

The Network Architect's Guide to Adopting a Zero Trust Network Access Service

Best practices for using ZTNA
as an alternative to VPN



With private applications moving to cloud and users working remotely, enterprises need a service that can ensure private apps are accessed securely while delivering a frictionless user experience. Even with the buzz around zero trust security some enterprises attempt to use incumbent network-centric architectures, which rely on next-gen firewalls built for access to the network, as a way to now limit user connectivity to applications. These incumbent architectures are a mismatch for today's needs and were not designed to connect authorized users to specific apps. They force users to be placed on-net and often lead to risk of lateral movement to other apps, and IP addresses exposed to the Internet and DDoS attacks via VPN concentrators that sit at the edge of the network and listen for inbound pings.

Many enterprises are considering zero trust network access (ZTNA) services as an alternative to VPN. In fact, Gartner believes that by 2021, 60% of enterprises will phase out their existing VPN for a ZTNA service. But, the reality is that in any large (global) organization even a small change in how users access applications can become a huge task. This document will help you understand where to begin so that you can embrace ZTNA, quickly, and without disruption to the business.

Within this guide we will cover the following:

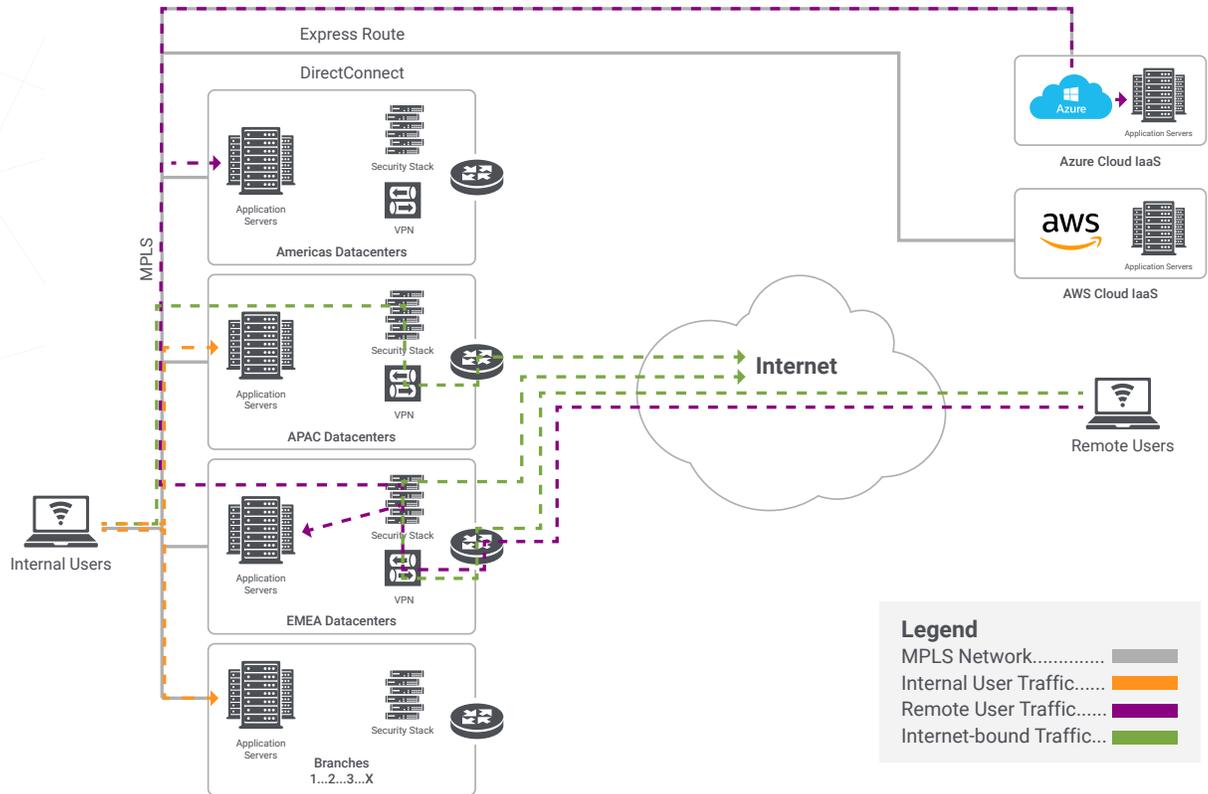
- Architectural differences between incumbent access technology and ZTNA
- A look at a reference architecture for deploying ZTNA
- The three phases to consider when adopting ZTNA within your company
- Pro-tips and considerations for getting the most out of your ZTNA deployment

Before we begin, please take a few moments to read "Mitigating Risk via the Software-defined Perimeter." The blog provides an initial overview of zero trust network access services.

