

Overview

Zscaler enables any end user, from any place, using any device, to experience a rich Internet experience while enforcing security and business policy. Rather than buying appliances such as web proxies, URL filtering and anti-virus gateways at each Internet egress point, Zscaler offers an on-demand service through a globally distributed cloud infrastructure. Zscaler provides risk mitigation and policy enforcement for businesses through this service, while enriching the user's Internet experience. To leverage Zscaler's service, companies simply define their corporate security control and compliance policy by accessing the Zscaler administration portal (once provisioned) and redirect all web traffic leaving the network to one of the 40+ data centers in Zscaler's global infrastructure. Based on an organization's policy, traffic is blocked, throttled, or allowed to access the Internet. As the browser retrieves the web pages, Zscaler scans it for a range of malware threats and delivers clean traffic to the end user. The true power of Zscaler's cloud lives within both its innovative cloud security technologies as well as in the methods by which it authenticates and tracks users.

Authentication Challenges of a Cloud Delivered Security Service

To truly have a complete security solution, security vendors must provide user-based Authentication, Authorization and Auditing (AAA). Without user-based AAA, there is no way to tie a specific user to a specific transaction. Traditional security appliances have had it easy when it comes to enforcing user-level policy and providing user-level granularity in reporting; a traditional web proxy or other security appliance sits within a network perimeter making access to authentication and authorization information easy. Strategies vary based on the vendor and technology, but generally one of three scenarios are selected:

- A software "connector" on a dedicated machine for brokering authentication¹
- Direct LDAP-based lookups from the security device to the existing directory infrastructure
- Software client installed on end users' devices²

In all three of these cases, while organizations likely raise an eyebrow to additional software being required to make the solution work, the network and security teams usually don't concern themselves with how the data traverses their network (assuming its encrypted within an SSL tunnel, as a minimum bar). The lack of concern is simply based on the premise that a user's credentials (obfuscated or not) never leaves their zone of control.

Authentication is one of the hardest problems for cloud providers for this exact reason; credential information must exist (even if it's simply identifiers) outside a customer's perimeter. User-based AAA is needed to enforce policies and report at a user and group level granularity. The difference between a traditional security appliance vendor and a cloud delivered service lies in the fact that a cloud must have some method of synchronizing that which already exists within a customer's AAA infrastructure. This can be a touchy subject within the confines of an IT Security organization especially since there are only two real options: allow the vendor to reach in or actively push the information out. The sophistication and efficacy of these solutions, again, depend on the methods employed. However, at the end of the day, the cloud vendor **MUST** have identical user and group info as that which exists within the customer's existing environment. Once this information is synchronized, then the cloud provider must ensure that each user's session is authenticated, authorized (policy enforced) and auditable (for later inspection and forensics). Zscaler enforces corporate (Internet usage) policy in both directions—inbound and outbound. Once authenticated, when end users attempt to access web resources, Zscaler's first step is to validate policies

¹ Also referred to as a "Mini-Squid"; SQUID is an open source, software-based caching proxy for the Web. Connectors are typically mini-proxies.

² Also referred to as a "Fat Client"; Fat refers to how large some of these clients have reached in install size (<50MB) and reach (within the OS).

regarding the destination, requests the content of the origin content server, inspects what is returned for threats. Assuming all steps occur with no errors, policy violations or threats, the content is then delivered back to the user for viewing.³ To ensure user-level AAA is available, Zscaler takes advantage of persistent cookies.

What is a Cookie?

A cookie, also known as a web cookie, browser cookie, or HTTP cookie, is a piece of text stored by a user's web browser. Traditionally, web cookies have been used for session management functions including authentication, website personalization and user tracking. A web cookie consists of one or more name-value pairs and is sent as an HTTP header by a web server to an end user's web browser. This web cookie is then sent back, unchanged by the browser, each time it accesses that server. A server can set a web cookie with or without an expiration date. Those without an expiration date exist until the browser terminates; the browser holds those with an expiration date until the expiration date passes. Users may also manually delete web cookies anytime. As text, web cookies are not executable. The problem with normal web cookies lies in their lack of persistence; given the ease at which cookies are deleted, Zscaler needed to leverage an approach that ensured end users were not constantly being asked to re-authenticate every time a site was visited. As a result, Zscaler uses an innovative approach that leverages Local Shared Objects (also known as "flash cookies").

Zscaler's Use of Flash Cookies

The Adobe Flash Player plugin (formerly developed by Macromedia) uses a type of cookie known as Local Shared Objects (also known as "flash cookies" or LSOs). Flash cookies can be used in a way very similar to normal cookies and yet are an attractive choice for web developers because they maintain a client-side persistence beyond what normal web cookies can. Flash cookies are also more sophisticated in that they are capable of storing up to 100 Kilobytes of data versus the 4 Kilobyte limit of a normal text-based web cookie. Additionally, unlike normal web cookies, flash cookies aren't stored within the browser; flash cookies are created and stored independent of the browser (multiple installed browsers will look to the same location for flash cookies as its operating system dependent versus browser dependent). A user has to know about LSOs and then take the necessary steps to actually clear them out. These persistent client-side cookies are paramount for accomplishing user level AAA within the Zscaler cloud because they enable far more flexible and yet sticky expiry periods while also ensuring users are tracked effectively and transparently.

