Using SDP to Secure Access to Private Apps Across Multi-Cloud Environments

Written by

Nathan Howe

ZPA Architect, Zscaler

zscaler™

Today the majority of enterprises are running their private applications on a variety of infrastructure types. This spans from traditional data center hardware to public cloud, or IaaS providers, such as Azure and AWS. In working with customers, we've realized that it's rare for an organization to go all in on just one IaaS provider. In fact, about 50 percent of enterprises today have deployed two or more cloud providers. This means that there is a real need to be able to manage and secure access to applications across multiple service providers. But how?

For years, enterprise security teams have treated their on-premises environment and public cloud as different IT environments. We believe this is due to two factors. The first is that security professionals are comfortable with the idea of managing their security applications versus having to rely on security technology hosted by someone else. Second, 44 percent of the time, security of the cloud is managed by someone outside the actual security organization. Rather than taking a holistic security strategy, this forces teams to rely on completely separate security technologies. This not only increases the amount of time and energy required to manage an unnecessary number of security technologies, but also cost. When you factor in the user, who now has to think about whether or not an application is running on AWS, Azure or the data center, the end result is a more complex IT environment and a frustrating user experience.

Legacy methods:

- Build a singular access path (gateway) for your users to connect to and then route them across some sort of backend network between various locations.
- Build multiple access paths, one for each location, and have your users connect on demand to the location they require.

Both of these options have their benefits and challenges, of which neither offer the end user the simplicity they deserve, or need, when trying to get their work done.

At Zscaler, we're working with some of the world's largest organizations to help them navigate this new space of multicloud and hybrid cloud, avoid common mistakes and ensure that their cloud adoption doesn't outpace their security capabilities.

Key to this is educating and enabling enterprises to embrace what Gartner calls a software-defined perimeter (SDP). This is a new set-up that replaces inbound VPN gateway appliances, redefining how users connect to apps by decoupling application access from network access altogether and without exposing the app to the internet. The SDP architecture serves as a faster and more secure alternative to the incumbent, network-centric processes. In the past, this involved setting up additional site-to-site capabilities, deploying virtual firewalls, and managing ACLs and policies for each one of their appliances and IaaS services.

This Architectural white paper, we will cover the following:

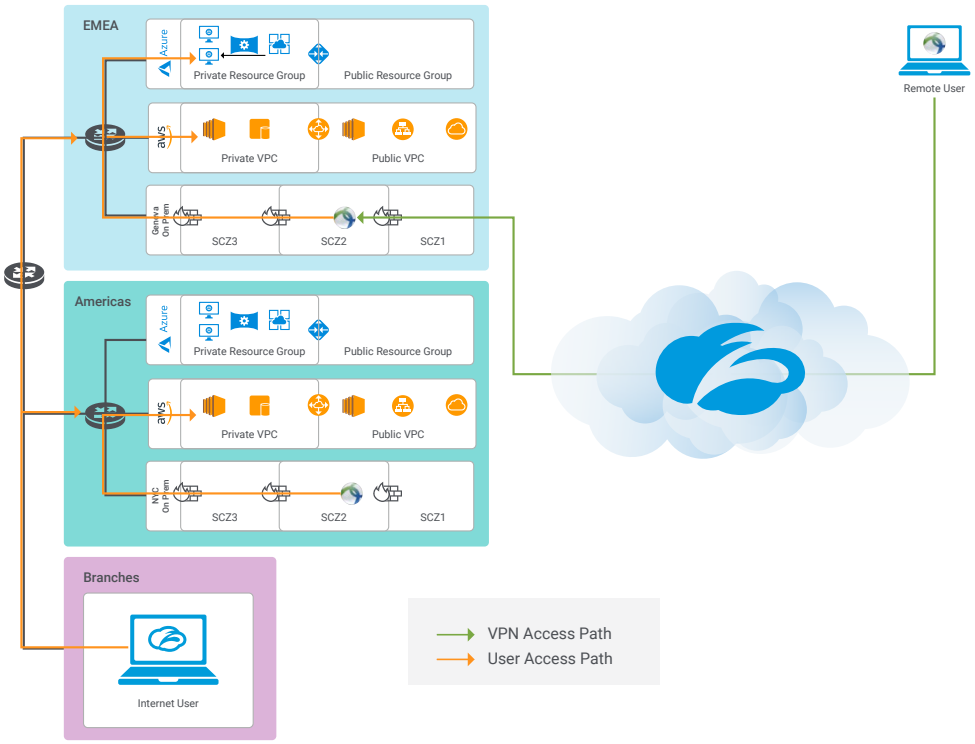
- Architectural differences between legacy network security (i.e remote access VPN) and modern user-defined security architectures (i.e SDP)
- Showcasing what a true zero trust security architecture looks like for multi-cloud
- Phases to go through as you look to secure access to private apps running on multi-cloud