Now, it's time to explore the ZTNA architecture as means of connecting authorized users to specific private applications, without ever placing them on the network.

Where are you today? - Taking a look at VPN within the enterprise

The architecture we are finding to be very common across many organizations can be depicted in this high level diagram. Yes, I realize the number and location of datacenters, routers, firewalls, VPN concentrators, and MPLS network will not be identical to the diagram but believe it provides a close-enough depiction of the components. There are many other network and security devices organizations have deployed, including inline proxy, sandboxes, L7 firewalls, AV and DLP solutions, etc. For sake of simplicity I have consolidated the entire Internet-bound security concept as Security Stack in the diagrams.



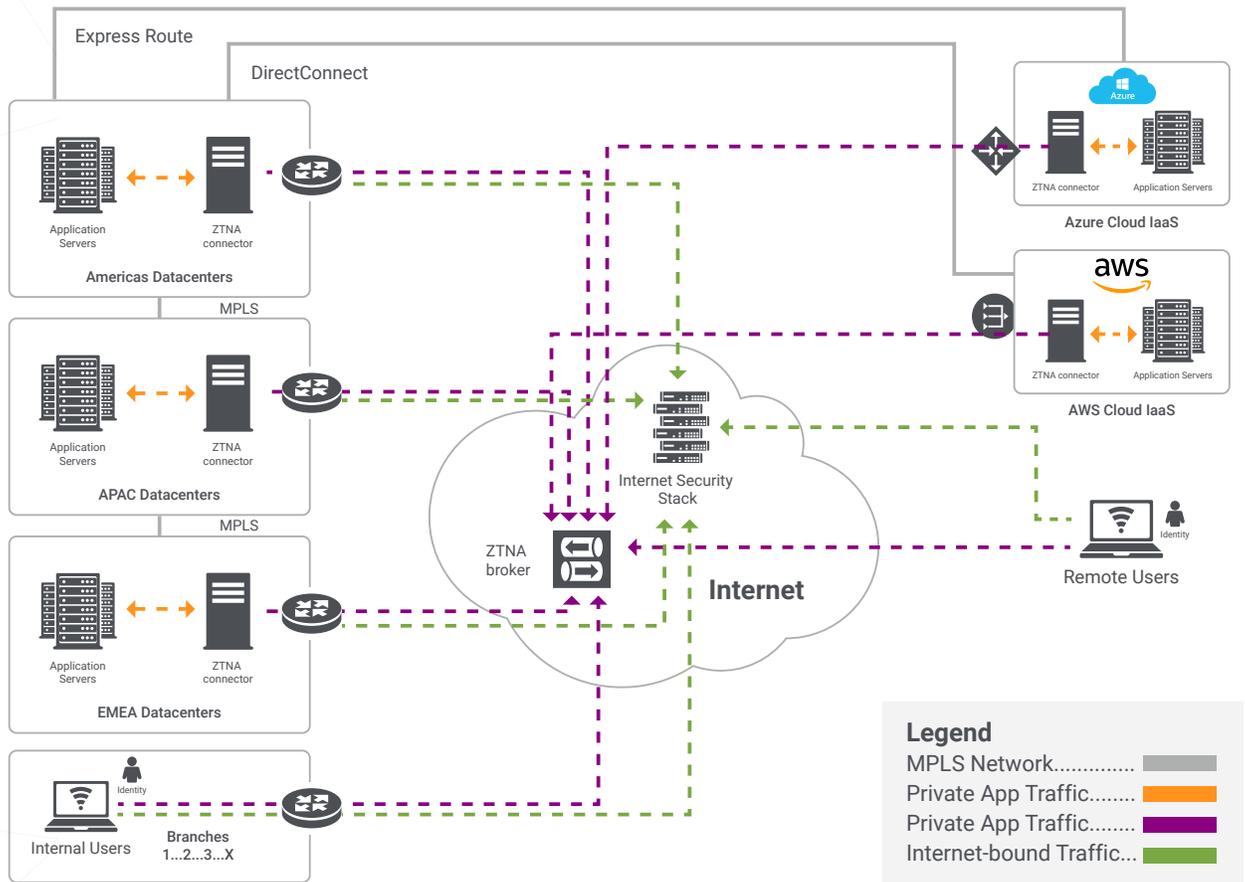
There are a few things I like to point out in this type of traditional architecture:

