Protecting Private Applications with ZPA AppProtection

Reference Architecture — Zscaler for Users

Contents

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**).

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

Terms and Acronyms Used in This Guide

Acronym	Definition
AUP	Acceptable Use Policy
CRM	Customer Relations Management
CVE	Common Vulnerabilities and Exposures
DC	Data Center
ERP	Enterprise Resource Planning
HTTP	Hypertext Transfer Protocol
OS	Operating System
OWASP	Open Worldwide Application Security Project
SQL	Structured Query Language
TCP	Transmission Control Protocol
TLS	Transport Layer Security
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
XSS	Cross-Site Scripting
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.

Zscaler Zero Trust Exchange

Zscaler App Connector

Laptop with Zscaler Client Connector Installed

Hacker Laptop

Data Center

AWS Cloud

AWS Application or Workload

Azure Cloud

Azure Application or Workload

Generic Application or Workload

Data Tunnel

Positive / True Badge

Negative / False Badge

Introduction

In today's cloud-connected world, our applications and the data they contain are under constant attack. These attacks on applications continue to grow in sophistication and number, targeting your highest-value digital assets including databases, file servers, customer relationship management (CRM), and your enterprise resource planning (ERP) systems.

The malicious actors who deploy these attacks cover a wide range of technical ability, from nation-state teams and professional criminal organizations, to "script kiddies" using downloadable exploit tools. These malicious users can also be legitimate users within your organization looking to cause harm, or they can be legitimate users with a compromised machine or credentials.