In this guide, we will use a typical enterprise scenario as an example. Our fictional enterprise will be called Geo-Enterprise SA, a manufacturer based in Geneva, Switzerland. Geo-Enterprise has traditionally hosted its applications within one of two on-premise locations, Geneva and New York, U.S., and is looking to adopt a multi-cloud strategy. This will allow it to reduce costs and improve scale at a level that is unachievable by relying solely on the data center or even a single IaaS provider. Historically, Geo-Enterprise has routed access to its cloud applications over its private links to the respective application location. User traffic would either route over the private link from a branch or from a remote access point and through the access gateway in the on-premises data centers. As cloud app locations are optimised for internet paths, generally the private links would be overwhelmed in peak periods and complicated to maintain.

Current Architecture of Geo-Enterprise

Geo-Enterprise is beginning to address how users access private applications across its environment. Approximately 40 percent of the workforce is fully mobile remote workers, with another 30 percent being campus-based users with work-from-home and travel remote access requirements.

Geo-Enterprise is currently running private apps on Microsoft Azure (Azure) and Amazon Web Services (AWS), as well as the incumbent data center infrastructure it has been using for years. This integration is done through an interconnected private MPLS network divided into two global regions – the U.S. and Europe.



Recently executives have led a push to have the majority of the company's applications migrated to its cloud environments in the upcoming 12 months, with a goal to complete the migration to cloud during the next 36 months. The company views this as a critical way of reducing overall costs to the business. However, certain applications will remain in its data center environments with the goal of consolidating its on-premises presence to a minimal data center in Geneva.

Beyond the shift of the application locations, there is a vision for a future model of separating user and service network functions as well as moving towards a zero trust networking strategy.

Geo-Enterprise is currently using Cisco AnyConnect as its remote access VPN technology. But, the company would like to reduce its dependence on VPN to access private applications running in data center and public cloud (IaaS / PaaS) environments.

Traditionally, all remote access traffic has been routed back to VPN concentrators in the primary data centers, including traffic destined for private apps hosted on public cloud providers.This presents a significant latency challenge for remote users. The complexity and expense of distributing and scaling VPN appliances, along with the complex infrastructure required to support and protect them, is an ongoing challenge.

Objectives

The company needs a simpler architectural model to enable direct access to internal applications in emerging cloud environments as well as existing data center environments.

The primary business driver is a readiness for the next generation of application delivery to a modern workforce as applications migrate to the cloud. Geo-Enterprise has already adopted cloud-hosted solutions and has defined a strategy. But as it looks to execute on this strategy, Geo-Enterprise anticipates that end users will continue to demand a more flexible and seamless application delivery service. To deliver this, Geo-Enterprise will need to break away from data center appliances and embrace a globally distributed cloud-hosted security platform that allows for centralized management.

As more private applications move to public cloud providers, it will be critical to enable simultaneous user access to apps in data center and cloud environments, while avoiding slow and costly backhaul wherever possible. As the network transforms, so too must its security control, while maintaining a cost-effective deployment model. More-effective tools to optimize the configuration and monitoring of access to apps running in public cloud will be needed to ensure additional risk is not introduced.

As part of this journey, Geo-Enterprise wants to embrace a zero trust model for it's private apps. In this new model, clients would be provided access to applications based on their identity and posture, rather than by the network from which they are connecting. This architecture fundamentally changes the cost structure and management burden of securing the enterprise networks, as all user endpoints can now be considered as "outside" the traditional security perimeter regardless of whether they are remote or on-premises.

Traditional remote access VPN infrastructure presents a risk to Geo-Enterprise, as it broadens the threat surface by always putting a user on the network. Zscaler Private Access (ZPA) overcomes this risk by following four key security tenets:

- Connect users to private applications without bringing them on to any internal networks
- Inside-out connections ensure private apps are never exposed to the internet
- Application segmentation without relying on complex and costly network segmentation
- Using the internet as a secure network transport without relying on L3 IPsec VPNs

This secure approach means that there can be no lateral movement to other Geo-Enterprise applications. Furthermore, those applications that the user does not have access to remain completely dark and cannot be discovered via port scans or any other mechanism, either locally or from the internet directed at the hosted environment as applications do not receive any inbound connections directly from users.

Vision

Geo-Enterprise has already reached a state where road warriors need secure access to applications, and office-based employees need to connect directly to cloud-based applications. To achieve better results, ZPA provides:

Global, Unified, Secure and Simple Access: Seamless, fast access to applications, regardless of user and application locations, is critical to the function of businesses. The increasing diversity of infrastructure on which private apps are running means that security and access controls for all users should be designed and applied globally, and be capable of ensuring the most efficient route for users to access apps. Applications would be equally accessible for on-premises and remote users from desktop and mobile devices. This would revolutionize the Geo-Enterprise connectivity experience.

