (<https://help.zscaler.com/zia/about-cloud-app-categories>).
- To learn more about administrator roles, see [Administration & Role Management](https://help.zscaler.com/zia/authentication-administration/administrator-role-management) (<https://help.zscaler.com/zia/authentication-administration/administrator-role-management>).

Interactive Reports

ZIA provides extensive reporting and near real-time information on what is happening with your organization's users and data. The Zscaler Central Authority provides a consolidated view of all the logs for your organization into a single pane of glass. Extensive drill downs let you discover which applications, departments, and users have the most violations, website visits, and application uses.

ZIA provides many predefined reports to address common reporting needs, including SaaS reporting and the shadow IT report. You can also build your own reports for custom use cases. Reports can be restricted by an administrator's role, limiting visibility and scope of the reports. Reports can be run on-demand in an interactive mode or scheduled for delivery. You can define multiple reports to support different parts of your IT team according to their needs.

If you need to provide reporting access to auditors or similar, it's possible to create limited admin roles that only allow users to manage reporting. This can be limited to read-only, or read and create reports.

Find out more about configuring and running reports at [Dashboard Reports and Logs](https://help.zscaler.com/zia/documentation-knowledgebase/analytics/dashboards-reports-and-logs/reports) (<https://help.zscaler.com/zia/documentation-knowledgebase/analytics/dashboards-reports-and-logs/reports>).

Application Discovery and the Shadow IT Report

One of the most important reports from a data protection viewpoint is the shadow IT report. In this report, you can search and filter applications to find those who violate policy. This report gives you the ability to filter by:

- Application category – Report on all applications or just specific categories.
- Risk index – All applications are given a risk index from 1 to 5, with 5 being the riskiest.
- Sanctioned state – Filter by all, sanctioned, and unsanctioned.
- Certifications – ZIA reporting supports many certification programs. You can select any certifications that apply to your organization to see a list of applications that are approved or unapproved.
- Hosting and security characteristics – This report filter lets you look for weakly protected or suspicious applications by how the application operates. This is a combination of technical capabilities and application actions that have been observed for an application in use. For example, you can check for TLS 1.1 support, but also if the application has suffered a reported data breach in the last three years.

This report can help you quickly understand unauthorized application access and the extent of the shadow IT issue. Zscaler recommends routinely checking this report for unsanctioned applications and taking the necessary remediation steps.

Insights Logs

As with most systems, ZIA provides extensive logging capabilities that provide visibility into the entirety of your organization. Zscaler uses a proprietary mechanism to highly compress logs from the ZIA Public Service Edges. No logs are stored on the ZIA Public Service Edge devices. Instead, logs are compressed and sent to log routers in the Zero Trust Exchange cloud, and then forwarded to the log server in your organization's region of choice.

ZIA's log compression means that extensive logging is possible for all users and transactions. Logs are available for the following groupings:

- Threat insights – 2D map or 3D globe representation of threats against your organization and their geographic origin.
- User insights – Provides visibility into top users and web destinations they visit.
- Web insights – Information on web browsing, including OS, agent, IP addressing, etc.
- Mobile insights – Information on mobile devices, applications, usage, and more.
- Firewall insights – Logs of firewall requests and transactions.
- DNS insights – Logs of DNS requests and responses.
- Tunnel insights – Provides log visibility into events around GRE and IPSec connections to Zscaler.

All logs are accessible from the Zscaler Central Authority. This provides a consolidated view, based on your permissions, of your entire organization or the portion that the admin can view. Logs are stored for 180 days in the ZTE cloud.

To view all available insight logs variables and more details on each category, see the [Help site \(https://help.zscaler.com/zia/documentation-knowledgebase/analytics/dashboards-reports-and-logs/logs\)](https://help.zscaler.com/zia/documentation-knowledgebase/analytics/dashboards-reports-and-logs/logs).

Using Inline CASB Identified Applications in Policy Definition

Building access policy involves analyzing your application usage and user traffic patterns. Your policy rules are based on a combination of your acceptable use policy (AUP) and approved application list. Policy definition can match against multiple user data points to allow, restrict, or deny access to applications. The destination can be expressed in several ways:

- Fully qualified domain name (FQDN)
- Application category or application identify
- IP address or range

Using application categories and the individual applications within them is a simple way to put controls in place over a range of cloud-based applications. In this model, you identify specific applications from within a category to apply your allow permissions. These are your sanctioned applications. When your approved applications are allowed, you can block or caution users on the rest of the category.

The criteria for allow can require that a user be on a device with a secure connection, using a current browser and OS, etc. You might limit it further, to say your finance tools can only be accessed from certain offices within the network. Together the user, device posture, destination, and user action (sometimes) result in a yes, no, or caution.

When your users attempt to access a resource, the rules are evaluated by comparing their posture and action to the policy rules to find the most specific match. When a match is made, the policy is applied. This might be different than what you have experienced with other tools, where first match wins. With Zscaler, the most specific policy condition match wins.

Applications are classified by Zscaler CASB into categories of applications with a similar purpose or functionality, such as business communication or social media. The application categories fall into one of two groups: those that support allow or block, and those that support additional user controls.

Application Categories That Support Allow or Block

The following list of applications supports allow or block options:

- Collaboration & Online Meetings
- Consumer
- DNS Over HTTPS Services
- Finance
- Health Care
- Hosting Providers
- Human Resources
- IT Services
- Legal
- Productivity & CRM Tools
- Sales & Marketing
- System & Development

Application Categories That Support Action-Based Controls

The following list of applications supports allow and block options, in addition to specific controls based on the applications themselves. For example, the webmail category supports separate controls for reading, sending, and attaching files to mail messages.

- File Sharing
- Instant Messaging
- Social Networking
- Streaming Media
- Webmail

Policy Rule Selection

Zscaler recommends putting your rules into order of most specific to least specific. This often makes the policy easier to understand for new administrators. For example, if an organization wants to set up rules for allowing OneDrive from the corporate tenant, but blocking personal tenants and other file sharing, you can build out the following rules:

Rule Order	Source	Device	Destination	Outcome
1	User	Org Issued	OneDrive(ORG)	Allow
2	User	Any	OneDrive(Personal)	Deny
3	User	Any	File Sharing Category	Deny

The rules are interpreted in order from lowest to highest. In this case, our most specific rule is to allow access to the organization's OneDrive account first. Users on an organization-issued device have access to the organization's OneDrive account. The next two rules specifically rule out other services and non-corporate issued devices at the same time. The first rules out access to personal OneDrive accounts from any device. The second rules out all other file sharing services, including access to the organization's OneDrive account with BYOD devices.

Zscaler Cloud DLP

The technologies used in data loss prevention monitor and inspect your organization's data to prevent exfiltration. ZIA Zscaler Cloud DLP examines the data both in motion and at rest in your cloud applications. You can view a complete list of supported cloud applications at [Adding SaaS Application Tenants \(https://help.zscaler.com/zia/adding-saas-application-tenants\)](https://help.zscaler.com/zia/adding-saas-application-tenants).

The same data match and indexing engines are applied to all data moving through a ZIA Service Edge or via API scanning of a cloud application (see [Out-of-Band CASB](#) in this guide). The data is examined in several ways:

- Files are compared against an indexed set of files for your organization.
- For any data in motion or files that don't have an indexed equivalent, the Zscaler Cloud DLP engines examine the file contents and any submitted HTTP form data in any application, looking for organizational data being exfiltrated.
- Exact data match and custom dictionaries can be created to specifically match your personally identifiable information (PII).

