

Zero Trust at Scale

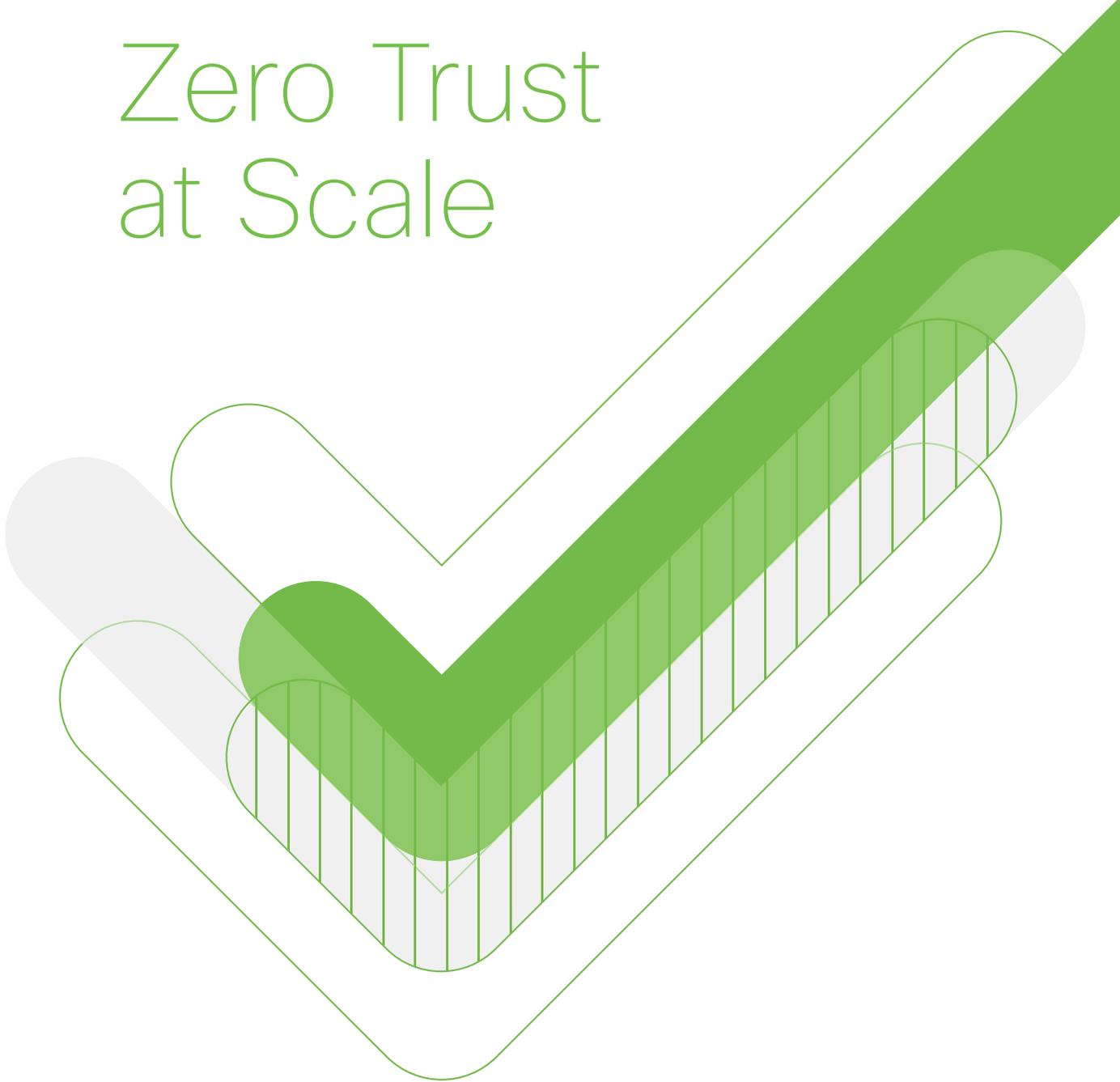


Table of Contents

Introduction	3
Why Zero Trust?	4
Why is this ambitious?	4
The Plan	5
Getting buy-in	5
Core team	6
Technical goals	7
Zero Trust Architecture	8
Setting up	8
SaaS apps	9
On-premises	9
Timeline	10
Month 1 (July–August)	10
Month 2 (September)	11
Month 3–4 (October–November)	11
Month 5 (December)	11
Lessons Learned/Best Practices	11
Team approach	11
Executive sponsorship	11
Pilot for proof	11
Create demand for Zero Trust	12
Full transparency and regular communication	12
Cisco Today	13
Deployment	13
Future of Zero Trust at Cisco	13

Introduction

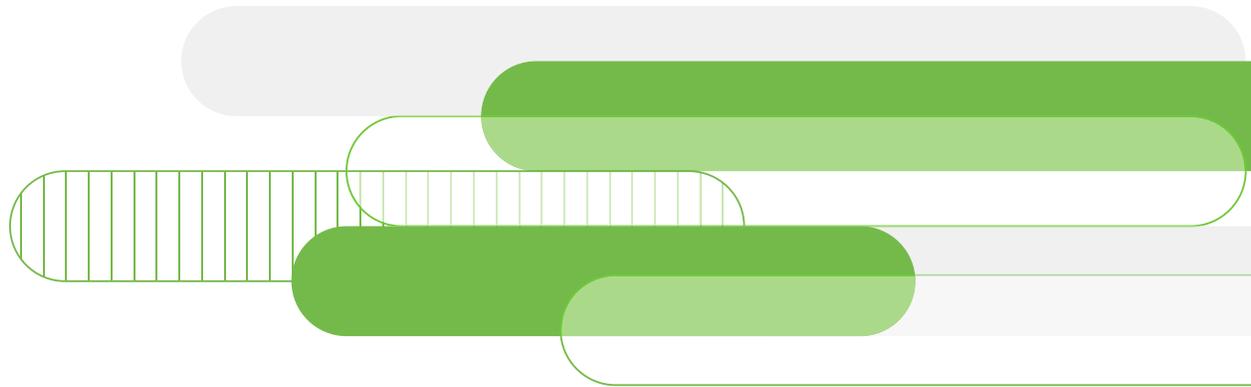
In 2020, Cisco set out to move from a traditional network-based perimeter and VPN model to a Zero Trust framework. Dubbed ‘borderless’ internally, the core goal was to give users a secure, uniform experience accessing applications, wherever the user or application is located. Using the features of Duo Beyond, our team set out to improve security and create a better experience for our 100,000+ users – a fundamental shift that took place in less than five months.

What does Zero Trust at Cisco look like?

When we think about Zero Trust at Cisco, four things need to happen every time someone tries to access an application:

1. We verify the user
2. We confirm that the device is up-to-date and healthy
3. We validate that a Cisco-managed device is being used
4. The application can be accessed without the VPN

And when we say every time, we mean it. Not once a day, not for one application, but continuously.



Why Zero Trust?

Let's first take a step back and talk about why Zero Trust is important. A modern security solution is essential to keeping up with the evolving threat landscape.

Identity	Application	Environment
<p>According to the 2022 Data Breach Investigation Report by Verizon, 82% of breaches involved the 'human element' – either via the use of stolen credentials, phishing scams (to steal credentials), or errors and misuse of resources.</p> <p>This suggests that passwords remain a target for attackers, making secure access easy for users is an effective way to reduce risk.¹</p>	<p>More than 50% of all global web sites were vulnerable to at least one serious exploitable vulnerability throughout the entire year.²</p> <p>The rising ubiquity of SaaS and the shift of valuable data to the cloud, including email accounts and business-related processes, means unsecured applications can be a hidden invite for exploitation.</p>	<p>IDC predicts there will be almost 56 billion connected devices worldwide by 2025, with 75% connected to an IoT platform.</p> <p>These smart devices, while useful, can be a risky target for attackers to access your network.</p>