Increased Security Visibility and Control: All traffic flows to internal applications would have controls to ensure only authorized users can access applications. Comprehensive visibility into who is accessing which apps ensures understanding of current workflows and enables design of granular access policies to meet the needs of IT and end users.

Cost Avoidance: Security and network infrastructure that was once required to enable your remote users access and mitigate the associated exposure could now be reduced or even eliminated as part of this project. As a result, network infrastructure and operational costs could be minimized.

Cloud Readiness: As applications shift to the cloud and users are increasingly mobile, the only scalable way to apply security is by connecting users to applications, rather than connecting endpoints to networks. Leveraging the Zscaler platform, Geo-Enterprise would be empowered to continue to adopt cloud applications without being constrained by current and future infrastructure costs.

Zero Trust: Shifting to a model of application access that uses identity and posture to provide access, rather than network connectivity, would help protect enterprise networks and information against lateral movement, unauthorized access, and insider as well as external threats.

Zscaler Private Access (ZPA)

Zscaler Private Access (ZPA) is a cloud service that provides users with seamless and secure access to private applications without placing users on the network and exposing them to the internet. Instead of the network-layer tunnels used in legacy remote access methods, ZPA provides authorized users access to specific applications via encrypted, per-session microtunnels that are only created upon demand. This software-defined perimeter (SDP) service delivers on Gartner's Continuous Adaptive Risk Trust Assessment (CARTA) framework for remote access, which builds on the concepts of zero trust networking first developed in 2008.

Before examining the solution and its various elements, it is useful to consider the goals of the system design:

Security – While security is a consideration in the design of almost any networking equipment, it was the fundamental requirement in the creation of ZPA. The private, internal applications to which ZPA provides access typically hold an organization's most sensitive data, and this fact is a paramount consideration. Every element of ZPA, including device architecture and inter-device communications, has been built to ensure the security of data in motion and data at rest.

Reliability – Any system designed to provide access to mission-critical internal resources must be highly reliable and available at all times. The ZPA solution delivers reliability with infrastructure that is globally distributed and maintained by Zscaler 24/7. Individual components of the infrastructure are designed to be able to fail gracefully with minimal user impact and without compromising security protections. Additionally, individual elements of the infrastructure can be incrementally upgraded without any end-user impact.

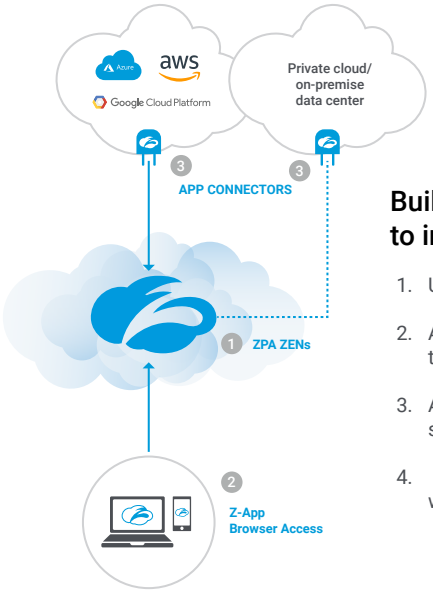
Scalability – ZPA has been designed to support users from many globally distributed organizations all accessing a wide variety of applications at the same time. The global ZPA solution scales across organizations and grows with them, seamlessly handling traffic spikes without interruption.

Manageability – This quality encompasses a collection of attributes that combine to make the ZPA system useable day-to-day. ZPA interacts with a wide variety of customer configurations and does not depend upon any specific computing environment. The solution is easy to deploy, manage, monitor, and troubleshoot — all essential attributes given the sensitive nature of the assets that ZPA protects. Because ZPA is a service rather than an appliance, there is no hardware to deploy, test, manage or upgrade. Finally, the solution brings simplicity to enterprises' infrastructure as a whole by abstracting applications from the networks upon which they reside, simplifying changes in application location, and easing moves from the data center to the cloud.

Innovative design

1. **ZPA ZENs** - secure user to app connection
   Cloud Policy engine - user to app access rights
2. **Z-App / Browser Access** - request access to app
3. **App Connectors** - sit in front of apps - outbound-only connection

Zscaler cloud brokers a secure connection between the Z-Connector and Z-App



Built for zero-trust access to internal applications

1. User are never on your network
2. Apps are invisible-never exposed to the internet
3. App segmentation without network segmentation
4. Use Internet as a secure network without remote access VPN

Before examining how ZPA handles the security risks implicit in remote access, it is useful to first consider the risks posed by remote access VPNs. The use of a remote access VPN requires that enterprises expose infrastructure to the internet to terminate a request. Regardless of the encryption strength of an IPsec tunnel, the "open ends" of the tunnel greatly increase the potential for an attack. It also introduces the possibility for an infected device to connect to the private network and attack other devices on that network. ZPA has created an architecture that addresses each of these risks, as follows:

The end device is never directly connected to the network on which the applications reside. ZPA instead delivers functionality similar to a forward and a reverse proxy acting together. This ensures that networks and applications cannot be infected or exploited by open L3 network tunnels.