To find your data, the Zscaler Cloud DLP service must first be trained to recognize it. While some data is relatively common, such as credit card numbers, others are unique to your organization. This is a process that occurs in phases as your data recognition becomes more refined and additional features are enabled. This policy refinement helps reduce your false-positive matches. As your results become increasingly accurate, your actionable matches increase. There are four types of DLP dictionaries available:

1. Patterns and Phrases – Using a combination of regular expressions and exact phrases, Zscaler DLP monitors for leaks of information to unauthorized applications.
2. Index Data Match – This dictionary allows you to fingerprint your most sensitive documents which are then leveraged by Zscaler DLP to catch documents leaving the organization.
3. Exact Data Match (EDM) – With EDM, you can fingerprint your structured data, such as your customer or user PII, giving you the ability to watch for leaks of specific data.
4. Microsoft Information Project (MIP) Labels – Zscaler DLP is capable of leveraging MIP labels in your Microsoft 365 data stores and taking actions based on those labels.

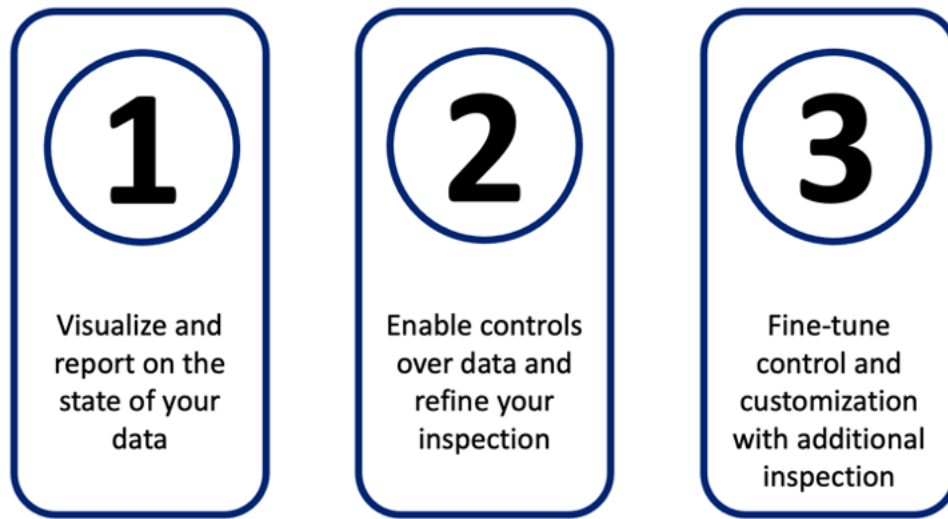


Figure 9. The three phases of Zscaler Cloud DLP deployment involve learning to recognize your data and then act

1. Visualize and report on the state of your data.
2. Enable controls over data and refine your inspection.
3. Fine-tune control and customization with additional inspection.

In the first phase, you leverage Zscaler Cloud DLP to learn and understand where your data is stored and how it moves through your clouds. You won't be enforcing yet. Through reporting and altering, you'll begin to set up workflows for notifications of violations. For this phase, you'll leverage preconfigured dictionaries and engines. Zscaler Cloud DLP comes with multiple dictionaries trained to recognize PII, such as government identity number, credit cards, and passport numbers. Finally, you'll configure file indexing using Zscaler's Index Tool to create Indexed Document Match (IDM) templates to fingerprint your files.

In the second phase, you enforce and refine your policies and restrict users' ability to share your confidential data. You'll begin to develop custom dictionaries that match your organizational data norms. False-positive alerts decrease with more data fine-tuning. You'll start to build policy to restrict data from leaving your organization to cloud destinations and applications that are not authorized.

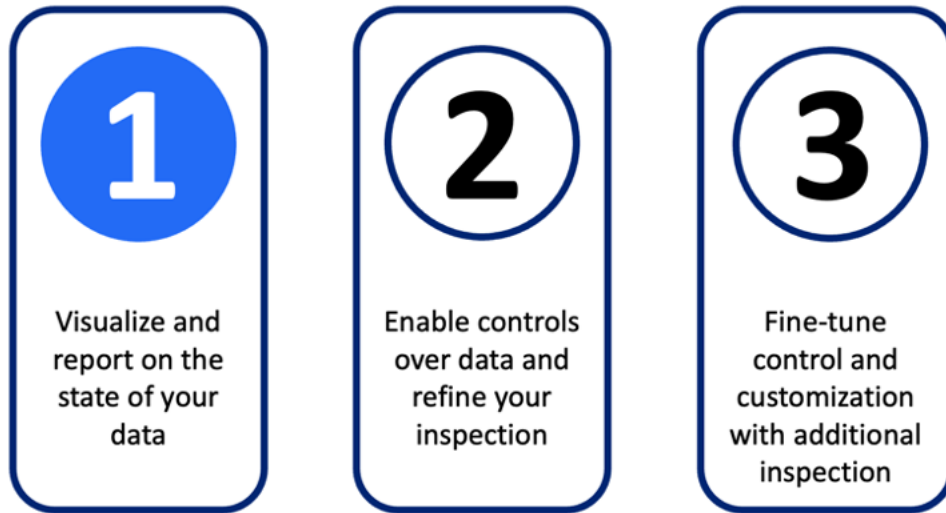
The third phase begins to move into advanced DLP tools and topics. Exact data match allows you to go beyond templates by indexing your PII data directly. This gives you the ability to look for specific data within documents to enforce compliance rules. Optical character recognition (OCR) of image files allows you to find data hidden in non-text files. At this stage, dictionary refinement comes in the form of additional regex improvements. Block policies can also be adjusted if you are seeing many false positives.



Some features discussed in this chapter require additional licensing. To see what is included at each licensing level, see [Zscaler Data Protection at a Glance \(https://www.zscaler.com/resources/data-sheets/zscaler-data-protection-benefits.pdf\)](https://www.zscaler.com/resources/data-sheets/zscaler-data-protection-benefits.pdf).

To learn more about Zscaler Cloud DLP, see the [Zscaler Cloud DLP data sheet \(https://www.zscaler.com/resources/data-sheets/zscaler-cloud-dlp.pdf\)](https://www.zscaler.com/resources/data-sheets/zscaler-cloud-dlp.pdf).

Visualize and Report on the State of Your Data



Preventing data exfiltration with Zscaler Cloud DLP requires an understanding of what your data looks like and where it currently resides. In this phase, you enable monitoring of your network using built-in dictionaries and tools for classification. This helps you identify any PII that is moving through your systems.

In this phase, you define workflows for DLP violations. This includes configuring an incident receiver to receive records of DLP violations. You can choose to leverage either an internet content adaptation protocol (ICAP) as your receiver, or a Zscaler Incident Receiver. The Zscaler Incident Receiver lists violations and their details for use in your existing workflows. The information contained in the report includes:

- The device IP and username via ICAP X-headers.
- A copy of the HTTP POST request that contains the file that violated the DLP policy, or if the content is from HTTP form data, a copy of the content that violated the DLP policy. The host URL to which the user was attempting to send content is also included.

Next, you'll build out policy to monitor traffic using default engines by deploying the Zscaler Indexing Tool to build signatures for your existing files that can be matched against files being transferred. When complete, you'll monitor your logs for violations for approximately one week to gain an understanding of a false positive and an actual threat.

Add Zscaler Cloud DLP to Your Workflows for Violations

With Zscaler Cloud DLP enabled, you should be prepared to receive and act on violations that occur. Integrating reporting and auditing of violations into your workflow likely already exists from a procedural standpoint if you've enabled TLS/SSL inspection. With DLP, you are provided extensive logging and reporting within the Zscaler Central Authority interface.

Many organizations' workflows are centered around a centralized incident receiver. This might be an ICAP server that is already in place, such as a legacy on-premises DLP server. If you do not already have an on-premises sever, you can deploy the Zscaler Incident Receiver to receive reports. Either of these tool sets allows you to examine and record violations, but they are not required. You'll select your incident receivers as part of your policy definition.



The Zscaler Incident Receiver does not store DLP files locally and requires that you provide a storage location for violation information to reside.