1: Source: <https://www.verizon.com/business/resources/reports/dbir/>

2: Source: <https://www.whitehatsec.com/blog/appsec-stats-flash-2021-year-in-review/>

Additionally, we know that vulnerabilities and their exploitation continue to be the root causes of most information security breaches today. While zero days receive a lot of attention, they historically account for only 0.4% of vulnerabilities. In fact, vulnerabilities unattended and unaddressed for months and even years are the most common vectors. Many of these can be avoided simply by keeping devices current with software patches and operating system updates. Therefore, validating device health (i.e., is it up-to-date and patched) becomes very important when making a decision about whether or not to allow access to a resource.

With the Zero Trust model, you gain better visibility across your users, devices and applications because you verify their security states with every access request. It isn't an either-or situation – it's equally important to ensure that both users and devices accessing applications meet your security requirements.

Why is this ambitious?

When most of our workforce went remote, it was confusing and taxing for users to know how to access different applications. For example, some apps required a VPN, while others could be accessed directly, causing a lot of frustration for our users.

Historically, organizations haven't spent a lot of time worrying about applications that require a VPN connection versus those that don't.. We know that IT leaders are bearing part of this burden – they're finding themselves having to manage laptops and VPNs for an increasingly remote workforce. Like many other companies, Cisco invested in VPN expansion in order to support employees working from home. In the Asia Pacific, Japan, and China region, we expanded the capacity including bandwidth and IP pools. In Europe, Middle East, and Africa, we found our overall capacity was good but needed additional resiliency. In the Americas, we increased resources at our San Jose campus and Research Triangle Park, North Carolina. We also made necessary changes to our VPN access points to automatically redirect and globally distribute traffic as needed.

However, there are other factors that we need to consider for a flexible workforce. The line between work and home life is blurring for many people. Family members may share computers for small tasks including schoolwork. It's also normal to expect that people want to keep their personal browsing and work separate from the corporate VPN. Switching VPN on and off multiple times can be frustrating. In addition, using a VPN when the workforce is almost fully remote can be inefficient, especially when we're sending data back over the corporate network only to have to eventually go back to the cloud. It can also be tiring for users to keep track of which applications need VPN and which don't, ultimately having a negative impact on their productivity.

This was the start of transforming how we work. We hear about Zero Trust everywhere these days, and really, it means something different for everyone. At Cisco, we've been invested in applying Zero Trust principles across the organization in all facets of our infrastructure processes.

To enable this new way of work, we needed a new approach. We decided to apply Zero Trust principles across the organization so that users could work anywhere, from any device, and do so securely without friction.

The Plan

Getting buy-in

The first step Cisco IT took was developing a plan to get executive buy-in. One common mistake we saw with other initiatives was that high complexity made them hard to understand and sponsor. Our goal was to make the message simple, specific and timebound, so it would be memorable and easy to repeat to others.

When we kicked off the program and met with our executive sponsors, we focused on three key benefits:

1. **Improved Authentication Experience:** We often heard users say, "I feel like I'm authenticating all the time." We set out to address that concern twofold, by enabling an experience less reliant on passwords, and by starting to apply Zero Trust principles to remote access use cases, removing the implicit trust granted by the traditional perimeter.
2. **Borderless Application Access:** We know that you can't implicitly trust a user simply because they're on the network. By the same token, if we are able to establish trust, then we can allow access to some of these applications or resources for which you traditionally require VPN access. In doing so, you provide a more seamless experience to users accessing these applications without having to be aware whether they're on VPN, the corporate network or elsewhere.
3. **Increased Security:** Access is not just based on the user and whether we trust the user because they're providing their username and password. We're also able to say, "Does the endpoint they're coming from meet a certain security posture? Is it healthy?" Duo gives us increased security and visibility to restrict access based on that level of context.