The user's visibility of available applications is limited to those to which the user is authorized to connect. ZPA has been designed to provide named user to named application access only. This allows the benefits of micro-segmentation and access control.

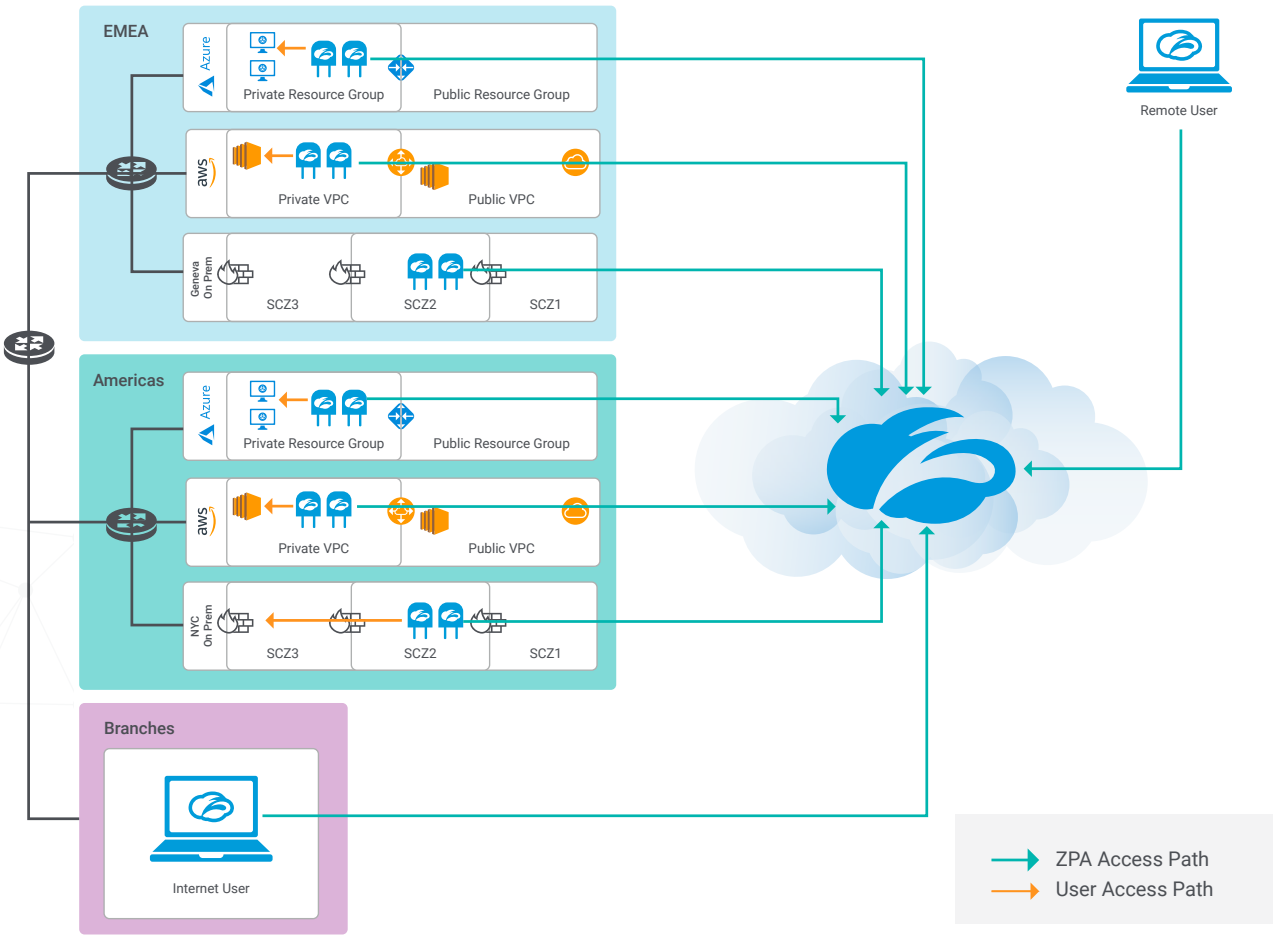
Elements on client and application sides are certificate-pinned. Zscaler acts as its own Certificate Authority (CA), and the Zscaler App on the client side and the ZEN Connectors on the application side are certificate-pinned to eliminate the risk of an attacker impersonating the CA to hijack the session. In zero trust mode, the customer can provide the signed certificate to ensure complete privacy.

