Zero Trust in Resilient Cloud and Network Architectures

JOSH HALLEY
DHRUMIL PRAJAPATI
ARIEL LEZA
VINAY SAINI

ciscopress.com

Zero Trust in Resilient Cloud and Network Architectures

Josh Halley, CCIEx3 No. 11924

Dhrumil Prajapati, CCIEx2 No. 28071,
CCDE No. 20210002

Ariel Leza

Vinay Saini, CCIE No. 38448,
CWNE No. 69, CCDE No. 20240032

Cisco Press

Bandwidth Planning, Congestion, and Oversubscription

Today's networks do not have the same problems of high congestion, thanks to the cheaper and more accessible high-bandwidth Internet connections. With standardization of the Ethernet across all media and platforms, bandwidth has been increasing exponentially every few years. Whereas 100 Mbps connections were normal in the mid-2000s to 10 Gbps by 2015, in 2024, 400 Gbps connections were normal on most campus and data center switches. Bandwidth evolution has taken QoS out of the picture, and almost all of the switches and routers today support line rate throughput for normal IP traffic. Back in the day, switches were oversubscribed in terms of bandwidth, but with better ASICs and advancement in technologies, oversubscription and congestion are things of the past. QoS only kicks in once there is congestion. If there is no congestion, there is no need for the QoS.

No matter how much enterprises plan to future-proof their network infrastructure, there is an event in the world or industry that changes all of their planning. One such event in 2020 was the global pandemic, when the workforce started to work from home. All of a sudden, all organizations had to come up with a plan to continue their businesses remotely. Whole dynamics of traffic flow and pattern were shifted. There was a high demand in residential Internet traffic, and enterprise Internet and VPN firewalls were overloaded with thousands of employees trying to access their work resources from outside. This situation resulted in the adoption of cloud technologies at a faster rate that gave an ease to the bandwidth requirements. Increasing a 1 Gbps connection to a 10 Gbps connection is not easy because many things need to be changed—from the interface to optics, and in some cases, the fiber itself. However, shifting some of the important workload to the cloud helps in offloading external user traffic to the cloud without increasing local bandwidth or at least giving some breathing room for expansion.

Today, bandwidth planning is crucial. Historical and current traffic patterns and traffic flows are taken into account, and higher access interfaces with lower bandwidth caps are preferred as insurance. For example, most Internet connections at a large site are 10 Gbps access with a 1 Gbps or 2 Gbps bandwidth cap. This ensures that businesses do not pay high costs up front, but in the event that they need more bandwidth, they simply have to ask their service provider to raise the cap.

Network Monitoring and Optimization

Once the traffic analysis is complete, and all management and planning are done, it's time for monitoring and optimization. This is a fairly simple task but still an important one. Monitoring the network and application performance can lead to optimization in the network. As businesses grow, they will have a newer set of applications and requirements. Some may replace legacy on-premises applications with newer cloud- and SaaS-based applications. If an application consisted of about 15 percent enterprise-wide traffic, as the adoption of this SaaS-based application increases, so will the traffic shift. That 15 percent or more traffic will now start shifting toward the Internet or cloud connections. That shift needs to be taken into account, and optimization needs to be addressed.

How the application fails over and the traffic shifts will need to be captured from network monitoring.

The task of monitoring and optimizing the network is an ongoing cycle and needs to be part of an everyday process.

Policy and Security

Among the last components of traffic engineering are policy and security. In the earlier underlay and overlay sections, we discussed secure routing protocols and use of NAC, but at an overall network level, we need to understand how to secure entire networks. There are many physical and logical attributes related to network security. From a device perspective, the following aspects of the network need to be secured:

- **Control Plane Policing:** You need to prevent the device's processor from being subjected to a distributed denial-of-service (DDoS) attack, making it stop forwarding data or slow down convergence.
- **Device Access:** You need to use the right amount of RADIUS or TACACS access with multifactor authentication (MFA) to ensure only authorized users are allowed to access a device.
- **Interface Protection:** We recommend adding an ACL to stop taking inbound connection requests, especially on public Internet-facing interfaces. This prevents inbound sniffing attacks and exploit vulnerabilities.
- **Time of Use Access:** After the devices are set up, they should not be accessed with full privilege access without a change control process. This is true for core or backbone switches because any misconfiguration can take down a large chunk of the network.
- **Security Audits and Firewall Rules:** A regular audit of firewall rules must be warranted to ensure there are no potential holes that can harm the network and overall system.

Global Internet

One of the last bits of the traffic engineering mechanism is the global Internet. Today, with IPv6 being adopted at a faster rate than before, almost all organizations are able to connect directly to Internet service providers (ISPs) and get full Internet routing tables. Direct access to Internet routing tables is good, but this access can also be dangerous if not planned properly. Internet peering is best if done with two or more ISPs. This approach provides protection for the organization's public IP space; in this way, an outage on one ISP will not constitute an outage for the organization. The routing will take care and fail over to the secondary ISP. In planning such architecture, you must make sure that the proper route filtering and policies are in place. An organization does not want to inadvertently become a transit for Internet traffic. Policies need to be in place that advertise only the organization's own prefixes and nothing else. The organization can choose to receive the entire or a partial Internet routing table and traffic-engineer prefixes of one ISP over the other. Planning Internet peering, although it may look simple, needs to be thought out carefully.