Challenges	Solutions	Results
<ul style="list-style-type: none"> • Protect worldwide access • Expand access to users securely • Secure all users and devices • Consistent experience for users 	<ul style="list-style-type: none"> • MFA prevents fraudulent login attempts • Trusted Endpoints limits access to managed devices • Trust Monitor detects abnormal login attempts • Duo Device Health app ensures devices are safe • Adaptive access policies block risky login attempts • Duo SSO simplifies access with one username and password • Duo Network Gateway provides access without a VPN 	<ul style="list-style-type: none"> • 100,000+ users and 170,000+ devices secured • Deployed in 5 months • <1% of users contacted help desk • 5.76 million health checks per month • 86,000 devices per month remediated • 410,000 fewer auths per month via DNG (VPN-less)

Executive sponsorship was a pretty large departure from how things were done in the past. Instead of waiting until the project was fully planned out before sharing it, stakeholders were brought in from the beginning. The project was sponsored by multiple executives from security and IT who could get buy-in at their level. They also provided air cover in case the team needed to move quickly and inadvertently broke things. While speed was deemed more important than perfection, everything actually worked out as intended.

The core team consisted of a representative with decision-making power from:

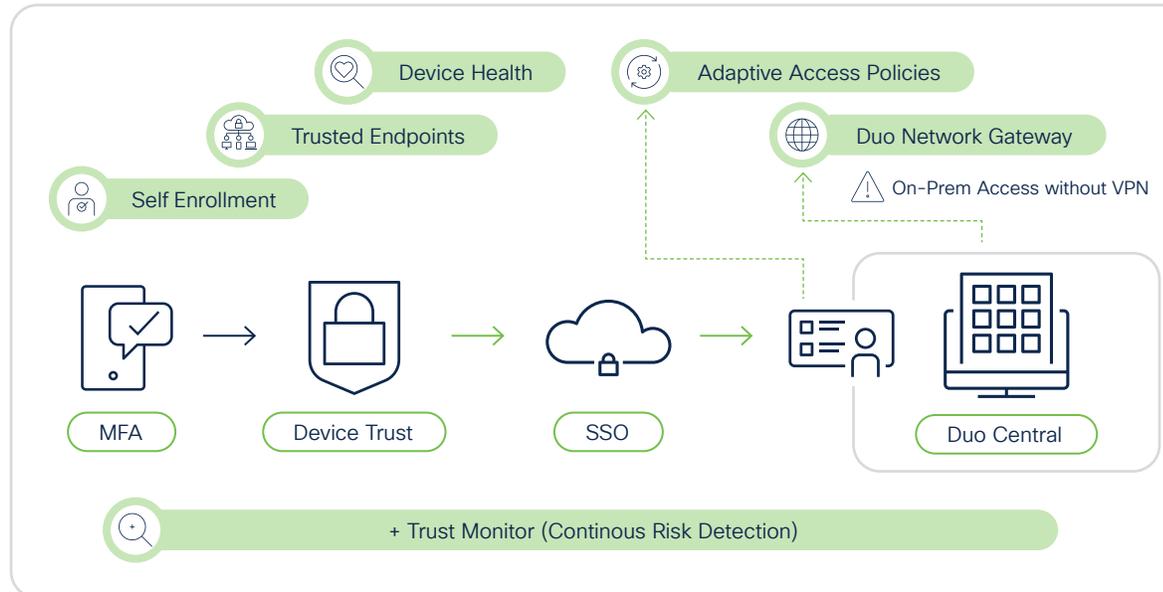
- Infrastructure
- Endpoint
- Identity
- App Trackers
- Comms
- Information Security

Core team

The team behind this expansive transformation was more than just one department or task force. Given the aggressive timelines, we decided from the beginning that we'd have a small core team driving the initiative with a representative from each function. This small but mighty group of engineers and architects were a powerful influence on how we moved forward. They were tasked with going back to the respective organizations and either sharing information, soliciting feedback, or asking them to move forward with various decisions made. Executive sponsorship provided reinforcement and helped remove barriers when needed.