To learn more about configuring ICAP incident receivers, see:

- [About ICAP Communication Between Zscaler and DLP Servers](https://help.zscaler.com/zia/about-icap-communication-between-zscaler-and-dlp-servers) (<https://help.zscaler.com/zia/about-icap-communication-between-zscaler-and-dlp-servers>)
- [About Zscaler Incident Receiver](https://help.zscaler.com/zia/about-zscaler-incident-receiver) (<https://help.zscaler.com/zia/about-zscaler-incident-receiver>)

In addition to automated receivers, you can also receive notifications via email to an auditor's address. The receiver can either be hosted or external to the Zscaler system. A hosted account is an admin who can log into the Zscaler Central Authority. An external account is an email address to receive the notifications.

Email notifications are controlled by a DLP Notification Template. The templates are made up of macros that are substituted for violation data when sent. This allows you to build up emails by configuring items such as subject lines for easy filtering and the ordered-by content. You can modify the templates to include items such as the URL to meet your auditor's needs.

You'll create separate templates for inline and out-of-band DLP violations, as each offers different macros based on their function. To view a full list of macros and see example email templates, see [Configuring DLP Notification Templates](https://help.zscaler.com/zia/configuring-dlp-notification-templates) (<https://help.zscaler.com/zia/configuring-dlp-notification-templates>).

Enable DLP in Monitor-Only Mode

The initial setup of Zscaler Cloud DLP is in monitor-only mode. In this mode, you log violations of the built-in DLP dictionary matches that occur. This does not change how your cloud storage and applications are used by your end users. Violations are logged and leveraged to help set up violation workflows, as well as map out your data storage locations and users.

A DLP dictionary is a set of patented algorithms and rule sets that are run against a data set to find matches. A DLP engine is a grouping of one or more DLP dictionaries and is used in policy definition. Together, dictionaries and engines allow you to build up a set of rules that match common PII or your organization's private information.

Using built-in DLP dictionaries enable you to get a handle on common PII that might be in use in the system. In the later phases, you'll modify or construct custom DLP dictionaries that match your specific organizational data. For now, you'll want to review the existing DLP engines against the kind of data that is likely to exist in your network. For instance, if your organization sells directly to customers, you might have financial data on file such as customer credit card information. In this case, you would want to ensure your DLP engine is leveraging the Credit Cards dictionary. For a complete list of available DLP dictionaries, see [Predefined DLP Dictionaries](https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries#predefined-dlp-dictionaries) (<https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries#predefined-dlp-dictionaries>).

Each predefined dictionary contains a confidence score threshold, which is used to define how many violations are required to trigger a match. This metric is a method of helping eliminate false positives by increasing the score required to trigger the match. The more violations matched in the definition, the higher the score. You can select from low, medium, or high threshold, each of which adds additional required violation criteria to the dictionary to trigger a match. If you see false positives too often, you can increase the confidence threshold to require a higher score and compare your results. You can find details on what is required to match at each confidence threshold by expanding each dictionary name at [Predefined DLP Dictionaries](https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries#predefined-dlp-dictionaries) (<https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries#predefined-dlp-dictionaries>).

You should look through the descriptions of the libraries you plan to use to understand what the various confidence thresholds require to trigger a match. With this information, you should modify the built-in dictionaries as needed to match your desired outcomes. You might start with low, ensuring the most matches, and then increase your level to reduce false positives. You can learn how to modify confidence levels at [Editing Predefined DLP Dictionaries \(https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries#predefined-dlp-dictionaries\)](https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries#predefined-dlp-dictionaries).

The policies you configure are evaluated in ascending numerical order, starting with 1 and continuing until a match is made or all rules have been exhausted. Unlike a firewall, there is no default “deny all” policy at the end. If the DLP engine doesn't find a match, the data is not considered confidential and is forwarded assuming other ZIA services don't intervene. A policies order in the list is determined by your administrators, and because order matters, care should be taken when building policy.

When a policy matches, the system stops looking for additional matches further down the rule set. However in an allow scenario, you might find that an allow policy match doesn't catch everything. This is especially true where certain users are expected to handle certain PII, such as your finance team uploading credit card information to a payment application. However if you Match Only on what is allowed (credit cards), you won't catch a file that contains both credit cards and social security numbers. To check against your entire rule set, enable Match Only on your allow policies.

If your finance user sends a file with both credit cards (allowed) and social security numbers (blocked), you want to ensure that the file is blocked. With Match Only enabled, the file matches the rule allowing credit cards, and then continues down the policy until it encounters the deny social security number rule, at which point that triggers the deny action and the transaction is blocked. Zscaler recommends enabling Match Only on allow rules. To learn more about configuring Match Only, see [DLP Policy Configuration Example: Match Only \(https://help.zscaler.com/zia/dlp-policy-configuration-example-match-only\)](https://help.zscaler.com/zia/dlp-policy-configuration-example-match-only).

Your initial policy can be as simple as allowing all traffic and using all engines. Note however that this has the potential to generate many logs, especially in organizations handling PII information directly, such as those in the medical or financial fields. In a very large organization, or one handling PII regularly, it is recommended to onboard a department or location first to see the impact on your logging systems.

Configuring individual policy upfront has the potential to save time if you know you plan to add controls by group or category. For example, you can build a policy to monitor file sharing sites for all users, knowing that later you plan to restrict access to approved sites and possibly approved users. In the short term, it allows you to see what other sites are in use and by who so that you can have a discussion with those groups ahead of enforcement.

When building a policy, you have the option of having content inspection on your traffic, or simply matching and forwarding the traffic on to a third-party DLP. While the policy is essentially the same, in the second case you need to specify an ICAP server to receive the forwarded information. Note that Zscaler Cloud DLP does not respond to messages from third-party ICAP servers. To view a video comparing these two options, see [Configuring DLP Policy Rules with Content Inspection \(https://help.zscaler.com/zia/configuring-dlp-policy-rules-content-inspection\)](https://help.zscaler.com/zia/configuring-dlp-policy-rules-content-inspection).

When configuring a policy, you can choose to limit your inspection or expand it to all traffic. Zscaler recommends scanning 100% of your traffic where possible. The following options are available for you to select from to limit or expand the scope of your inspection:

- Which engines are applied – You can select one or more engines, or all the configured and enabled engines. Limiting engines is useful when you need to allow certain PII from certain groups or users.
- URL categories – Include all categories or select specific categories including custom categories.
- Cloud applications – Include all applications or specific cloud applications.
- Outbound data types – Search all file types or only specific types of files, such as PDF for images.
- Users – Up to four individual users or all users, with one of two states:
 - Include – Apply the rule on up to 4 selected users, or all users.
 - Exclude – Apply to all users except those selected, up to 256 users.
- Groups – Groups of users can be selected, and users can be a member of multiple groups. The groups control has one of two states:
 - Include – Apply the rule on up to 8 selected groups, or all groups.
 - Exclude – Apply to all groups except those selected, up to 256 groups.
- Departments – Departments of users can be selected, and users can be a member of only one department. The departments control has one of two states:
 - Include – Apply the rule on up to 8 selected departments, or all departments.
 - Exclude – Apply to all departments except those selected, up to 256 departments.
- Locations – Apply the policy on up to 8 specific locations, or all locations.
- Location groups – Apply the policy on up to 32 location groups, or all location groups.
- Time intervals – Apply all the time or only during a time interval. Up to two intervals are supported per policy.
- Protocols – You can apply your policy to HTTP, HTTPS, native FTP, or all protocols.
- OCR – Enable optical character recognition. Contact your Zscaler Account team for more information.

In addition to scope, there are a few other decisions to make:

- Zscaler Incident Receiver or ICAP server (optional) – As previously mentioned, if you have either a Zscaler Incident Receiver or an ICAP-capable server, the violation information can be forwarded to that device.
- Configure email notification (optional) – Select a hosted account to act as the auditor such as an admin, or an external account and provide the email address for the account.
- Allow or block – If the selectors all match, choose to allow or block the traffic.

Initially, you want to see everything that is going on in the network, so you can leave everything in the “any” category and allow as the rule action. If you already know some of your policy needs, you can start building test policy as well. For instance, if you plan to allow members of finance to upload credit card information, you can build that allow policy now. Monitoring this ensures matching the correct users and destinations before you begin enforcement.