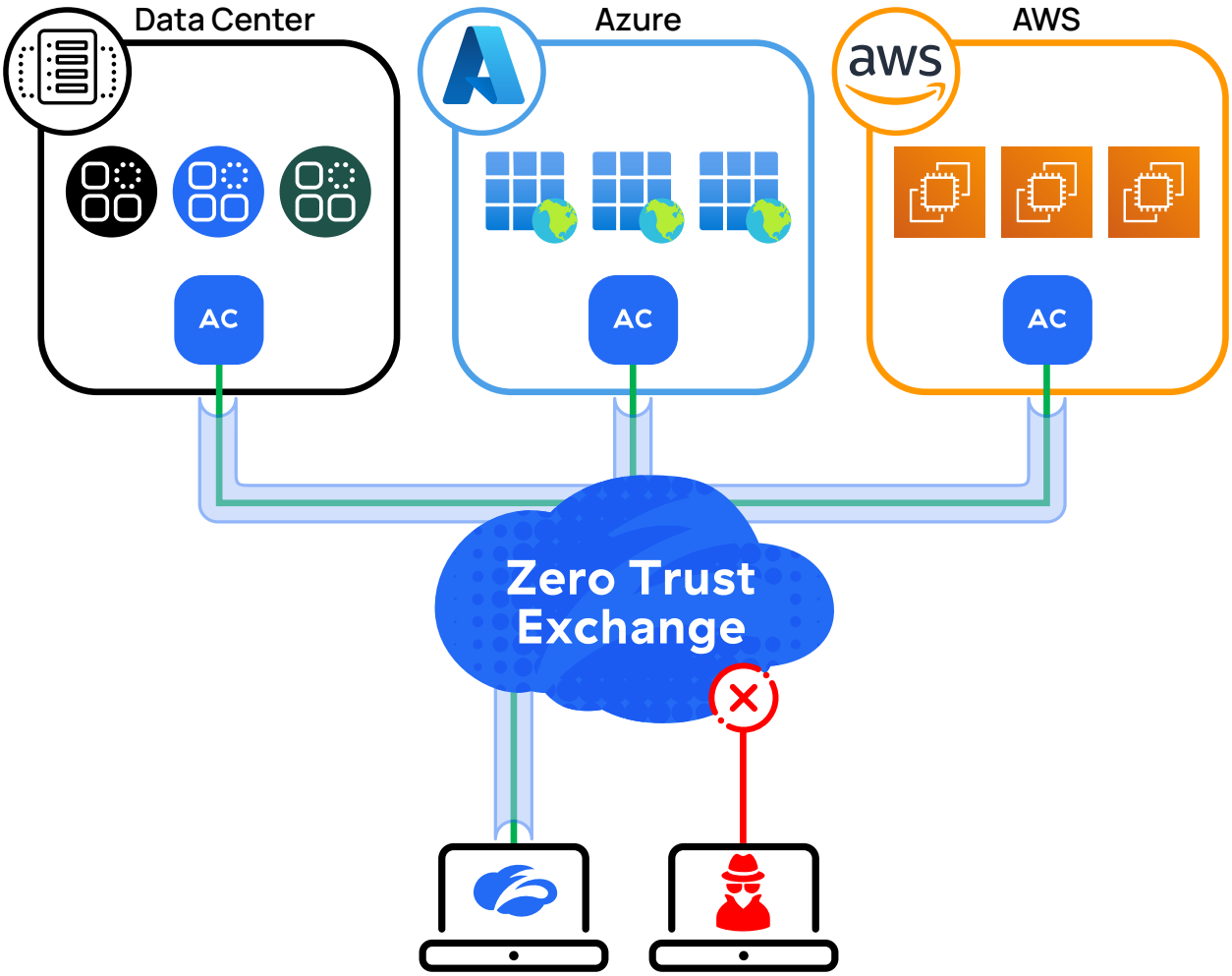

Figure 1. Protect your private applications in your data center, private cloud, or SaaS apps with ZPA AppProtection

Zscaler Private Access (ZPA) protects your private applications from unauthorized access by making them undiscoverable to the internet. Your users securely access only the applications they have approval to use. ZPA provides secure access directly to approved applications, not to the network itself.

Traffic from authorized users is also inspected when accessing private applications. ZPA AppProtection controls enable you to protect your internal application traffic from many types of attacks. AppProtection allows you to do inline inspection of application flows in real time, and to inspect them for malicious content. This includes HTTPS traffic through real-time inspection of encrypted traffic.

AppProtection provides you with built-in controls for the most common attack types. These include the Open Worldwide Application Security Project (OWASP) Top 10 list covering common attacks such as Structured Query Language (SQL) injections and cross-site scripting (XSS). The system also watches for network discovery tools such as port scanners. These controls enable you to quickly eliminate the most common types of attacks on your system.

Zscaler's ThreatLabZ team of security researchers is constantly adding new controls as attacks emerge. This includes Zscaler's security research findings, as well as building controls for recently discovered vulnerabilities that are reported through the Common Vulnerabilities and Exposures (CVE) reporting mechanism.

The ThreatLabZ controls also enable you to automatically subscribe to signature updates from the ThreatLabZ team as they are released. Keeping up with new attacks positions you to protect your applications against zero-day attacks before they become commonplace. This allows you to take advantage of new protections without having to manually configure the system every time an update is made.

WebSocket controls are the third type of built-in control. WebSockets are web API channels that enable two-way communication over a single TCP channel. Controlling and inspecting this traffic ensures that the protocol conforms to the standard and watches for errors in the protocol stream.

Outside of these built-in controls, you might have additional custom controls that must be applied to your applications. Zscaler allows you to build custom controls for both web-based applications and WebSocket applications to meet your deployment needs. By combining built-in and custom controls, you can build out comprehensive security for your internal applications.

When AppProtection is deployed, it monitors browser sessions and alerts you to users with a high number of policy matches. You can also closely monitor your high-risk users to ensure they are not acting in a manner contrary to their roles. This includes additional policy measures and fingerprinting of the user's browser and device. User monitoring gives you detailed visibility into user activities through detailed diagnostics and dashboards that can detect signs of compromise.