Taking time early on to plan the project scope, measure and affirm milestones along the way, and identify core team members helped considerably. Having several specialized teams allowed us to break up the deployment into multiple workstreams, with each function free to focus on their tasks. Weekly meetings were arranged to share updates, with one person from each workstream present – this kept the team small and efficient.

Technical goals



From a technical perspective, our primary goal was to implement an architecture that would allow secure, VPN-free access to some of our most-visited internal and SaaS applications. Secondly, we wanted to validate user and device trust, doing this on a per-app basis with an ability to set per-app access policies. Third, we aimed to improve our authentication experience by reducing the burden on users. Finally, we wanted to build this transition seamlessly, requiring zero user action, and without any outages or distractions. The core thought behind zero trust is to validate user and device trust every time a resource is accessed, rather than implicitly trusting the device simply because it's connected to your corporate network. What does this actually mean?

First, we need a way to verify user trust. Historically, this was achieved with a combination of username and password, and more recently with a second authentication factor. If you recall, one of our objectives with this program was to reduce the friction for users when it came to authentication.

To achieve this, we leveraged a user-identity certificate securely deployed to our managed endpoints by our device management suite. This certificate then acts as the first factor of authentication, saving the user the step of having to type in their username and password. This also reduces the likelihood that a user will simply save their corporate identity and password in their browser for convenience. We want to discourage this behavior to help prevent third parties from gaining access to our user credentials. Plus, by employing the username-password form less often, users indirectly become more vigilant against phishing. Next, the user completes a second factor of authentication. For most users, this means acknowledging a push notification sent to their phone on the Duo Mobile app.

After establishing user trust, we need a way to validate device trust and health. This is a new area that the Zero Trust architecture model aims to address. Here at Cisco, we start with the assumption that if a device is managed by our corporate device management platforms, then it must have a good baseline security posture. This includes guarantees like the latest software patches, screen lock, firewall and encryption enabled. Therefore, to identify a trusted device within our Zero Trust framework, we use device certificates that are pushed to each endpoint by our device management platform.

Finally, we go one step further to verify device trust. While the device may well have been managed at some point when the device certificate was first deployed, it's possible that a particularly tech-savvy user, which we have plenty of at Cisco, could have disabled some of the protections that our management platform enforces. To mitigate this, we do an additional device health check during every authentication transaction to ensure that the device still has the latest software, screen lock, disk encryption, firewall, and anti-virus agent running. This real-time health check is conducted by Duo's Device Health app, which continuously operates on the device in the background.

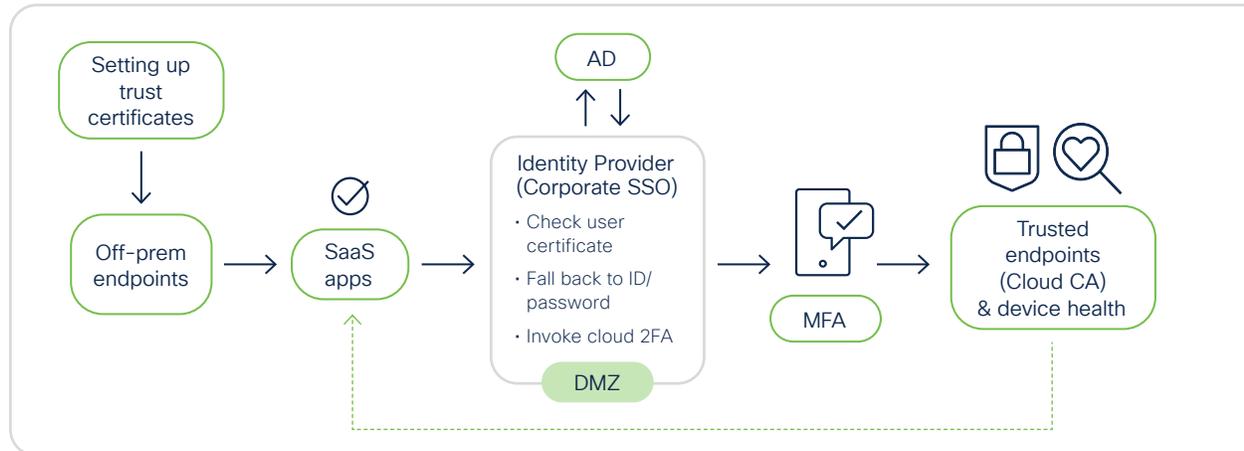
Putting this all together, when a user tries to log into an application, our corporate SSO identity engine checks the user and device certificate, does a real-time health assessment of the device, and finally triggers a second-factor notification before allowing the user access. We do this regardless of whether the application is on-premise, behind the corporate firewall, or a SaaS app. Now that we've seen the high-level components of this architecture, let's take a closer look at our end-to-end architecture.