A useful set of benchmarks can be derived from requirements, such as PCI-DSS and HIPAA, both of which mandate documented security controls to protect data in motion and data at rest. Another criterion, which is crucially important in ZPA, is the availability of mission-critical assets accessed through a remote access solution.

01 Geo-Enterprise's Future SDP Architecture

Leveraging the ZPA platform, Geo-Enterprise can experience decentralized user access to private apps in disparate environments. In the initial production phase, this access will be across the core Geo-Enterprise locations in the Americas, Europe and Asia and will incorporate access for administrative users to critical applications, such as SAP.

In subsequent phases, Geo-Enterprise will explore additional use cases, including access to cloud-migrated applications for both on-premises and remote users, and an M&A experiment hosted in Azure. During these phases, it is critical that users are able to access private apps in public cloud provider environments directly, rather than backhauling to regional or central data centers. This requirement will hold true of other cloud applications as they are deployed.



Decentralized user access to private apps in disparate environments is the critical next step that Geo-Enterprise must take on the journey to the cloud. One of the first and most critical applications to make this journey was SAP. It is critical that users are able to access private apps in public cloud provider environments directly rather than backhauling to regional or central data centers. This requirement will hold true of other cloud applications as they are deployed.

Simplicity and agility are key to success; using the ZPA cloud platform to implement security controls will provide not just improved security, but also the agility and simplicity required to migrate apps to the cloud. As the simple diagram above shows, Zscaler will provide an optimized path to private applications in data centers or in cloud environments across a global cloud footprint. Geo-Enterprise users will connect - via Z App or browser-based access - to Zscaler's global cloud platform, and Zscaler then becomes the one path for all private app traffic. This brings all policy controls, reporting, and visibility into a single unified platform. Diversity and failover will be provided by the distributed Zscaler cloud as well as App Connector groups and redundancy. The key is that during this transition, Zscaler provides the consistent end user experience and policy controls that are required by the business.

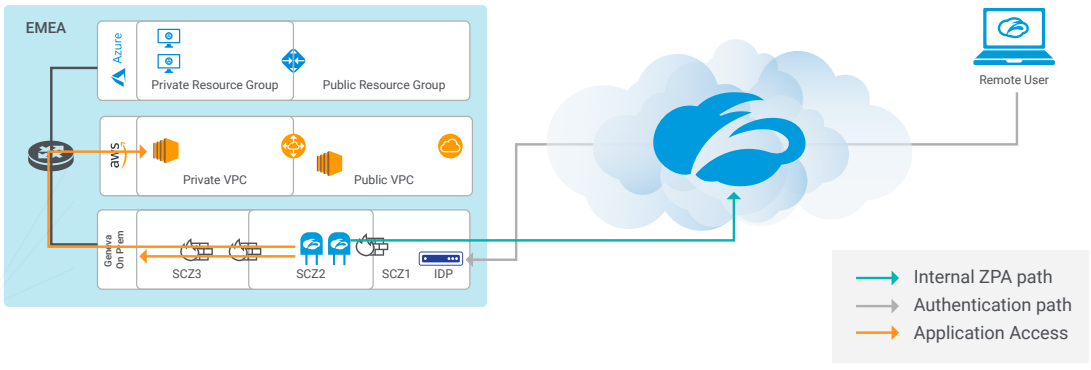
The Zscaler App and browser-based access methods will be used to bring all internal traffic to the Zscaler cloud for distribution to private apps in data centers and cloud environments, via the App Connectors. As this will be a phased approach, after Geo-Enterprise addresses remote access with Zscaler, you can transition over time to the full zero trust approach supported by the Zscaler platform. This will include on-premises users as well as remote users and will eventually allow Geo-Enterprise to forward all user traffic through ZPA for visibility, control, and distribution of access.

02 A Phased Journey to Adopting an SDP Architecture

Phase I Initial Use Case - Prioritize Accessibility

The first stage in this journey will be to leverage Zscaler to enable secure application access and replace the legacy remote access VPN solution, Cisco AnyConnect. In many cases, enterprises will not have a comprehensive view of which apps are currently being accessed or which users are accessing those apps. Several ZPA features support this phase, offering initial visibility with minimal disruption:

- Selective ZPA enrollment
- Zscaler App co-existence with current VPN client
- Trusted network detection
- Application discovery
- Log streaming
- Flexible access policies
- Custom re-authentication policies
- Multiple access methods



For organizations that are already delivering Zscaler Internet Access to their users via the Zscaler App, selective enrollment allows complete control of which Z App users participate in ZPA. ZPA service entitlement can be configured only for users in specific groups, facilitating a smooth transition from proof of concept to pilot to expanded user communities.