- 01 Remote users will VPN into one of the datacenters and be placed on the corporate network. In my experience with many organizations, the network is relatively flat, ACLs are rather limited, thus, exposing all the corporate datacenter infrastructure and networks to each remote user.
- 02 Remote users will have all Internet-bound traffic backhaul to the datacenter for inspection of using the (hardware) Security Stack the organization owns. This is known as a Full Tunnel VPN; ideal for security teams that need to ensure users are safe when off the corporate network, but can negatively impact user experience when all Internet/SaaS applications are backhauled instead of locally egressing. Many users now have home broadband Internet connections that are faster than some corporate WAN pipes (Even in somewhat rural Tennessee, I have a 1Gbps fiber connection through my ISP)!
- 03 Internal users are typically on device/user networks, whether physical or wireless, but are still typically able to route/connect to all the datacenter networks since these networks are traditionally "trusted". Access to internal applications routes over the LAN, and Internet/SaaS applications will go through the Security Stack before egressing to the ISP. The problem with this situation is the incorrect assumption that just because you "own" and control the network, that you should automatically trust all users and devices on it.

Take note of the inbound access required from the internet (VPN) for remote access, and the ability for internal users to directly communicate with all application servers regardless of identity.

A reference architecture for providing access to internal apps without increasing risk

The end goal with a software-defined architecture is to decouple application access from network access. Users will no longer need to be placed on the network, private applications are only accessible to authorized users, IP addresses never exposed to the internet and the complexity of managing network segments, FW policies and ACLs goes away. The following diagram shows simplified look at the end result.



With this new software-defined architecture, you will notice that there is a clear separation of datacenter/application networks, remote users, and internal users. It doesn't matter if your organization has just two US-based datacenters, a dozen global datacenters, some Azure/AWS/GCP environments, etc... the results are pretty straight-forward:

01

Private networking, such as MPLS or even site-to-site VPNs, should only be needed between datacenters and Cloud IaaS environments where server to server communication is required. If your organization has moved the www site web tier to AWS but the backend SQL database is still in a physical datacenter, you still require private connectivity (low latency, high bandwidth) between those locations.

02

Remote access no longer requires inbound connectivity for users, such as `vpn.company.com`! This architecture puts the orchestration (control) plane in the cloud where communication from users is terminated. The gateways, known in the Zscaler world as ZPA App Connectors, do not require inbound listening ports, a public IP/DNS record. These Connectors communicate outbound over TLS to the SaaS-based orchestration plane. Internal applications are only brokered once a user's identity has been verified and matched against access policies.

- If a user is allowed to access an internal application/resources, then the orchestration plane stitches the outbound TLS connections together between the Connectors and the user devices. This user is not placed on the network however, so DNS based applications are obfuscated; meaning that the true private IP addresses of application servers are not exposed to the user devices. Rather, a synthetic IP address is dynamically created on the client for each app they are accessing.
- If a user is not allowed access to an internal application, there will never have been any network traffic generated coming into the datacenter. The request would be blocked in the cloud, thus eliminating the risk of even letting users get to the "front door" of critical application servers. The easiest way to look at this is to cut users off at the cloud prior to being able to establish an SSH or RDP session to a server. Even though the user would most likely not be able to authenticate the SSH/RDP session (aside from brute force or stolen credentials), this architecture eliminates this risk. The best part? Each of these attempts are logged allow your security organization to proactively (and reactively) monitor what users are attempting to do. One example would be to send all the logs to your SIEM, such as Splunk, and create an alert if any user generates X number of blocked policies in X number of minutes on the same servers/ports; such as trying to SSH into `sap.company.com` 20 times in 5 minutes. If the user is blocked via policy then you are safe and can proactively reach out to see if the user's device is compromised or if the user had malicious intent. If the user was not blocked via policy then the SSH sessions would have been brokered but the server was rejecting incorrect credentials; meaning this user was authorized but forgot the admin (root) password!