Key Features and Benefits

- Powerful built-in application protection identifies potentially malicious user traffic that is destined for private applications. AppProtection also prevents attempts to exploit vulnerabilities or abuse application logic. Admins can tailor protection against any threat or vulnerability, or implement business-specific security policies with our easily customizable rulesets.
- OWASP Top 10 prevention provides comprehensive coverage for the most common types of attacks used by cybercriminals—including SQL injection, XSS, environment and port scanners, and cookie poisoning.
- Positive security controls reduce your application attack surface by allowing only known good traffic and enforcing access and inspection policies based on identity and context.
- Inline traffic inspection analyzes every HTTP/S transaction between users and private apps, providing a level of visibility into the application layer (L7) that is not possible with traditional network security controls (L4).
- Zero-day threat defense provides predefined controls from the Zscaler ThreatLabZ research team to protect against the latest security threats.
- Browser session protection identifies high-risk users by examining the number of unique fingerprints generated by user browser activity and flagging users with an abnormally high fingerprint count.
- Easy deployment and scalability are possible with one-click activation from the ZPA console, with no new components to install in your environment.

New to ZPA and AppProtection?

- Learn the basics of ZPA, a service that provides secure access directly to applications that your users are approved to access. Unlike a VPN, no network or general browsing access is granted. For a quick demo of ZPA, see ZPA Secure Private Access (https://www.zscaler.com/products/zscaler-private-access).
- For information on configuring AppProtection, see AppProtection for Private Application Traffic (formerly Inspection) (https://help.zscaler.com/zpa/appprotection-private-application-traffic-formerly-inspection).
- Learn more about the Open Worldwide Application Security Project at OWASP (https://owasp.org/).
- The Zscaler ThreatLabZ team maintains a blog at Zscaler (https://www.zscaler.com/author/threatlabz).

Understanding ZPA and AppProtection

The AppProtection controls operate within the ZPA service, providing access to private applications for authorized users. Unlike a VPN that provides access to a network, ZPA provides direct application access to authorized users. For any user who is not authorized to use the application, it will appear as if the application does not exist.

Your applications are not visible to internet users because they are not directly on the internet. Instead, your private applications are behind a Zscaler App Connector or Cloud Connector, a virtual machine that sits in front of your applications. Zscaler App Connector and Cloud Connector allow the applications sitting behind them to be reached only via the ZPA service. Both the Zscaler App Connector and Cloud Connector are outbound-only devices, launching a Transport Layer Security (TLS) tunnel to the Zscaler Zero Trust Exchange (ZTE). Communication is handled only via the tunnel to the ZTE, and any inbound requests are dropped.

Your users run a small agent on their devices that connects them to all of the Zscaler for Users services to which you are subscribed. Zscaler Client Connector makes connections to the ZTE to provide Zscaler services and inspection. This can include ZPA, Zscaler Internet Access (ZIA), and Zscaler Digital Experience (ZDX), depending on your organization's subscription.

Zscaler Client Connector also creates a TLS tunnel to the ZTE cloud. The ZTE acts as both a broker to the applications and as an inspection and enforcement point for policy. When a user attempts to access an application, ZPA first checks the user's policy to see if the user is allowed to access the application.

Figure 2. Connections to private applications with ZPA are secured between users and applications

If the user is allowed to access the application, Zscaler Client Connector launches a connection to the application. That connection is secured inside a Microtunnel, carried within a Zscaler Client Connector tunnel. The ZTE also signals the App Connector or Cloud Connector for the application, and it establishes a Microtunnel with the ZTE for that session.

AppProtection in ZPA Sessions

When the user's application transactions reach the App Connector or Cloud Connector, the data and requests are inspected, terminating any TLS sessions, and inspecting the underlying request. This is where AppProtection inspection and control occurs.