The Zscaler App can be provisioned to endpoints that already have a remote-access VPN client installed without conflict. Once enrolled in ZPA, users can be instructed to terminate their VPN connection and access applications as they normally would. If any configuration or access issues arise, users can turn off the ZPA service and reconnect to the VPN, ensuring minimal downtime as ZPA is deployed throughout the user community. Once all applications are working as desired over ZPA, the remote access VPN client can be removed from the end-user devices.

ZPA traffic forwarding can be controlled based on whether the user is on or off the trusted network. Trusted network detection can leverage a variety of network criteria; a common initial configuration is for ZPA to forward traffic when the user is off network and bypass when the user is on network, mirroring existing remote access VPN behavior.

ZPA application discovery acts as a cloud-access security broker (CASB) for internal apps, enabling an administrator to discover previously unknown applications (Shadow IT) running in a public cloud (or data center) and apply granular access controls for them to minimize risk. Teams start broadly with their initial access policies by allowing access to an entire application domain or subdomain, as well as ranges of IP addresses, similar to the open connectivity of a remote access VPN. As users access specific resources, those resources are captured in a discovered applications database, and granular details of the access (which apps, what users, where in the environment) are logged and can be streamed via syslog to back-end SIEM or log correlation systems for visibility and reporting. The discovered apps database and detailed user activity logs provide the foundation for development of more granular, app-centric access policies.

Certain critical applications may require more restrictive access controls to ensure that only authorized users access them, even during the initial discovery phase. These applications can be defined in parallel with application discovery, and granular access policies applied to ensure that only authorized users on appropriate devices access these apps from day one.

For key infrastructure services, such as Active Directory domain services, exemptions can be made to the standard re-authentication policy to ensure that users are always able to access domain controllers, ensuring seamless access to distributed file shares (DFS), integrated Windows authentication (IWA), and internal Kerberos SSO.

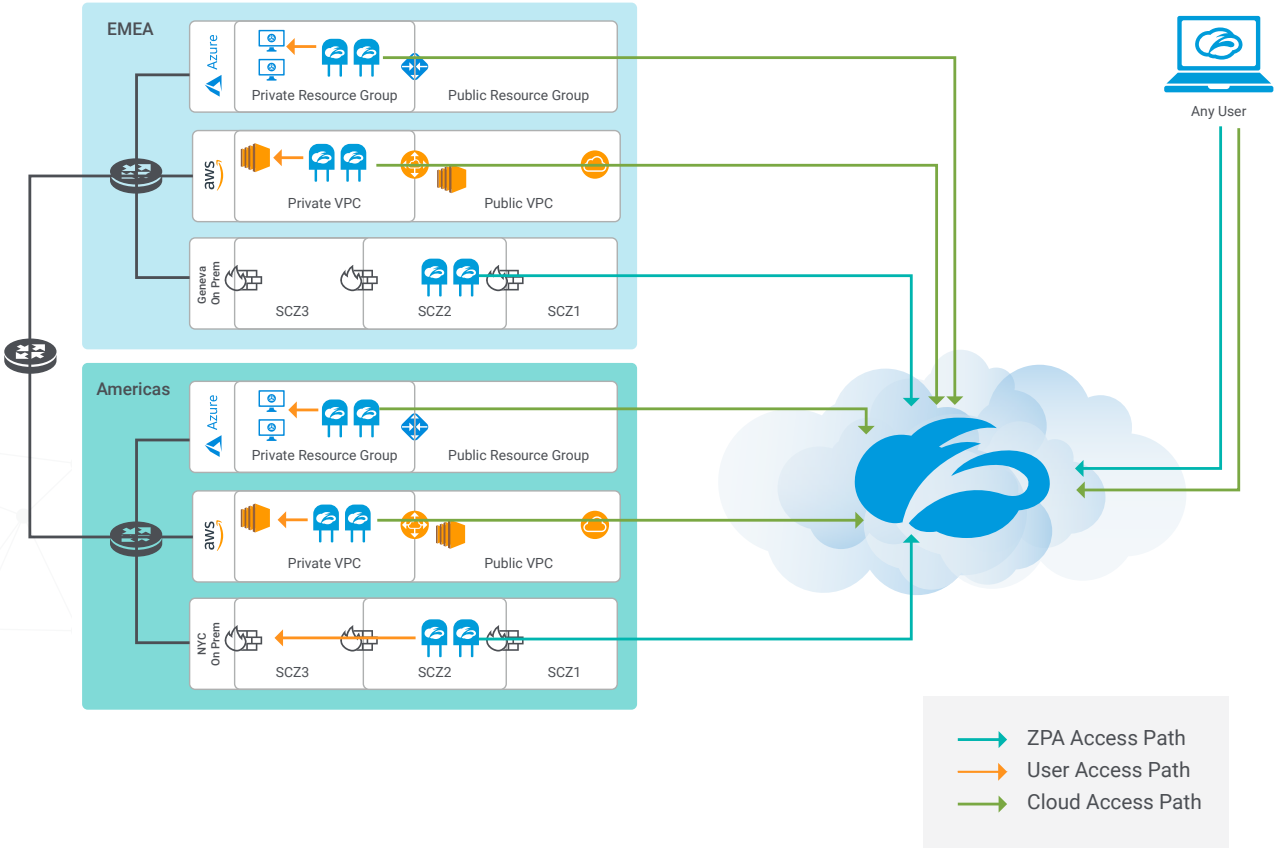
For users requiring access to web apps and other client-to-server TCP and UDP access, the Zscaler App provides seamless, transparent access to applications with a consistent user experience whether on-premises or remote. For users requiring access only to HTTP/HTTPS applications that can be accessed via a web browser, browser-based access provides simple access to applications without requiring installation of software on the user's endpoint.