Zero Trust Architecture

Setting up

The first piece of this architecture is getting the endpoints set up properly. We have a fleet of more than 170,000 endpoints, which includes a near-even split of Mac and Windows, and more than 60,000 mobile-class endpoints, most of which are iOS but also with a sizable number of Android users. We leveraged our device management platforms to deploy the user and device trust certificates to endpoints. We also pushed the Device Health app to user desktops. Finally, we deployed configurations for the main browser types to select the right certificate automatically, rather than prompting the user to choose. We use Jamf for managing our Mac fleet, SCCM and Intune for managing our Windows, and Cisco Meraki for managing our mobile endpoints.

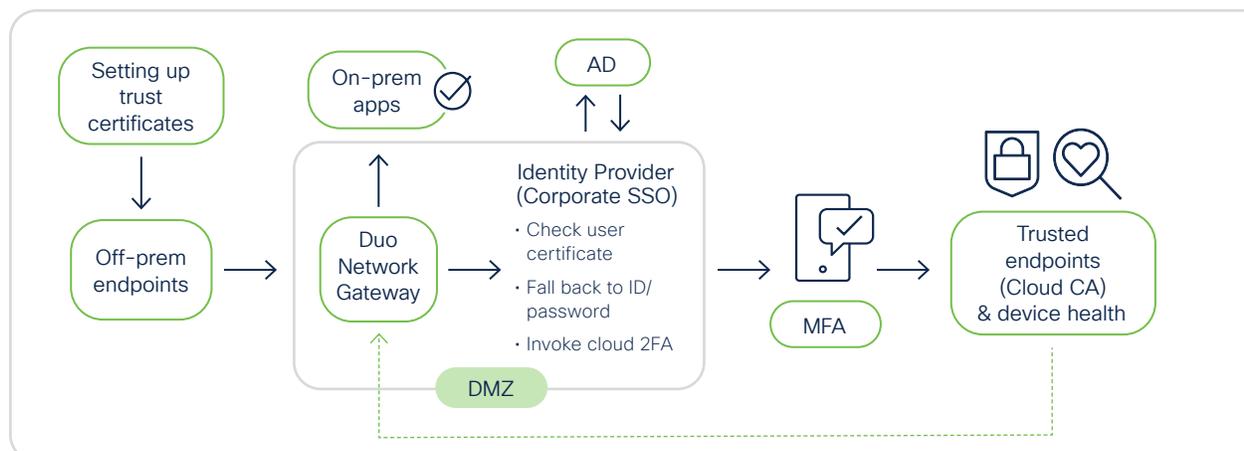
SaaS apps



The architecture for SaaS apps starts with the endpoint requesting a cloud web application. This starts a SAML handshake with the identity provider (IdP). Our IdP tries to retrieve the user identity certificate from the endpoint, and upon receiving a valid certificate it verifies that the user is still an active employee by looking up their information in Active Directory. Once all these checks complete, the user is prompted to complete a second authentication factor, usually via a push notification from the Duo Mobile app. We also support other authentication methods like TouchID and YubiKeys.