For more information on building out policy, see [Configuring DLP Policy Rules with Content Inspection \(https://help.zscaler.com/zia/configuring-dlp-policy-rules-content-inspection\)](https://help.zscaler.com/zia/configuring-dlp-policy-rules-content-inspection) and [Configuring DLP Policy Rules without Content Inspection \(https://help.zscaler.com/zia/configuring-dlp-policy-rules-without-content-inspection\)](https://help.zscaler.com/zia/configuring-dlp-policy-rules-without-content-inspection).

Monitor for One Week

After you have enabled your Zscaler Cloud DLP policies in allow mode, you want to monitor your logs and monitoring tools for incident matches. You then compare the output to your list of approved applications and users to understand where the gaps are in your controls. The information you gather is often organization specific. There are several common items to identify as you baseline your network, including:

- Top data violations by user
- Top data violations by destination
- Top data violations by file type
- Top data violations by file category
- Where legitimate data is going
- Who legitimately is moving that data

These common violations can drive conversations with the affected users and groups. In some cases, it might be a training issue or a process you don't yet fully understand. It could be that some users are engaged in shadow IT to avoid organizational controls or a process that is too cumbersome. In others, it might be a legitimate bad actor. With the reporting information, you are equipped with the tools to discover the source of the violation and prepare to enforce policy to put a halt to it.

Enable Controls Over Data and Refine Your Dictionaries



In this phase, you begin to refine your matches and enforce blocking policies. The goal of refining your matches is to work to eliminate false positives. These matches occur when the criterion for a match is overly broad or is being applied to incorrect users or groups. This in turn reduces the number of false-positive alerts, allowing your teams to focus on real issues.

To do this, you index your documents and begin building custom dictionaries using phrases and keywords that appear in your proprietary data, such as product code names. Unlike the generic dictionaries used to this point, you now start to specifically match your specific data versus generally available PII. However, you won't dispose of the built-in dictionaries. These are still extremely useful for matching things like credit cards or government ID numbers.

Next, you'll begin blocking unauthorized traffic. You'll begin to break out allows and denies based on the user's role in your organization. There will also be some level of optimization in the blocks as you look at reduce your false-positive matches.

Creating Custom Dictionaries

Each organization has terms and phrases that are specific to that organization. By creating custom dictionaries, you can match these terms and strings within your data flows and reduce your false-positive hits. Your custom dictionaries are added to DLP engines, just like one of the built-in dictionaries. You can create up to 160 custom DLP dictionaries.

There are four types of dictionaries you can choose to build:

- Exact Data Match (EDM) – This template indexes a CSV file of your organization's data, such as employee IDs. This template is covered in the section [Enable Exact Data Match](#).
- Indexed Document Match – This template is leveraged by the Zscaler Index Tool to create templates of your critical docs for matching or partial matching. This template is covered in the section [Create Index Document Match Templates with the Zscaler Index Tool](#).
- Microsoft Information Protection (MIP) – MIP provides sensitivity labels, which you can use to identify and protect files with sensitive content. MIP integration is beyond the scope of this guide. You can find more information on MIP at [About Microsoft Information Protection Labels \(https://help.zscaler.com/zia/about-microsoft-information-protection-labels\)](https://help.zscaler.com/zia/about-microsoft-information-protection-labels).
- Patterns & Phrases – You can specify patterns as regular expressions (regex) or quoted exact phrases that you specify.

Using Phrases to Build Custom Dictionaries

Building a custom dictionary using phrases allows you to build out lists of keywords and strings that match your internal data. This could be strings that include project code names that appear regularly in confidential reports. Phrases are a set of words that must appear together to trigger a match. Each dictionary can have up to 256 phrases, and each phrase can contain up to 128 characters.

Words in phrases are matched in any order unless they are contained within a set of double quotes. Double-quoted phrases must match in the exact order. If you quote a set of terms in double quotes such as:

```
"security service"
```

The system only matches if those words are seen in that exact order. If the system encountered the string:

```
service security
```

There would not be a match. If the quotes were removed, the phrase would match both security service and service security. The system is also aware of non-words in matching patterns. It ignores punctuation, HTML tags, and other display-related information when matching. The system also ignores case when matching, so Security Service would match both of our previous example patterns.

When your phrase is defined, you also must specify an action to go with the phrase, and each phrase will have its own selection. This drop-down menu has one of two options:

- **Count All** – The dictionary counts all matches of the pattern, including identical patterns, toward the match count. For example, you create a dictionary with a pattern for US phone numbers and choose Count All. When the dictionary scans content containing three instances of the same exact US phone number, it counts all three instances as three matches.
- **Count Unique** – The dictionary counts each unique match of the pattern toward the match count only once, regardless of how many times it appears in the content. For example, you create a dictionary with a pattern for US phone numbers and choose Count Unique. When the dictionary scans content containing three instances of the same exact US phone number, it counts all three instances as one match.

To learn more about creating DLP phrase matching dictionaries, see [Defining Phrases for Custom DLP Dictionaries \(https://help.zscaler.com/zia/defining-phrases-custom-dictionaries\)](https://help.zscaler.com/zia/defining-phrases-custom-dictionaries).

Create Index Document Match Templates with the Zscaler Index Tool

The Zscaler Index Tool is a virtual machine that creates templates of your documents or data for inspection. Initially, you'll focus on indexing your documents. This creates a fingerprint the system leverages when looking for data leaks. The DLP engine can then use those fingerprints to look for full or partial document matches while your data is in transit. In the third phase, you'll use this same tool to create templates of your data itself for exact data match.

The output of the indexing tool is a custom DLP library. This can be added to a new or existing engine for enforcement. As your policy executes, the data or documents being inspected are compared against your fingerprints for a full or partial match. For example, you might index forms that are routinely filled out by your employees. This blank form triggers a partial match whenever a completed form is sent, because the Zscaler Cloud DLP engine knows the format of the original form. You might also index items such as internal announcements or memos to prevent exfiltration of sensitive internal data.

- Learn more at [About the DLP Index Tool \(https://help.zscaler.com/zia/about-index-tool\)](https://help.zscaler.com/zia/about-index-tool).
- Learn more at [About Indexed Document Match \(https://help.zscaler.com/zia/about-indexed-document-match\)](https://help.zscaler.com/zia/about-indexed-document-match).

Building Blocking Policies for Users

At this stage, you're ready to begin enforcing policy. Because this directly impacts your user's ability to access resources, it's recommended to start with specific user groups that need access to a set of well-defined applications. This group could be someone like your finance or development teams.

First, begin by identifying which groups have a legitimate need to move PII or organizational IP to cloud applications. This should list specifically:

- What kind of information is moving – Use the provided information to build custom dictionaries that match this data, used for both allow and deny policies.
- Where the data is moving to – List what specific applications, storage, or other services are approved to host this information.
- Who is moving the data – Which group(s) of users should have the ability to move the data.

Using department and group membership is an ideal combination to strictly limit which users can interact with PII in the cloud. Finance users for instance likely handle monetary transaction reports, paying vendors, credit card numbers, and purchase orders. But in many large organizations, these are specialized groups within finance. Your policy can allow a certain group under finance to transfer this information to specific cloud applications or storage, and then you can create a second policy to block other users from doing the same.

Continue this pattern of building in allow policies and a corresponding block policy for the rest of your organization as needed.