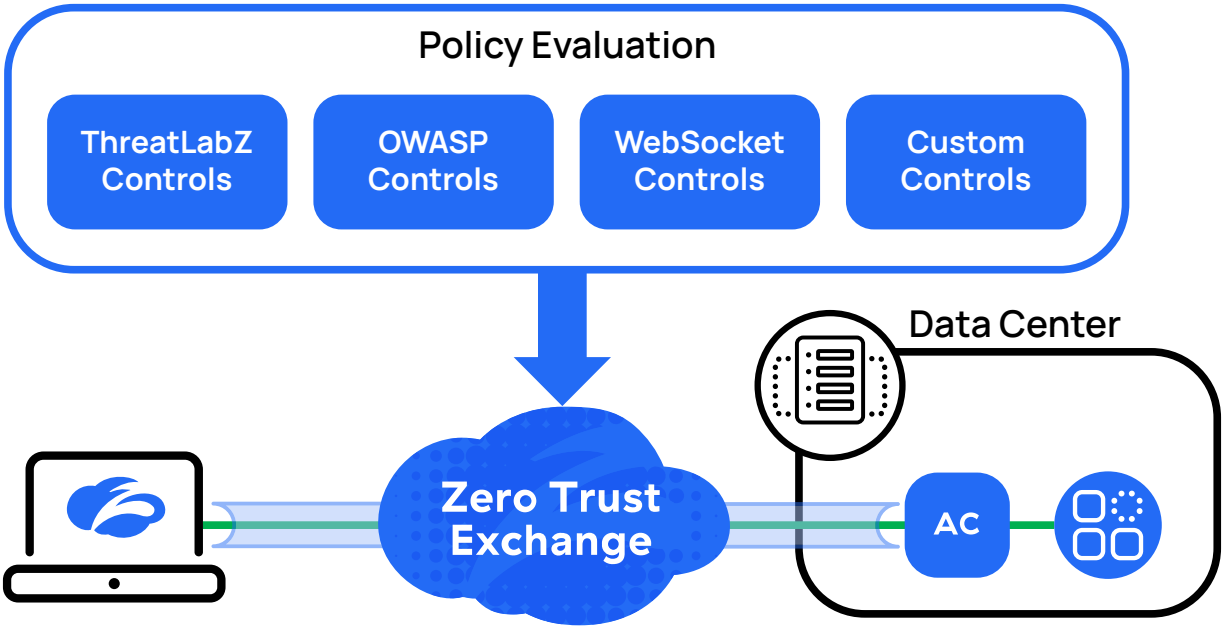

Figure 3. Traffic inspection occurs at the App Connector, where AppProtection policy is applied

The App Connector applies its inspection tools to the transaction. The AppProtection policies that you specify for the transaction are applied at this stage. This could be a single policy applied to all traffic or custom policies that specifically target the application.

To change the number of possible matches, you can adjust the value of the paranoia level for the control. The paranoia level allows you to set a value from 1 (highest concern) to 4 (least concern). As you increase the paranoia level for your control, more inputs are considered when making a match determination.

More inputs can lead to an increase in false positives and the prevention of legitimate traffic. As you tune your controls, it is recommended that you increase or decrease the paranoia value in single-step changes and then observe the results. Zscaler recommends that you set the paranoia level as high as possible without interfering with legitimate users and traffic.

When there is an AppProtection policy match, you select the action that the system takes. This can be to allow the traffic, or to block and redirect the user to a different URL such as your acceptable use policy (AUP). These actions can be set at the common level affecting all controls in the category of control, or for individual rules using specific controls. The categories are as follows:

- ThreatLabZ controls – Developed by Zscaler's security research team ThreatLabZ to address CVE-reported attacks.
- OWASP predefined controls – Web application firewall rules for detecting common attacks.
- WebSocket controls – Controls for inspection of WebSocket traffic.
- Custom controls HTTP and WebSocket controls – Custom regular expressions to match web and WebSocket transactions.

If the request is allowed, it is then forwarded to the second Microtunnel that was initiated from the App Connector or Cloud Connector for that user's session. Even if multiple users connect to the same application, each user session resides in its own set of Microtunnels dedicated to that user and that session.

Understanding AppProtection and Browser Protection Profiles

With AppProtection, you apply controls to traffic and browsers of specific users by building AppProtection profiles. The profiles contain instances of one or more of these control categories: ThreatLabZ Controls, OWASP predefined controls, WebSocket controls, and all custom controls. The AppProtection profile can then be used as a policy object for applying the different controls to a transaction.

Your application and policy dictate the controls you include and the actions to take when a match occurs. These actions can be based on the entire category of controls such as all OWASP controls, or you can have different actions for controls within the category.