Having verified user identity, we proceed to validate device trust. This starts with Duo polling the device for a device trust certificate to confirm it belongs to the same user. After successful certificate validation, Duo invokes the Device Health app running on the endpoint, which reports its health status in real time. If no issues are reported, the authentication flow is completed and a SAML assertion is issued by our IdP to the SaaS application. If a device is out of compliance, the app tells the user how to update and fix the issue. The user is then permitted access to the application.

On-premises



The flow slightly differs for users accessing an on-premises application. A key component making this possible is the Duo Network Gateway, a reverse proxy that acts as the main control point for allowing VPN-free access. When a user tries to request an on-premises application in their browser (for example, our intranet website), the first thing that happens is our DNS servers return the IP address of the reverse proxy instead of the actual application host. This mechanism is commonly referred to as split DNS.

After the browser is redirected to the Network Gateway, it initiates the same authentication flow previously described for SaaS applications. The user identity is validated first, and then the device trust is checked. If both succeed, then relevant session cookies are issued and the user's browser is proxied through to the application.

Timeline

As mentioned previously, this was a five-month program with clearly defined goals, commitments and deadlines.



Month 1 (July - August)

We dedicated the first month or so to conducting proofs of concept with dummy applications. This was the first time we tested the entire architecture end-to-end on all our device platforms.

One of the learnings from this exercise was that getting consistent browser behavior across four different OS platforms is trickier than you might think. Even though the browser configurations for automatic cert selection were well documented, we found lots of edge cases – especially on the Windows side of things – where our policies conflicted with other enterprise group policies, and cert collection and selection didn't always work as expected. In fact, it's safe to say that getting certificates deployed to the endpoints, and then having our identity engine collect those certificates consistently required the most iterations before we had it working like we wanted.

During this period, we also examined our web and VPN access logs across multiple systems to empirically identify our most-used web applications, both SaaS and on-premises. This step was important because it allowed us to focus on the big wins instead of trying to boil the ocean. We narrowed our list of apps to about 20 and then started engaging with the application teams. In some cases, we had to do a security audit of the selected application to ensure it was meeting all our enterprise security standards before opening it up for remote access.

This first month was also when we had to get buy-in on the architecture. This involved approval from our InfoSec, Endpoint, and Infrastructure teams. After completing this step, we started preparing our production infrastructure.

Month 2 (September)

The next big milestone was our product deployment of the Duo Network Gateway, a prerequisite for making on-premise applications available via remote access. After this step, the entire setup was unofficially live. Any device with the right certificate and trust would be able to access the included onprem applications without having to use VPN.

Month 3-4 (October - November)

We started with a small-scale pilot of about 2,000 users within our IT organization. This meant deploying the endpoint configurations and trust certificates to this group and having them try out both the onprem and SaaS authentication flows. The feedback was mostly positive, which allowed us to proceed on time with our planned first wave program scope of roughly 35,000 users spread across four different organizations within Cisco.

Month 5 (December)

The time invested during our proof of concept and pilot phase really paid off. Our initial deployment to 35,000 users was so successful that we proceeded to do a full enterprise rollout two months sooner than planned. Since then, we've continued onboarding more applications to the Zero Trust program, and our users are loving it.

Lessons Learned/Best Practices

Team approach

A core team made up of one representative from each workstream was empowered to make decisions for their organization. The core team met on a regular basis for updates, which helped keep people focused on what they needed to do and be agile in decision-making. There were times in the project where we had to deviate from the project timeline and either delay or accelerate some efforts – but by having a lean team, we were able to agree on the adjustments quickly and effectively.

Executive sponsorship

Both our CIO and Chief Security and Trust Officer offered their full support for marching ahead with our Zero Trust initiative. Our transformation project required efforts across IT and security. Ensuring there was alignment from our executives empowered our team to make decisions and move quickly. There was no doubt that this was an important initiative across the organization, and leadership support helped prioritize it.

Pilot for proof

A phased rollout helped keep Zero Trust both top-of-mind and manageable. Rather than roll out to all apps and all users at once, the team started with a subset of apps and departments. That helped to prove out the process, identify any issues, and fix them. As a result, the full rollout was able to go live months ahead of schedule.