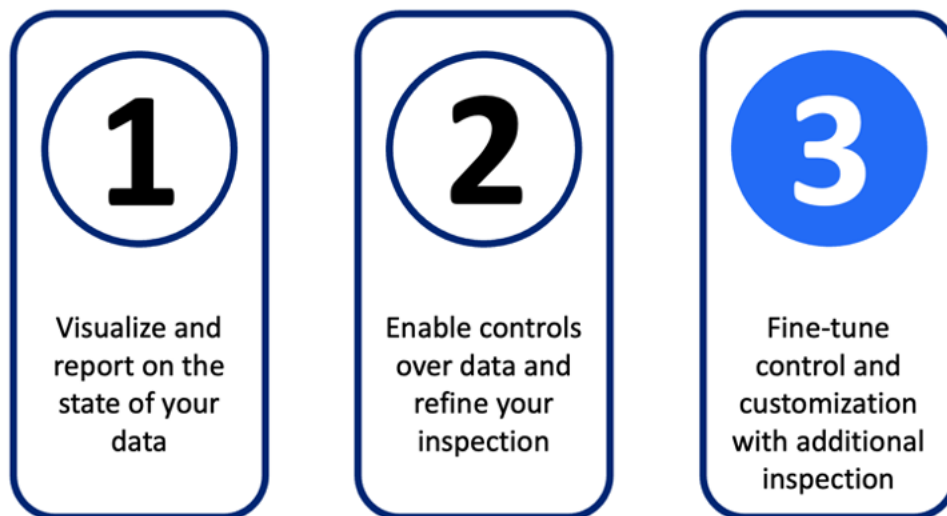
Monitor for One Week

As in phase 1, you'll want to monitor your logs and monitoring tools for incident matches, especially for blocked actions. The list of common items to monitor as you begin enforcement should include:

- Top data violations by user
- Top data violations by destination
- Top data violations by file type
- Top data violations by file category
- Where legitimate data is going
- Who is legitimately moving that data

You'll still be monitoring for gaps in your controls, as well as legitimate users that are being accidentally blocked.

Fine-Tuning Control and Additional Inspection Tools



In the final stage, you continue to refine your policy and engines to reduce false positives and ensure compliance with reporting and alerting. Monitoring traffic and reporting continues. Examining false positives to create and refine dictionaries should become a part of your operating process. You might also encounter cloud destinations that are exempt from inspection. You'll learn about exception policies for bypassing DLP inspection. With Exact Data Match, you can fingerprint your organization's data so it can be found when sent in different formats. Using optical character recognition (OCR) on images can stop data exfiltration using images such as screen capture tools.



Some features described in this section might require additional subscriptions. For a complete overview of what is included in the different ZIA subscription levels, see [Pricing and Plans](https://www.zscaler.com/pricing-and-plans) (<https://www.zscaler.com/pricing-and-plans>), or contact your Zscaler Account team.

Using Patterns to Build Custom Dictionaries

Building a custom dictionary using regular expressions (regex) patterns allows you to create patterns that specifically match your organization's data. This can be something like an employee ID, or perhaps a client ID that is used in your internal communication. It can be a pattern matching your hardware serial numbers, or a driver's license pattern for a country where your employees operate. These patterns are typically unique to an organization or country and require customization to match effectively.

Much like the built-in pattern matching for credit cards or government ID numbers, regex matches work to create patterns that expand to match data. With these patterns, syntax and order matter a great deal, so it's advisable to seek help from experienced programmers with a regex background. The patterns you create are meant to match patterns you expect to find in your data. These can be patterns of numbers, characters, or both.

A single dictionary can contain up to 8 patterns. This is useful because it takes more than a single pattern to match all the variations in a data set in some cases. For example, when a country's passport numbering system changes the pattern of numbering, but older passports are still valid, you need to match both the new pattern and the old one. These patterns are placeholders for both upper and lower characters, digits, and repeated patterns of numbers. For example, a Taiwan ID number would be matched by the following pattern:

```
\b[a-zA-Z][12]\d{8}\b
```

This regular expression breaks down as:

- The shorthand class, `\b`, matches a word boundary, the starting or ending of a word.
- The first square brackets match any one letter from a to z or A to Z.
- This is followed by either a 1 or a 2 in the second square bracket.
- The class `\d` matches any single digit from 0 to 9, and the pattern expects 8 digits in a row due to the repeats specified with curly braces `{8}`.
- The final word boundary `\b` that ends our string.

This powerful string allowed us to match all ID numbers for Taiwan without having to specify exactly what number we were looking for. When your pattern is defined, you also must specify an action to go with the pattern, and each pattern will have its own selection. This drop-down menu has one of two options:

- **Count All** – The dictionary counts all matches of the pattern, including identical patterns, toward the match count. For example, you create a dictionary with a pattern for US phone numbers and choose Count All. When the dictionary scans content containing three instances of the same exact US phone number, it counts all three instances as three matches.
- **Count Unique** – The dictionary counts each unique match of the pattern toward the match count only once, regardless of how many times it appears in the content. For example, you create a dictionary with a pattern for US phone numbers and choose Count Unique. When the dictionary scans content containing three instances of the same exact US phone number, it counts all three instances as one match.

To learn more about creating regex patterns, see [Defining Patterns for Custom DLP Dictionaries](https://help.zscaler.com/zia/defining-patterns-custom-dictionaries) (<https://help.zscaler.com/zia/defining-patterns-custom-dictionaries>).

Increase Match Count in Built-In and Custom Engines

In certain inspection engines, a match count is required to determine how many instances of a violation must be found to trigger a match. You'll want to adjust your expressions match counts where applicable, typically increasing the counts required. This has two effects on your data:

1. This policy will be tuned to ensure that sites such as ecommerce aren't blocked by requiring high match counts. This won't stop a single user's worth of information and credit card data to a retailer but would stop a mass exfiltration of data.
2. You will greatly reduce the number of false positives due to a single piece of PII included in a communication.

It is recommended to typically match against a high number of violations. Zscaler recommends 25+ matches for most engine expressions. If you are unsure that this is the right level for your organization, instead begin by increasing your match count in small increments each week and monitoring the effect on your reported violations and false-positive rate. You'll eventually reach a point where you feel that you are catching actual violations and not generating high rates of false-positive alerts.

You can view the steps for configuring match count at [Editing Predefined DLP Engines \(https://help.zscaler.com/zia/editing-predefined-dlp-engines\)](https://help.zscaler.com/zia/editing-predefined-dlp-engines).

Increase Confidence Score Threshold in Dictionaries

Each predefined DLP dictionary has a confidence score threshold assigned to it. The confidence score threshold has three states: Low, Medium, and High. When a threshold is set to Low, the dictionary is more aggressive in its matching and alerting of violations. As you increase the threshold, the number and types of matches required to trigger a match increase. If you find that a particular dictionary is reporting too many violations, you can increase the threshold to reduce matching.

For a complete listing of built-in dictionaries and what each confidence threshold requires to match, see [Predefined DLP Dictionaries \(https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries#predefined-dlp-dictionaries\)](https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries#predefined-dlp-dictionaries).

Add Phrases to Built-In and Custom Dictionaries

When you set your confidence score to High, the Custom High Confidence Phrases field appears and you can add additional phrase matching to the existing dictionary. This can be used if you need additional matching to reduce false-positive matches. This refinement adds your custom phrases to the dictionary check and works with both phrase and pattern-matching dictionaries.

To learn more about adding phrases to an existing dictionary, see [Editing Predefined DLP Dictionaries \(https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries\)](https://help.zscaler.com/zia/editing-predefined-dlp-dictionaries).



Phrases have limitations for the number of characters and sensitivity. Before adding custom phrases, see [Using Phrases to Build Custom Dictionaries](#) in this guide for relevant details and guidance.

Refining Block Policies

At this stage, your block policies should be tuned to find both common PII and your organization's data as it transits the network. Your initial groups of users who had specific needs to handle PII should be able to function as expected. Your next step is to block policies for all other users, cloud storage, and cloud applications.

Using your universal reporting rule developed in phase 1, you can find out which departments and users are likely to be impacted. Work with these users to understand why they are triggering violations and if this is appropriate to their role and function. If these users are performing roles that involve this type of data, you can approve an exception for that user or group.

Much like your previously identified groups, you can build an allow policy and place it above any block policies in your DLP policy list. Whenever possible, leverage groups over lists of individual users for ease of management.