Once the ZPA service is in production, all remote access traffic flows to private applications will be facilitated by Zscaler, increasing productivity for end users and visibility for administrators. This will allow Geo-Enterprise a consolidated view of all remote access traffic globally. This will also significantly reduce backhaul traffic between data centers and cloud environments, and allow for unified control of policy.

As this stage completes, Geo-Enterprise will have rich visibility into the private applications that are being used today, enabling analysis of end-user activity, as well as granular access controls for critical applications. Once app traffic is completely migrated to Zscaler for an initial group of users, the next steps are to expand the use cases/user communities and to continue developing a more granular access policy.

Phase II Initial Use Case - Enable Cloud Access

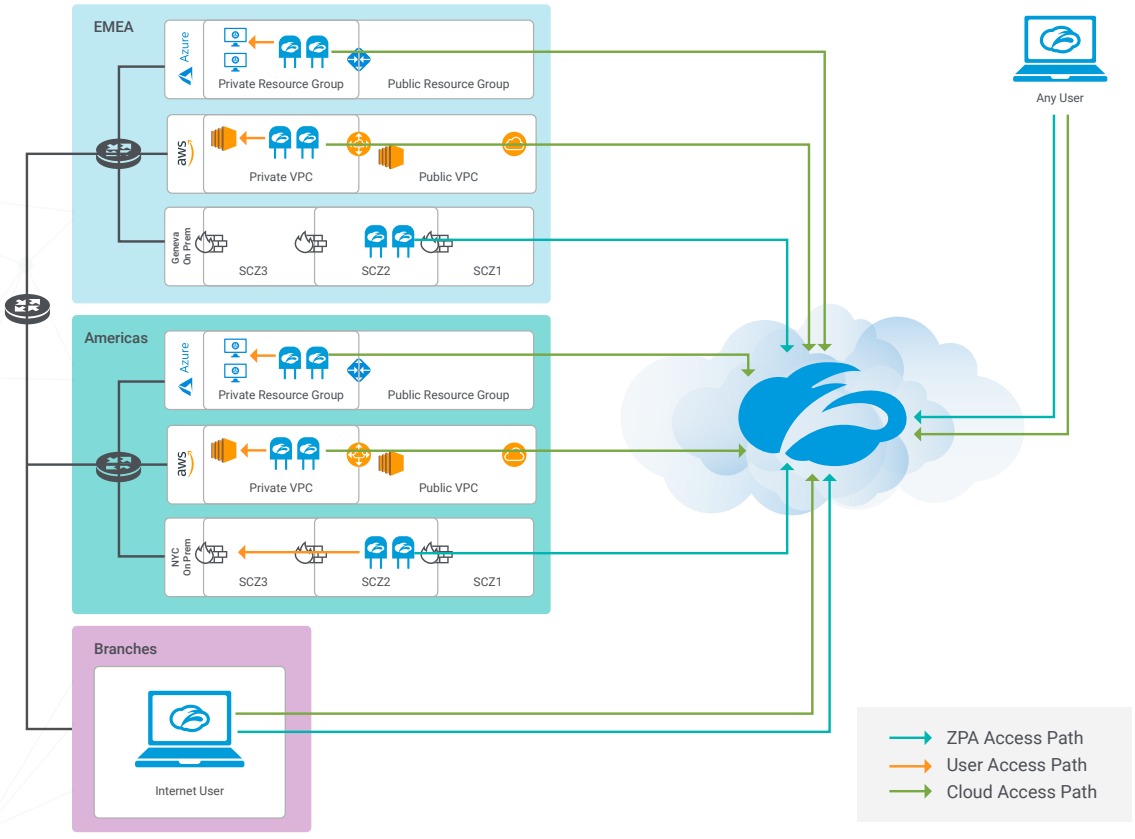
Once the Zscaler service is in production within the main app locations within the Geo-Enterprise, ZPA will be able to simply enable access to all private apps on Azure- and AWS-based locations (ZPA is not limited to web apps only). This can be accomplished by rolling out the ZPA connectors within the necessary Microsoft Azure regions or resource groups, depending on the Geo-Enterprise architecture. By leveraging the default function of ZPA, application access will dynamically find the most direct path for the user to the application, thus steering the Azure and AWS access traffic via the available connector.