Our approach continues to be gradually introducing Zero Trust principles across the organization. Instead of making drastic changes that will impact productivity and generate chaos, we introduced Zero Trust for the workforce, but we focused on the remote worker use case, offering a new borderless method of access.

Create demand for Zero Trust

We took a methodical approach to introduce Zero Trust principles across our organization. The team started by enabling all users with the technology of Duo and then added apps to the Zero Trust architecture over time. This helped build demand for the program and made the process effortless for users. To help with prioritization, we invited users to nominate applications to include in the program. Additionally, we developed a process for application owners to request Zero Trust for their apps.

When socializing with our workforce, we emphasized, at its core, this was an initiative about improving the overall security posture at Cisco. We also explained that the experience would benefit them by providing frictionless secure access if they had a device validated as managed and healthy. Improving user experience while increasing security is something that's rarely achieved. We believe, by applying Zero Trust principles, you can actually achieve that.

Full transparency and regular communication

We can't overstate the importance of actively communicating to your stakeholders, leadership and employees. We had several channels to keep people in the loop. A weekly update newsletter went out to anyone who signed up for it, a SharePoint site was created to explain what was happening, forums were available for public comments as well as emails, articles, and guidance for leadership communication.

Once we agreed on when to go live with our project, we decided to send our stakeholders, senior leadership and interested parties a weekly newsletter which prominently showed how many days remained before our project go-live date. Consisting of accomplishments performed that week, along with open issues, the newsletter served several purposes. First, it kept our teams accountable. We felt a sense of urgency to ensure the project remained in the green status. Next, it gave us a single platform to inform stakeholders about challenges and issues. After sharing challenges for a given week, it was pretty common for a leader to reach out and ask how they could help resolve any of the roadblocks we were facing.

Being engaged with our employees was also very important. While you may try to provide the best experience to your employees, the reality is that you will negatively impact them, or at least some users, along the way. We created several forums, such as our SharePoint site and team alias, where our users can provide feedback. We also met with our support organization on a weekly basis to better understand if there are areas we need to address, whether it's end-user documentation, support training opportunities, or simply awareness of new issues cropping up.

Cisco Today

Deployment

- Deployed Duo configurations to more than 180,000 endpoints, including our entire fleet of Cisco IT-managed iOS, Android, Mac and Windows devices – and that number continues to grow as all newly enrolled devices automatically receive the necessary configurations.
- Introduced this new borderless method of access for more than 120 key corporate applications. Since announcing the availability of borderless access, we've received an overwhelming response to enable other applications.
- Engaged with app owners on how their users can take advantage of this new method of access by providing a documented list of requirements, allowing them to test in staged environments, and offering the form for them to reach out if they have any questions.

Future of Zero Trust at Cisco

Since incorporating controls for device health and trust at the application layer, our ability to react to device risk has substantially improved. For example, there are approximately 5.76 million device health checks automatically conducted per month. This has resulted in finding 86,000 devices in one month that users self-remediated. That's 86,000 potential compromises effortlessly averted.

While there were some concerns of increased support call volume when introducing the device health checks for borderless access, we've seen that less than 1% of our users have contacted the help desk for support. We feel that the easy-to-follow remediation steps within the Duo Device Health App have played a key role in keeping support numbers down.

Now, while we've made significant strides in a short amount of time, our team realized that there's a lot more to do in our Zero Trust journey. We're already planning areas to expand such as streamlining the process for app onboarding. These processes include validating that an application meets login requirements, and making the necessary ACL changes and Duo Network Gateway configurations. Automating this allows us to accelerate the app onboarding process and reduce human error. Finally, we plan to incorporate Duo's upcoming passwordless solution and further simplify and secure the user experience. Cisco IT is also focused on expanding the usage of Zero Trust beyond Cisco's immediate ecosystem to use cases such as the extranet partner landscape and onboarding acquisitions in a more seamless, less infrastructure intensive way.