When all legitimate exceptions are accounted for, you can move on to building a robust block policy. You should now have your dictionaries adjusted to increase match counts. Check that you are running multiple dictionaries in one or more engines to ensure that you are matching widely. For your global block, Zscaler recommends a rule to match against all cloud destinations and applications. This rule must be set below any allow rules for data exceptions, otherwise you will block your users with legitimate needs. When this policy is enabled, you should closely monitor your help queue and block notifications to your audit team.

Scanning Images with Optical Character Recognition

Enabling optical character recognition (OCR) checks against image files allows you to discover data that might be hidden in something like a screen capture. By enabling this checkbox in policy, the selected DLP engines are run against images as well as other files. Check with your Zscaler team to see if OCR is enabled for your organization.

Exempt Cloud Applications and URLs from DLP Evaluation

If you need to have certain cloud applications, URLs, URL categories, or URL encoded data exempted from inspection, you can specify things in the advanced DLP section. When you exempt cloud apps or URLs from inspection, you remove them from all inspection. The exemption list overrides any rules configured for inspection. Zscaler recommends you only exempt apps where there is a regulatory requirement or legal agreement requiring inspection bypass.

To learn more, see [Exempting Cloud Apps and URLs from DLP Evaluation \(https://help.zscaler.com/zia/exempting-cloud-apps-and-urls-dlp-evaluation\)](https://help.zscaler.com/zia/exempting-cloud-apps-and-urls-dlp-evaluation).

Enable Exact Data Match

The exact data match (EDM) feature goes beyond pattern matching with regular expressions, and instead allows you to fingerprint your organization's PII. These fingerprints are then leveraged by DLP rules to find your data in forms or applications that differ from your internal documents or formats. Zscaler recommends considering the following before creating an EDM index template:

- As you review the DLP policy you want to create, consider the data that must be included in your EDM index template.
- Try to create a template where your data records need to be indexed once, and avoid the need to re-index where possible.
- Review your data records to avoid potential duplication.

The fingerprinting takes in structured data uploaded as a CSV file to the Zscaler Index Tool. This will be the same VM you deployed previously to fingerprint documents. Here, you identify the data columns by the type of data they contain. The tool shows you the field name based on the title row of the CSV and some sample data from the file for each column.

The data must contain at least one and up to two primary fields that are unique in the data. These can be fields such as your user's government ID number and their organization-issued credit card number. In addition to the primary fields, you can include additional identifiers as secondary fields. Each field type must have a data type selected that indicates the type of data contained in the field, such as credit card numbers. You must select your field to be included in the template. If there is data that is not relevant for consideration in the document, deselect the "Include Field in Template" checkbox for that field.

When you've made your selection, you can index your data and build your template. The template and its fingerprinted data are then available for use to create a new dictionary. When creating a new dictionary, you select "Exact Data Match" as your dictionary type. You'll select your template, and from the drop-down menu select:

- Primary Field – Select a primary field from those specified in the template.
- Secondary Fields – Select any of the secondary fields from fields in the template.
- Secondary Match On – Select how you want to match secondary fields.

Secondary fields can be included either as a logical AND (default) or a logical OR along with the primary fields. You can also specify a match on count, such as much match 1 primary and 2 secondary matches to constitute a violation. If you have two primary fields, they are checked independently in separate rules with the secondary fields you have selected. You can also write a policy to check both primary fields together with secondary fields.

When saved, the dictionary is now available to be included in your DLP engines. Your policy will match exactly against your real user data as opposed to data that matches the pattern. You also need to keep your data synchronized on a regular basis.

EDM templates can be scheduled to update on a regular basis by reading a shared file. This allows you to automatically export your PII data and have it read and updated by the Zscaler Index Tool. Zscaler recommends doing automated scheduled updates so that your data is kept in sync.

To learn more about creating EDM templates, see [Creating an Exact Data Match Template \(https://help.zscaler.com/zia/creating-exact-data-match-template\)](https://help.zscaler.com/zia/creating-exact-data-match-template).

To learn more about EDM configuration, see [About Exact Data Match \(https://help.zscaler.com/zia/about-exact-data-match\)](https://help.zscaler.com/zia/about-exact-data-match).

Out-of-Band CASB

Zscaler's out-of-band CASB uses multiple technologies to scan your data at rest in SaaS applications tenants. This tenant is then scanned by [Zscaler Cloud DLP](#) and Zscaler Malware Detection via API to your SaaS applications. This scanning leverages the tuned DLP dictionaries and engines you create for inline DLP against the data in these applications. This is combined with Malware Detection to look for trojan files that might have been uploaded to your SaaS applications outside of ZIA inspection, such as through a partner or customer upload.

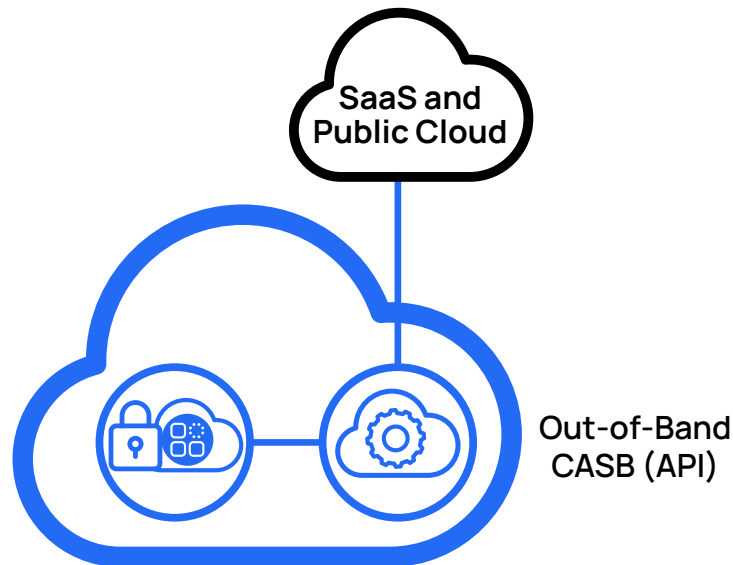


Figure 10. ZTE scans your SaaS application tenants for DLP violations and malicious content

The Zero Trust Exchange is configured to reach out to your SaaS application tenants to scan data stored in those applications. If a violation is found, you have multiple ways to be alerted about the violation. Your policy can also include automated actions that occur when a violation of policy is detected, allowing you to immediately take action to safeguard a threat as soon as it is detected.



Out-of-band CASB might require additional subscriptions to enable all features and applications. For more information on subscriptions, see [Zscaler Data Protection at a Glance \(https://www.zscaler.com/resources/data-sheets/zscaler-data-protection-benefits.pdf\)](https://www.zscaler.com/resources/data-sheets/zscaler-data-protection-benefits.pdf), or contact your Zscaler Account team.

Incident Receiver

Zscaler DLP policy supports sending violations to an incident receiver. This can be any ICAP-capable device, such as a legacy on-premises DLP server or the Zscaler Incident Receiver virtual machine. For information regarding DLP incident receivers, see [Add Zscaler Cloud DLP to Your Workflows for Violations](#) in this guide.

SaaS Application Tenants

SaaS applications that your organization sanctions and builds a scan policy for are called SaaS application tenants. To scan the files and content on your SaaS applications, you must first add that application as a tenant. When configured and authorized, you can use your tenants in policy for DLP and Malware scanning. Zscaler supports the following list of supported SaaS application tenants:

- Amazon S3
- Box
- Citrix ShareFile
- Dropbox
- GitHub
- Gmail
- Google Cloud
- Google Drive
- Google Workspace
- Google Workspace Marketplace
- Microsoft 365
- Microsoft Azure
- Microsoft Exchange
- Microsoft OneDrive
- Microsoft SharePoint
- Microsoft Teams
- Salesforce
- ServiceNow
- Slack
- Webex Teams

Based on this list, you can determine which SaaS applications your organization has in use and that you want to scan. To view configuration steps for each tenant application, see [Adding SaaS Application Tenants \(https://help.zscaler.com/zia/adding-saas-application-tenants\)](https://help.zscaler.com/zia/adding-saas-application-tenants).