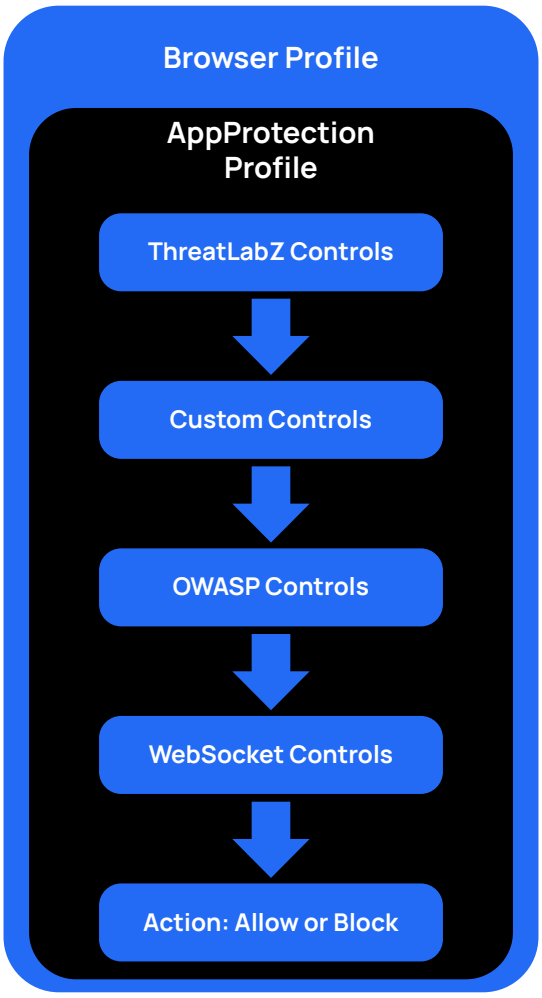

Figure 4. Profile hierarchies allow you to reuse components and provide flexible deployment options

Each AppProtection profile has an ordered set of controls within the profile, and each control category has an execution order. The order of execution is as follows:

1. ThreatLabZ predefined controls
2. Custom controls (HTTP and WebSocket), if any
3. OWASP and WebSocket predefined controls

It is important to understand the order of execution to prevent traffic from being allowed higher in the order that is blocked by a lower rule. When building a new profile, the admin interface shows the order of execution so that you can adjust it prior to deploying your profile.

As mentioned previously, each AppProtection profile contains a paranoia setting. This setting influences the number of matches by adjusting the paranoia level of the control. The paranoia setting allows you to set 1 (highest concern) to 4 (least concern). As you increase the paranoia setting for your control, more input sources and signals are considered when making a match determination. This leads to more matches and potentially false-positive results. Take care to adjust controls so that you do not interfere with legitimate traffic.

In addition to the paranoia level, many controls have a severity rating. The severity rating is factored in when the control is developed and used to help assign a paranoia level. This severity level is provided as an informational item to help you make informed policy decisions. These levels match the terminology used by the CVE process. The levels are Critical, High, Medium, and Low.

AppProtection profiles are also used in the Browser Protection policy. Browser Protection is applied to browser and operating system (OS) combinations to collect information about the user device. You choose the information to collect about a user's browser, device, and location. When a rule matches, it also generates a unique user fingerprint. When a user generates a high number of fingerprints, they will appear as an alert in the browser profile dashboard.

- Learn more at **About AppProtection Profiles** (**https://help.zscaler.com/zpa/about-appprotection-profiles**).
- Learn more at **About Browser Protection Profiles** (**https://help.zscaler.com/zpa/about-browser-protection-profiles**).

Understanding AppProtection Controls

AppProtection is broken into categories to make maintaining your policy easier. Each category of control can be combined into an AppProtection profile to match your policy needs. Each category and individual control has paranoia levels that you can modify to suit your policy needs. You can also subscribe to automated updates from Zscaler's ThreatLabZ team to automatically protect you from new threats to your applications.

If you have subscribed to the ThreatLabZ updates to your controls, you can only set common-level actions for the controls. You must apply the same policy to all controls, because controls are added to these categories as ThreatLabZ publishes them.