03

All user networks should be treated as Internet-cafes or guest WiFi networks. Whether the user is on the main campus at HQ, at a branch office, in a manufacturing plant, or simply traveling, there should never be a reason to put the user onto the network where they can explore/route to your application servers and datacenters. It is important to note that some branch sites may have requirements outside the user to app access. In such a case, IoT devices and server to server type communications would still need private network connectivity. However, even if such requirement exists it is best to separate these networks from the user networks.

04

Internet access, aka the Security Stack, should also be modernized to allow for the best security and user experience. When decoupling the users from the network, you should explore sending Internet traffic directly from users instead of to centralized datacenter for inspection. For branch offices, it can be as simple as using an existing router, firewall or SD-WAN device to steer all Internet-bound traffic to a cloud security solution, such as the Zscaler Internet Access platform. The full Security Stack is offered as a service, and with over 100 global locations, it means you can send all corporate locations to the closest Zscaler sites for inspection! Even if a user travels, the unified Zscaler App client, a lightweight forwarding agent deployed on user mobile devices and laptops, can provide the user experience needed (sending Internet traffic locally to the closest Zscaler node instead of backhauling), but still provide the IT team with the security controls and visibility needed.

Three phases to make adopting an ZTNA Architecture possible

Architects often ask “what is the best way to get started?” One of my favorites answers is “it depends”. I know many engineers and architects can relate to this because there are many outcomes that can be achieved based on specific needs, requirements, and configuration. However, it is our responsibility to provide the best practice recommendations with organizations on this journey. This is my notice that the phased journey approach discussed in this section is not a concrete set of steps each organization must follow. It is a high level approach we have found in many engagements to meet the current requirements, but at the same time enable the organization to embrace the zero trust networking concept. Trust is never implied and access adaptive, based on contextual policies set by admins - user, device, service etc.

The approach is almost like baby steps: start with remote users, develop segments, and then leverage ZTNA for access to private apps for all users, regardless of location. You'll need to consider the way users access applications and services, the distribution (quantities and types) of your locations (datacenters, Cloud IaaS environments, and physical locations where employees work from), and any project-based timeline. In many cases a VPN refresh could serve as a catalyst to embrace ZTNA rather than purchase a next-gen or “always-on” VPN which bring with it the same challenges of your VPN today.

Phase 1 Getting the ZTNA deployed for remote access and application discovery

In this phase, you will want to begin by replacing the existing remote access VPN solution. To do this you may have to deploy the ZTNA with similar access levels of your current remote access VPN. This is key as you want to make sure that your new initiative is not viewed as an inhibitor of remote user productivity.

You will also need to understand what private apps are running in your environment to reduce your attack surface and root out Shadow IT. There is a good chance there are far more apps than you are currently aware of. Our solution called Zscaler Private Access (ZPA) solves this with our Application Discovery feature. It is impossible to know every internal application/service each user needs to access, so Application Discovery allows you to essentially do wildcards, such as *.company.com, *.company.net, all TCP and UDP ports.

Once a user has successfully enrolled with the service, the client automatically detects when the user is no longer on the corporate network; all internal applications will now flow through the ZTNA when the user is off network. There is no longer a need to launch a VPN client and the user can access internal resources just like before. All these access logs are found in the ZPA admin console and can also be streamed in near real-time to your SIEM of choice, allowing for granular visibility into what applications are being accessed by users.

Add Application Segment
✕

1 Define Applications
2 Segment Group
3 Server Groups
4 Servers
5 Review
6 Policies

GENERAL INFORMATION

Name
Application Discovery Status
 Enabled Disabled

Description

APPLICATIONS

> *.companyintranet.com	<input type="checkbox"/> Browser Access	Add More
> *.oldcompanyintranet.net	<input type="checkbox"/> Browser Access	Remove

ZSCALER APP ACCESS

TCP Port Ranges

1	65535	Add More
---	-------	--------------------------

UDP Port Ranges

1	65536	Add More
---	-------	--------------------------

ADDITIONAL CONFIGURATION

Double Encryption
 Enabled Disabled

Bypass
On Corporate Network

Add Access Policy
✕

Name
Allow Employees App Discovery

Description

Action
Allow

SAML Attribute
Group Memberships Domain Users [+](#)

Posture Profiles
(Optional) Choose posture profiles

Message to User

Application Segments
[Choose Application Segments](#)

Segment Groups
✕ Application Discovery

Save
Cancel

As the internal private network (MPLS, site to site VPN) most likely still exist, the Zscaler App client will automatically turn ZPA off when the user comes back to the corporate network. Now, all access to internal applications happens on the LAN without Zscaler in the path.