Limits exist on the number of SaaS applications you can scan and might require additional subscriptions to scan all your organization's applications. For more information on subscriptions, see [Zscaler Data Protection at a Glance \(https://www.zscaler.com/resources/data-sheets/zscaler-data-protection-benefits.pdf\)](https://www.zscaler.com/resources/data-sheets/zscaler-data-protection-benefits.pdf), or contact your Zscaler Account team.



To enable Amazon S3, Google Cloud Platform, and Microsoft Azure for your organization, contact your Zscaler Account team.

DLP Policy

Building a policy for SaaS Security API Data Loss Prevention (DLP) scanning is built around application categories. These categories group together applications with the same or similar functions, such as collaboration. Your policy begins by selecting the application category rule. Each application category has controls customized to the application's function. For example, an application in the collaboration tool category has controls for messages and file attachments. This is different than a source code repository category application with file types and collaborators.

During the configuration of a DLP policy rule, the set of available controls will be different based on the application category. When you configure your policy, you select the SaaS application tenant that the rule applies to. This is limited to applications that are in the same application category as the policy. The categories of applications are:

- Collaboration
- CRM (customer relationship management)
- Email
- File Sharing
- ITSM (IT service management)
- Public Cloud Storage
- Source Code Repository

The controls are appropriate to the application function and allow you to build granular inspection policy. As you build out your policy, consider that data leaks can occur anywhere. Running all your DLP engines against all content types is most effective at detecting data leaks.

The other controls are focused on what order the rules should be evaluated in, to which users, and to what content the rules should be applied. There are variations in terms, such as author and sender in place of user, depending on the application. Together these rules allow you to scan as broadly or as narrowly as required for auditing and remediation purposes.

The rules attributes include:

- Rule Order – The numerical order of the rules is evaluated in numerical order from lowest to highest.
- Admin Rank (optional) – A control that enables you to prevent rules from being reordered or overwritten by admins with a lower organizational rank. To learn more about admin rank, see [About Admin Rank \(https://help.zscaler.com/zia/about-admin-rank\)](https://help.zscaler.com/zia/about-admin-rank).
- Rule Name – A name to help describe the purpose of the rule.
- Rule Status – Either enabled or disabled.
- Rule Label (optional) – Allows you to logically group rules to look for common policy enforcement. To learn more about labels, see [About Rule Labels \(https://help.zscaler.com/zia/about-rule-labels\)](https://help.zscaler.com/zia/about-rule-labels).

The criteria by which rules are applied is where you find the variation in controls by application category.

- SaaS application tenant – Select the SaaS application tenants that the rule will be applied to.
- Users/Senders/Authors – Up to 4 individual users or all users.
- Groups – Groups of users can be selected, and users can be a member of multiple groups. Apply the rule on up to 8 selected groups, or all groups.
- Departments – Departments of users can be selected, and users can be a member of only one department. Apply the rule on up to 8 selected departments, or all departments.
- DLP Engines – Select Any to apply all engines, or select up to 4 engines.
- Category specific criteria – To view specific controls by application category, see [Configuring the SaaS Security API DLP Policy](https://help.zscaler.com/zia/configuring-saas-security-api-dlp-policy) (<https://help.zscaler.com/zia/configuring-saas-security-api-dlp-policy>).

After the criteria are in place, you can define the actions taken by policy when a match occurs. The actions available are dependent on which application category the policy is based on. Depending on your SaaS application tenant, you might have additional functionality available. The kinds of functionality are dependent on the application purpose, such as collaboration or file sharing. For example, in a collaboration application you can take actions to report the violation or warn the user through a bot in their messaging application. With a file share, you might restrict or remove access to a file.

In all cases, you can also select a severity level that will be associated with this incident. Properly setting incident levels between low, medium, and high allows you to move quickly on major violations while monitoring lower priority issues.

You are also provided two options for incident reporting that are consistent across all the application categories. If you've configured a DLP incident receiver, you can select it here. You also have the option to configure email notification for the rule. This can be sent to a Zscaler admin or to any email address you choose.

After you've completed your policy configuration, your next step is to schedule it to run against your SaaS application tenants. Scheduling is discussed in [SaaS Security API Scan Schedules](#) in this guide.

To view the options for criteria and actions for each category type, see [Configuring the SaaS Security API DLP Policy](https://help.zscaler.com/zia/configuring-saas-security-api-dlp-policy) (<https://help.zscaler.com/zia/configuring-saas-security-api-dlp-policy>).

Learn more at [About Rule Labels](https://help.zscaler.com/zia/about-rule-labels) (<https://help.zscaler.com/zia/about-rule-labels>).

Malware Policy

Detecting malware residing in your SaaS application tenants requires building a policy for each tenant. During the configuration of a malware policy rule, the set of available controls will be different based on the application category. When you configure your policy, you select the SaaS application tenant that the rule applies to. This is limited to applications that are in the same application category as the policy. The categories of applications are:

- Collaboration
- CRM (customer relationship management)
- Email
- File Sharing
- ITSM (IT service management)
- Public Cloud Storage
- Source Code Repository

Unlike many other policies, malware policies have fewer constraints as they are applied to all data and users. Unlike PII and other sensitive data that certain groups of users have a legitimate use for, malware is an immediate danger to everyone in the organization. The various policies provide differentiated controls by application type. Some applications like file sharing or collaboration apps allow robust control via API calls. Others like email provide no control, only alerting that malware was seen.

When building a policy, there are four common components to the policy to select:

- Application – Select an application from the list of supported applications. See [SaaS Application Tenants](#) for more information.
- SaaS Application Tenant – Select the tenant for the application from a list of tenants you've configured.
- Status – Enable or disable the Zscaler service from inspecting data for malware.
- Rule Label – Select a rule label to associate it with the rule.
- Any application specific settings – Some categories have additional requirements for setup.

Under the actions section, the options vary depending on the application. There are only a few main types of action, but not all of these options are available for all application types:

- Remove Malware – Completely delete the malware from the application.
- Quarantine Malware – Move the malware to a quarantined area, such as a chat channel, private folder, or via permissions changes.
- Report Malware – ZTE reports the incident but does not do anything with the file itself.

Zscaler recommends taking the strongest action possible in each policy when malware is detected. Where you can't remove or quarantine the malware such as in email, you should ensure proper escalation of incident reports in those categories. By automatically removing or quarantining your infected files, you rapidly curtail the spread of malicious content amongst your users and clouds.

For more on configuring malware detection policy, see [Configuring the SaaS Security API Malware Detection Policy \(https://help.zscaler.com/zia/configuring-saas-security-api-malware-detection-policy\)](https://help.zscaler.com/zia/configuring-saas-security-api-malware-detection-policy).

SaaS Security API Scan Schedules

Now that you have configured your [DLP Policy](#) and [Malware Policy](#), you must set up scheduled scans of your SaaS application tenants. The term schedule is a bit confusing in that you don't schedule scans to be performed at a particular time. Instead, the scan is scheduled to run whenever the ZTE is notified of new data on the SaaS application tenant.

Your scanning schedule will combine a SaaS application tenant and a policy that can include both DLP and malware, but only one policy is required. The Data to Scan field specifies the amount of historical data versus active data to scan. Active content is any content created, modified, sent, or received after the scan's start date. There are three options for historic data:

- New Data Only – This scans only active data. Historical data remains unscanned unless modified.
- Data Created or Modified After – Select a date to scan as the start date, and scan all active data after that date.
- All Data – Scan all data in your application.

Not all applications support unlimited lookback on their API, so be sure to understand what is possible with your application. To learn more, see [Configuring a Scan to Inspect Historical Data \(https://help.zscaler.com/zia/understanding-saas-security-api-scan-schedules#configure-scan-inspect-historic\)](https://help.zscaler.com/zia/understanding-saas-security-api-scan-schedules#configure-scan-inspect-historic).