Zscaler Authentication Paradigms

Zscaler has two primary user identification scenarios:

- **Known Locations** – Zscaler has identified specific gateway IPs (static) and traffic is originating from this location. The organizations' administrator accomplishes this by inputting the public IPs of their Internet gateways into the Zscaler administration interface. Once Zscaler knows of these egress points, the cloud associates all traffic originating from these locations with that specific organization. This further enables "authentication" to be enabled or disabled—end users will or will not be required to submit a valid email ID and password based on policy defined in the administration UI.
- **Unknown Locations** – This scenario is defined as traffic coming from a location that is unknown to the cloud. Specific examples include mobile users and sites with dynamic IPs. In the case of unknown locations, all users must be authenticated via a valid email ID and password. Once authentication occurs, a cookie will be stored on the user's device, eliminating the need for multiple authentication challenges.

User Management

There are three available options for user management:

³ For more information about traffic flow with Zscaler, please refer to the Traffic Forwarding documentation.

- Hosted Database
 - With a Hosted DB, an administrator manages all creation, modifications, deletions and notifications for individual end users within the administrative UI. Names, email addresses, group information and passwords are all maintained on the cloud. Expiry dates and password strength are also managed within this interface. In addition to manually creating end users within the UI, administrators can import end users from file. The Hosted DB option also provides the ability to email temporary one-time tokens to new users such that administrators don't have to manually create and distribute passwords. This capability also enables users to request a token (once every 24 hours) in order to reset forgotten passwords.
- Active Directory Synchronization
 - Zscaler has the ability to connect to your Active Directory and import your user, group, and department information into your organization's profile within the cloud (end users are managed within Active Directory).
- LDAP Synchronization
 - Virtually identical to Active Directory; users managed within the existing directory infrastructure while appropriate identification information (user, groups, departments, etc.) are populated into the cloud

Zscaler also provides an option for an on-premise Authentication Bridge.

Federated Identity Systems & SAML

Federated Identity systems prescribe a method for portability of identity information across otherwise autonomous security domains. It is an essential unifying layer in helping user-authentication and authorization scale effectively for cloud deployments. By having Service Providers agree on federated identity standards, businesses do not need to create redundant user-directory information for cloud services, end-users do not need to be burdened by any new authentication methods, and IT and Security administrators do not have to stress out over having to change any of their existing policies. OASIS' Security Assertion Markup Language (SAML) is an XML-based framework for marshaling security and identity information and exchanging it across domain boundaries. SAML provides a means for communicating user authentication, entitlement, and attribute information by providing business entities a way to make assertions regarding a subject to other entities, such as a partner company or another enterprise application. As with most service providers whom have enabled integration with federated identity systems via the use of SAML, Zscaler provides a means for organizations to authenticate users without users' passwords leaving an organization's domain of control. Please refer to our brief on Federated Identity & SAML for more information.

One-Time Passwords

The One-Time Password (OTP) mechanism within the Zscaler service enables administrators to manually create and distribute passwords to new end users as well as providing end users a way to reset their own forgotten passwords. OTPs function in two paradigms, when enabled:

1. End User Requests - In the case of end users requesting an OTP, when an end user is prompted to enter his or her password after successfully submitting their user ID to the authentication prompt, a "Temporary Password" hyperlink is also displayed in password form. It will then generate a one-time password and send it to the email address just entered. The temporary password is only valid for 24 hours. If the end user does not use the temporary password within 24 hours, he or she will have to generate a new password by clicking the "Temporary Password" link once again.
2. Administrative Use - Within the administrator interface, when One-time Token is enabled, a new Send Authentication Email link displays in the admin interface. This link opens a pop-up window in which you may automatically send passwords to users, groups, or departments. All users provided with an OTP are required to change their passwords when they authenticate with the OTP the first (and only) time. While One-time Token is enabled, the Zscaler is responsible for handling the passwords for affected users. *While One-time Token may be used with Hosted User Database, Active Directory, or OpenLDAP authentication methods, it will over-ride the LDAP lookup in Active Directory or OpenLDAP*

when new users authenticate, or other users re-authenticate. The LDAP over-ride remains in place for those end users who authenticated with the One-time Token until Temporary Authentication is set back to None.

Pre-Provisioning

HTTP has a very rich set of methods to perform a variety of functions. Through the use of a Representational State Transfer (RESTful) API, Zscaler uses existing features of the HTTP protocol to enable administrators to download pre-generated authentication cookies (flash cookies) securely over HTTPS (in bulk if necessary). The purpose of this tool is to empower two use cases simultaneously:

- Enable transparent authentication via pre-provisioned cookies (prevents the user from being challenged to authenticate).
- Simplify deployments of the Zscaler service across large numbers of end user devices (again, where transparent authentication is required).

The API for requesting pre-provisioned cookies is set up to provide a granular set of queries for batched user selection based on a variety of attributes (including group, department and time). Please refer to Pre-Provisioning Cookies datasheet for more information.

Zscaler, Inc.

392 Potrero Avenue,
Sunnyvale, CA 94085

USA

+1 408.533.0288

+1 866.902.7811

www.zscaler.com