Phase 2 Leverage micro-segmentation to ensure least-privilege connectivity

In this phase you will need to define policies that separate private applications into segments, and provide access to these segments via user identity attributes

As large organizations may have hundreds or thousands of unique applications/services, many organizations may want to start with segmenting out management ports, such as TCP 22 (SSH) and TCP/UDP 3389 (RDP), and only provide access to these ports globally for IT users. There can always be one off requirements of course, but this segmentation can help reduce the surface of users connecting to servers they should not be able to even get to. For example, your sales folks most likely should not be able to get to TCP 3389 on a Windows Server that is hosting your SAP application; they should only be getting to the front-end web portion which would be the same servers but just on ports TCP 80/443.

Ideally, infrastructure servers which can be domain controllers/services, security software clients, software deployment clients, etc-- these can be easily segmented as the hosts are well known.

App Segmentation is an ongoing process, with a general recommendation being to prioritize applications that are most critical to the business and should only be accessed by known user types.

As you segment applications, they are removed from the "pool" of application discovery. This means that you can mix and match, to ensure that users can still get to applications in your domains you haven't explicitly defined, but can also get to known applications on the service ports required.

NOTE: Don't forget about Internet Security too

Although we are focusing on the private applications in this guide, it's important to realize that it is equally important to provide a Security Stack for all Internet-bound traffic as well. Many organizations are exploring a more modern inbound and outbound security stack that is fully cloud-based instead of relying on appliances or virtual appliances (i.e firewalls). At Zscaler our outbound cloud security solution is called Zscaler Internet Access (ZIA).

Phase 3 ZTNA for access to private apps for all users (not just remote ones)

You are now ready for the final phase. This means that going forward all access to private applications is based on precise settings that enable explicit, least-privilege connectivity only, by default.

ZPA provides this through inside-out connectivity via TLS double-encrypted micro-tunnels that are spun up on a per session basis and create a secure segment of one between an authorized user and a specific private app.

You may recall that I mentioned earlier how Zscaler App can detect the corporate network, right? This means that in ZPA, each application segment has a configuration option to (1) bypass ZPA when on the corporate network, (2) always bypass ZPA, or (3) never bypass ZPA. In phase 1 you deployed app segments using option 1, but what about secure access from not only remote users, but all? To do this you can simply switch the app segments to never bypass ZPA. This means that even when the users are in a physical office, all access to internal resources is brokered through this explicit trust architecture solution- never just routed on the LAN directly to the application servers in your datacenter!

Go from**Bypass**On Corporate Network **to****Bypass**Never 

Easy, right? Well I think the challenges involved with this switch remain outside of our platform itself. As you may recall the end goal is typically to completely remove the application server/datacenter networks from all user networks. This means no connectivity from any branch office, manufacturing plant, etc (to be clear I mean no connectivity from the USER networks at these locations) to the datacenter.

Final thoughts and pro-tips:

It might be easiest to start with a new small office that isn't on the network yet. Open that office with just a broadband Internet connection. Have all internet-bound traffic go to a cloud security platform (such as ZIA) and have any private application traffic flow through the ZPA platform.

Treat the new office like you are opening an Internet cafe! Again, keep in mind that today we are able to provide this connectivity for users to applications; some locations such as a manufacturing plant with sensors, IoT devices, and servers, will still likely need to communicate with your datacenters over private MPLS or VPN. Treat those location networks as datacenters, and just remote the users from them; all users will be on "guest WiFi" and internal app access will be brokered to authorized users.

In the end, there is a lot of excitement and buzz around ZTNA architectures, but the real goal is to deliver the experience users want, with the security it needs when it comes to private apps. It will take time for your organization to adopt this new method but you, as the network architect, can lay the foundation (platforms) that will enable it.

You can experience ZPA for yourself by signing up for a 7-day hosted trial at <https://www.zscaler.com/zpa-interactive>.