Learn more at [About SaaS Security API Scan Configuration \(https://help.zscaler.com/zia/about-saas-security-api-scan-configuration\)](https://help.zscaler.com/zia/about-saas-security-api-scan-configuration).



The amount of historical data scanned is limited by your subscription. For more information on subscriptions, see [Zscaler Data Protection at a Glance \(https://www.zscaler.com/resources/data-sheets/zscaler-data-protection-benefits.pdf\)](https://www.zscaler.com/resources/data-sheets/zscaler-data-protection-benefits.pdf), or contact your Zscaler Account team.

Optionally with file sharing applications, you can specify folders not to scan. To create a folder exception, you need the following information:

- Tenant – From the drop-down menu, choose the SaaS application tenant that the folder belongs to.
- Owner – From the drop-down menu, choose the user who owns the folder.
- Folder – Enter the full path for the folder you want to exempt from inspection.

Exceptions should be handled carefully, as you have no visibility for DLP or malware detection in those folders. To learn more, see [Configuring SaaS Security API Scanning Exceptions \(https://help.zscaler.com/zia/configuring-saas-security-api-scanning-exceptions\)](https://help.zscaler.com/zia/configuring-saas-security-api-scanning-exceptions).

SaaS Security Activities and Alerts

Monitoring SaaS application tenant activity by your users can help alert you to issues that you might want to investigate. ZIA provides four types of alerts that you can duplicate and modify to suit your organization's needs. These alerts appear on the SaaS Security Activity Alerts page, where they can be sorted and filtered.

The four types of alerts that you can configure are:

- **Default Impossible Travel Alert** – This alert identifies users who log in from multiple locations that are impossible to physically travel to in between logins. This alert applies to all tenants, and this activity must occur at least once within a day to generate an alert.
- **Default Failed Logins Alert** – This alert identifies users who have multiple failed logins. This alert applies to all tenants, and the activity must occur at least three times within a day to generate an alert.
- **Default Bulk Upload of Data Alert** – This alert identifies users who upload bulk amounts of data. This alert applies to all tenants, and the activity must occur at least 1,000 times (i.e., a user must upload 1,000 files) within a month to generate an alert.
- **Default Bulk Download of Data Alert** – This alert identifies users who download bulk amounts of data. This alert applies to all tenants, and the activity must occur at least 1,000 times (i.e., a user must download 1,000 files) within a month to generate an alert.

The alerts can be duplicated and modified. For instance, you could specify a different number of files downloaded over a different time frame or build custom alerts for specific tenants. You can also build in exceptions to avoid triggering alerts on legitimate traffic.

To learn more, see [About SaaS Security Activity Alerts \(https://help.zscaler.com/zia/about-saas-security-activity-alerts\)](https://help.zscaler.com/zia/about-saas-security-activity-alerts).

Cloud Browser Isolation

Cloud Browser Isolation is an advanced cybersecurity technique that provides an additional layer of protection by moving web browsing sessions to cloud-hosted remote servers. Cloud Browser Isolation separates browsing activity from endpoint hardware to reduce the attack surface of a user's device. When a user accesses a web page or app, it's loaded onto a remote browser that serves a rendering of the web page to the user. The page operates normally, but only pixels, not the active content, are delivered to the user. Malicious code that might be hidden is kept at bay, and the user experience is unaffected.

Cloud Browser Isolation security uses the same ZIA inspection for all traffic, and then adds an additional layer of separation. Your users are interacting with rendered data, not a real machine. Without additional permissions, they cannot save files or copy data from the session. To learn more, see [About Cloud Browser Isolation \(https://help.zscaler.com/zia/about-cloud-browser-isolation\)](https://help.zscaler.com/zia/about-cloud-browser-isolation).

How It Works

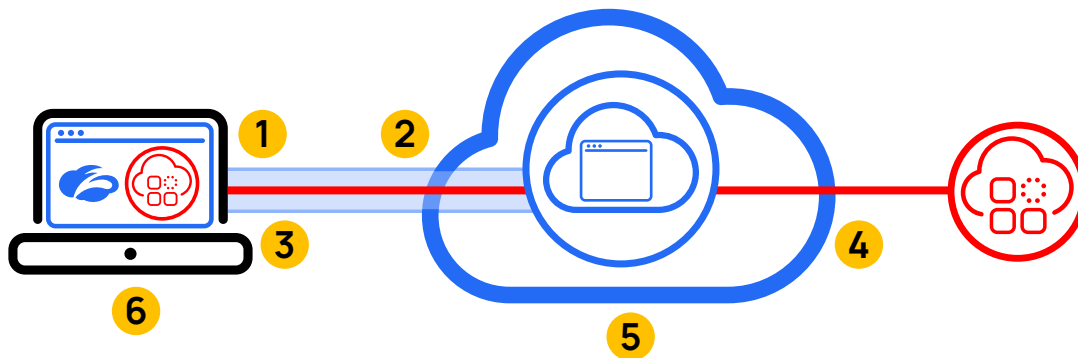


Figure 11. Zscaler Cloud Browser Isolation keeps malicious page content at bay and browsing continues as usual

1. When your request reaches a ZIA Public Service Edge, it is matched against policy. With Cloud Browser Isolation as its action, the traffic is redirected to an isolation profile URL.
2. The user's browser follows the redirect and makes a request to the isolation profile URL.
3. Cloud Browser Isolation accepts the request and assigns a temporary remote browser for the user.
4. The remote browser then makes a connection to the original URL that the user intended to access, and the web page is loaded on the remote browser. This request to the original web page is also routed through the nearest ZIA Public Service Edges, and the traffic is evaluated against all the policies defined for the user on ZIA by the admin.
5. If the policy allows the page to be loaded, it is rendered on the remote browser as a rendered page.
6. The original user sees the page rendered as pixels rather than HTML objects in their native browser.

Use Cases for Remote Browser Isolation

There are four main use cases for Cloud Browser Isolation:

- Zero trust key employee isolation – Easily set up user group isolation policy for high-value targets like your C-level executives, finance, accounting, or HR teams.
- Zero trust data isolation – Prevent data leakage or theft from web-based email, SaaS applications, and internal apps with granular policy control to prevent upload, download, copy, paste, and print actions.
- Zero trust app isolation – Without software installations on endpoints, secure access to SaaS and private apps for third-party unmanaged devices that require access but are difficult to control and constitute risk.
- Zero trust threat isolation – Stop zero-day vulnerabilities, ransomware, drive-by downloads, malvertising, and other sophisticated attacks from reaching end users by isolating web traffic, creating an air gap in front of all active web content.



Cloud Browser Isolation is an add-on subscription. For more information on subscriptions, see [Zscaler Plans and Pricing \(https://www.zscaler.com/pricing-and-plans\)](https://www.zscaler.com/pricing-and-plans) or contact your Zscaler Account team.

Summary

Zscaler's Data Protection suite offers robust protection for your organization's data. Unlike complex legacy approaches that can't follow users, Zscaler Data Protection is a simple but powerful way to secure all cloud data channels. Zscaler protects all users anywhere and controls data in SaaS and public clouds, all with a robust and intuitive data discovery engine.

About Zscaler

Zscaler (NASDAQ: ZS) accelerates digital transformation so customers can be more agile, efficient, resilient, and secure. The Zscaler Zero Trust Exchange protects thousands of customers from cyberattacks and data loss by securely connecting users, devices, and applications in any location. Distributed across more than 150 data centers globally, the SASE-based Zero Trust Exchange is the world's largest inline cloud security platform.

©2023 Zscaler, Inc. All rights reserved. Zscaler, Zero Trust Exchange, Zscaler Private Access, ZPA, Zscaler Internet Access, ZIA, Zscaler Digital Experience, and ZDX are either (i) registered trademarks or service marks or (ii) trademarks or service marks of Zscaler, Inc. in the United States and/or other countries. Any other trademarks are the properties of their respective owners.