Learn more at **About AppProtection Controls** (**https://help.zscaler.com/zpa/about-appprotection-controls**).

ThreatLabZ Controls

Zscaler's ThreatLabZ team is an in-house security research team. Leveraging experts in cybersecurity along with artificial intelligence and machine learning (AI/ML) tool sets, Zscaler can proactively protect your applications against new and emerging threats. In addition to its own research, the ThreatLabZ team builds controls for vulnerabilities reported through the CVE process, and rules link to the Zscaler security portal with more information about the control.

Each control is listed in a table. The controls each have a number, a control name, a severity level, a control action that occurs when there is a match, and the current version of the control as updated by ThreatLabZ. In addition, the control can be expanded to see the following information:

- A description of the control.
- The paranoia level for the control.
- A URL linking back to the Zscaler security portal with more information about the control and vulnerability.
- A list of which AppProtection profiles are currently using the control, allowing you to see where a control is being utilized.

Zscaler recommends enabling all ThreatLabZ controls and subscribing to the ThreatLabZ updates to keep your controls current.

Learn more at **About ThreatLabZ Controls** (**https://help.zscaler.com/zpa/about-threatlabz-controls**).

OWASP Controls

Zscaler leverages controls developed by the Open Worldwide Application Security Project (OWASP) core set of controls for application firewalls. These rulesets are built to detect attacks on web applications using the most common types of attacks. These predefined controls are built into subcategories:

- Preprocessors
- Environment and Port Scanners
- Protocol Issues
- Request Smuggling or Response Split or Header Injection
- Local File Inclusion
- Remote File Inclusion
- Remote Code Execution
- PHP Injection
- Cross-Site Scripting (XSS)
- SQL Injection
- Session Fixation
- Deserialization
- Issues Anomalies

Each control within a subcategory is listed in a table. The controls each have a number, a control name, a severity level, and a control action that occurs when there is a match. In addition, the control can be expanded to see the following information:

- A description of the control.
- The paranoia level for the control.
- A list of which AppProtection profiles are currently using the control, allowing you to see where a control is being utilized.

These controls represent the most common types of attacks launched by malicious actors. These tools are often available in prebuilt vulnerability kits that require little skill to operate. Zscaler recommends enabling all OWASP controls and subscribing to the ThreatLabZ updates to keep your controls current.

- Learn more at **About the OSWAP Predefined Controls Page** (**https://help.zscaler.com/zpa/about-appprotection-controls#predefinedcontrols**).
- Learn more about OWASP controls at **OWASP ModSecurity Core Rule Set** (**https://coreruleset.org/**).

WebSocket Controls

WebSocket connections are a bidirectional protocol that is compatible with HTTP. This protocol, like HTTP, is a Layer 7 protocol and provides a full duplex communication stream over TCP. It leverages the same ports (80 and 443) as an HTTP connection. When communication begins, WebSocket uses an HTTP Upgrade header to switch protocols from HTTP to WebSocket.

WebSocket is supported by major browsers. It is commonly used by applications that must have two-way communication without the need to open multiple TCP connections or poll the server for updates. This can include common productivity applications such as chat and shared document editing.

The WebSocket predefined controls are organized into various categories:

- WebSocket Handshake Headers Check
- WebSocket Framing Errors
- Frame Type Verification
- Client Mask Bit Verification
- Server Unmask Bit Verification

WebSocket handshake errors are verifiable through custom HTTP controls. This is done because at that stage, this is still an HTTP transaction. For more information on custom HTTP controls, see **HTTP and WebSocket Custom Controls**.

Each control is listed in a table. The controls each have a number, a control name, the control source either custom or predefined, a severity level, and a control action that occurs when there is a match. In addition, the control can be expanded to see the following information:

- A description of the control.
- The paranoia level for the control.
- A list of which AppProtection profiles are currently using the control, allowing you to see where a control is being utilized.

Zscaler recommends enabling WebSocket controls to protect your applications from common WebSocket attacks.

- Learn more at **About WebSocket Controls** (**https://help.zscaler.com/zpa/about-websocket-controls**).
- Learn more about the WebSocket protocol at **RFC 6455 The WebSocket Protocol** (**https://datatracker.ietf.org/doc/html/rfc6455**).