Phase III Initial Use Case - Add All Users and Eliminate Local Networks

Once the Zscaler service is in production and application traffic is carried via ZPA for these users, the next phase is to transition from open wildcard access to a more-granular, application-specific access policy. This can be accomplished by identifying critical applications, determining who is currently accessing those applications, and building application-specific policies to control access to the applications. These application-specific policies can co-exist with application discovery, enabling the gradual migration from open access to micro-segmented access.

- Key steps in this migration process include:
- Review discovered apps and sort by priority
- Review diagnostic logs for critical apps to identify which users are accessing what
- Identify policy approach - app-centric, user-centric, or mixed
- Create granular access policies for critical applications
- Decide whether device posture is a required element of access policy
- Repeat process for remaining known applications
- Review unknown applications to determine priority
- Determine whether to retain broad connectivity



The Discovered Applications dashboard provides a cumulative list of all applications accessed by any user over the life of the ZPA instance. A good first step to begin the process of migration from open access to granular policies is to review this list of applications and categorize known apps into critical, medium, or low priority, with Shadow IT apps in a fourth category.

Once the applications are categorized by priority, the diagnostic logs for user activity can be used to generate reports on which users are accessing what applications. These user lists can then be mapped to existing user attributes (group membership, role, organization, etc.), and/or new attributes can be defined to enable granular policy creation for specific user communities.

When defining granular access policies, it can be helpful to think of a conceptual security policy as a statement that has a subject and an object. The user and the app can either be the subject and the object, or vice versa. A user-centric policy takes the user as the subject and identifies what applications that user community is - or is not - permitted to access. An app-centric policy takes the application as the subject and identifies what user communities are - or are not - permitted to access. As existing user access is reviewed, most organizations find that the most flexible approach is a mix of user-centric policies for tightly scoped user communities (such as third-party access) and app-centric policies for resources with clearly defined access requirements (such as regulatory compliance for access to PCI data).

These conceptual security policies can be expressed in ZPA access policies by either mapping a single SAML attribute to a group of applications/application groups, or by mapping multiple SAML attributes to a single application/application group. An example of a user-centric policy would be a ZPA access policy mapping the Contractor group to a specific application or set of applications. An example of an app-centric policy would be a ZPA access policy mapping multiple employee roles to an internal resource, such as SAP. Granular access policies should be constructed above the general connectivity policy to enable a gradual tightening of the access requirements across various applications.

Device posture attributes may be considered in addition to user attributes when defining granular access policies. ZPA posture profiles enable an administrator to restrict access to an application based on whether the user's device is managed or unmanaged, and for managed devices, whether the device is a corporate asset or personal device. The posture profile tests include confirming the drive is encrypted, there is a trusted client certificate, or if the device has been jailbroken, to name a few. These then may be applied broadly to the application or more narrowly to a specific user community accessing an application.

Once granular policies are created to address critical applications and key user communities, a common next step is to repeat this process of review and policy definition for the remaining known applications. As the user-and app-specific access policy set continues to expand, less and less traffic is allowed via the general connectivity policy.

After granular policy is developed for all known applications, the unknown applications can be reviewed - following the same process - to determine whether the application is legitimate, and if so, who should be able to access it.

Eventually, once all traffic is characterized, the decision can be made whether to maintain the access policy allowing general access to all applications (i.e., anything not forbidden is permitted) or change it to block access to any undefined applications (i.e., anything not permitted is forbidden).

Conclusion

With the move to the cloud, network infrastructures will have to change. Organizations of all sizes in all industries from every corner of the globe have to navigate this new landscape of multi-cloud and hybrid cloud while trying to avoid the common mistakes that could doom your cloud adoption. And they also have to ensure that their cloud adoption doesn't outpace their security capabilities. None of this is a challenge that you want to undergo alone. This white paper provided some guidance on how to approach these cloud-adoption projects. But, there is so much more to cover.

As such, Zscaler developed a community specifically for architects working on cloud-first network projects. The Cloud-First Architect is an interactive community that will discuss a range of technical topics to help architects develop the skills needed to design networks that take advantage of cloud infrastructure. Read our blog to learn more about the Cloud-First Architect program, then join the community to discover best practices and deep dive virtual workshops to help you make the most of the cloud.

Zscaler, Inc.
110 Rose Orchard Way
San Jose, CA 95134
+1 408.533.0288
www.zscaler.com