HTTP and WebSocket Custom Controls

If you have enabled the built-in controls but still have concerns about applications that you are running, AppProtection provides the ability to define custom controls. Typically, these are custom or customized applications you have deployed that are not generally available to other organizations.

It is important to remember that any custom controls you implement are given second-highest execution priority. They are second only to Zscaler's ThreatLabZ controls. Custom controls also default to a paranoia level of 1, the highest level. You should test your rules with an informed test group before deploying across your organization.

Given the nature of the two protocols, the design of the custom controls for each is very different.

Custom HTTP Controls

Custom HTTP controls are built by selecting data within an HTTP request or response message. This is done by selecting the appropriate type and field where you will inspect for malicious payloads. The payload itself is matched with a regular expression string against the values in that field.

When defining a control, you must select either a request or a response control type. Each of those has subtypes available:

Request

- Request Header
- Request URI
- Query String
- Request Cookie
- Request Body
- Request Method

Response

- Response Header
- Response Body

After you select the target type and subtype to scan, you must then supply the custom regular expression to match the payload that you expect. Zscaler recommends engaging an expert on regular expressions to help build and refine your custom controls.

Learn more at Defining Regular Expression Values (https://help.zscaler.com/zpa/defining-regular-expression-values).

The Custom Controls page shows each control listed in a table. The controls each have a number, a control name, a control type listing the elements that make up the control, a severity level, and a control action that occurs when there is a match. In addition, the control can be expanded to see the following information:

- A description of the control.
- The paranoia level for the control.
- A list of which AppProtection profiles are currently using the control, allowing you to see where a control is being utilized.
- The regular expression that is run against requests.

Learn more at About Custom Controls (https://help.zscaler.com/zpa/about-custom-controls).

Custom WebSocket Controls

When you build custom WebSocket controls, they are not given a special display like custom HTTP controls. Instead, they are displayed with the prebuilt WebSocket controls on the WebSocket Control dashboard. See WebSocket Controls for more information.

When defining custom WebSocket controls, you must select either a request or a response control type, and a subtype of either Max Payload Size or Max Fragments Per Message. Each subtype requires you to add a value to act as the maximum for that control. You must also choose to allow or block requests that match your settings.

After you save your custom WebSocket control, it is available for use in your AppProtection profile.

Learn more at Configuring WebSocket Controls (https://help.zscaler.com/zpa/configuring-websocket-controls).

ZPA AppProtection Deployment Considerations

Because AppProtection takes direct action on your traffic, you must consider where and how you are applying your controls. You want to ensure that you are blocking actual threats and not your organization's legitimate traffic. As with any other security service, you should gradually roll out AppProtection so that you can assess how these controls affect your user traffic.

Zscaler recommends enabling AppProtection initially in monitor-only mode. This gives you an idea of how many application transactions would have been caught. It also enables you to check the legitimacy of the matches. You might find that the paranoia setting is too high or too low for a particular threat or category. During monitoring, you should continue to adjust settings to fine-tune your matches and traffic.

When you feel that you have a well-tuned system, you should begin deploying enforcement to a small group such as the IT department. Make sure to keep a separate system to enable you to recover the system if your changes lock you out of the Zscaler interface.

After deployment to the initial small group, select another small group or location in which you can deploy the changes and where, ideally, there is a broad spectrum of application usage. Make sure that users are alerted to the change and know how to immediately report any issues to IT.

Summary

Zscaler offers powerful application security built into its Zero Trust Network Access (ZTNA) solution, ZPA, to protect internal apps and infrastructure against the most prevalent cyberattacks. The ZPA AppProtection service provides high-performance, inline security inspection of the entire application payload to expose threats. It identifies and blocks known web security risks, such as the OWASP Top 10, and emerging zero-day vulnerabilities that can bypass traditional network security controls. With Zscaler's unique Zero Trust architecture, you get AppProtection as part of an integrated set of security services that is only available with advanced ZTNA solutions.

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.