Geo-routing

Depending on compliance regulations, sometimes an organization may want to restrict access to its applications to different countries. For example, if a local bank does not have any branches outside of the state and does not offer any internal investment or banking products, it may not want to allow people from different countries to access its banking application. This is done by geo-routing. Since the Internet Assigned Numbers Authority (IANA) is responsible for allocation of IPv4 and IPv6 addresses to all organizations per region and country, the IANA maintains a comprehensive list of IP allocations to all the countries. Organizations can use this list to update their prefix list so that all inbound connections from restricted countries are denied and their networks can be further secured.

Today, geo-routing is also used for streaming content. Video content such as Netflix or YouTube uses geo-routing to publish local available content based on the location of the user. If a user is in India, suggestions would be provided based on that market; the process works similarly for users in the United States with their exclusive content. To avoid geo-routing issues, many users use third-party VPN services to tunnel their traffic through another country and pretend to be from a different region to leverage different content and/or access restricted applications. As networks and technologies advance, there will be mitigations and ways to detect such patterns.

Summary

In this chapter, you learned that routing and traffic engineering are fundamental for scalable, resilient, and secure networks. We looked into underlay routing protocols and their advantages and disadvantages. We also looked at various fabric-based overlay solutions and at traffic engineering and what components are required for traffic engineering to be useful.

References

1. IPv4 CIDR Report: https://www.cidr-report.org/as2.0
2. IPv6 CIDR Report: https://www.cidr-report.org/v6/as2.0

Chapter 9

Authentication and Authorization

In this chapter, you will learn about the following:

- What identity is
- Different types of authentication methods
- Enterprise authentication (dot1x)
- How to monitor authorization of endpoints

Overview

Identity verification is a concept that has been common in society for millennia, with one of the first-known references to an identity document being recorded as early as 450 BC. This reference can be found in the Old Testament's book of Nehemiah, recounting how Nehemiah, a cupbearer to Persian King Artaxerxes I, sought to contribute to rebuilding Jerusalem and asked the king for letters to the governors of the provinces situated west of the Euphrates River, guaranteeing him safe travel to Judah (Neh. 2:7–9). This narrative is an early testament to the use of official documents for secure passage, paralleling today's passports.

King Henry V can likely be credited in this domain in a modern sense, with the first passport used for travel in 1414. Since these times in history, the ability to prove one's identity has become a critical component of trustworthiness verification, with identity being required for deeds of land, verification of age, confirmation of trade certification or skill, and permission to access or be briefed on confidential information.

Fast-forwarding to today, identity has evolved significantly from its humble beginnings to the use of digital ledger technologies such as blockchain, where in banking, supply chain, and certain military applications, the technology allows for nonrepudiation.

Identity is a multifaceted concept that is defined differently across different academic fields, typically referring to the characteristics that identify and differentiate an individual or an entity:

- **Personal Identity:** An individual's self-conception, including traits, values, and beliefs, often discussed in philosophical and psychological terms.
- **Social Identity:** Group-based aspects of identity, such as cultural or ethnic affiliations, which are commonly explored in sociology and anthropology.
- **Digital Identity:** In information technology, the information used by computer systems to represent an external agent—for example, a set of data that uniquely describes a person, an endpoint, or a system's service, as well as the means of controlling access to certain resources within a system based on that unique data.

The inception of IT authentication dates back to the 1960s with the advent of password usage, coinciding with the emergence of the first computers. These initial computers were notably large, expensive, and inefficient by contemporary standards. Their ownership was confined to a handful of universities and large enterprises; however, demand was significantly high. In response, academic institutions like the Massachusetts Institute of Technology (MIT) pioneered time-sharing systems, notably the Compatible Time-Sharing System (CTSS), to facilitate concurrent resource utilization by multiple users on a single machine. Passwords were implemented to avoid everyone having access to everything on those initial systems.

In modern IT architectures, the term *triple A*—henceforth written as *AAA*, which stands for *authentication, authorization, and accounting*—is synonymous with the use and monitoring of verifiable identity, providing the right levels of privilege to access key resources. While this technology is not new, having first been proposed as an IETF draft in 1999, the right set of capabilities, use cases, and scenarios related to its deployability has evolved significantly over the years.

Within IP networks, concepts such as network access control (NAC) using AAA capabilities to limit access to the perimeter of a computer network and restrict lateral movement by applying access restrictions using identity represent a strong foundation in applying the methodologies of zero trust within organizations, as introduced in detail in Chapter 1, "Zero Trust Demystified."

A Broader View of Identity

Today the concept of identity expands beyond the simple use of username and password, which historically was the method of "securely" accessing information systems that were either local or connected to the Internet. Over time, it became apparent that the username-and-password pair was maybe not the most effective way to maintain security. It also became apparent that password sprawl can lead to scenarios where users would need to document or write down their credentials somewhere, potentially resulting in further scenarios where the credentials could be stolen, or a breach could take place as a result of the user's password having existed elsewhere—perhaps even in a public service in the Internet that had been compromised in the past.

Surprisingly, even today, identity management is often split into separate teams within the security entity of an organization, leading to challenges with the flow of information. This often occurs due to historical reasons, or due to policies within the company prior to the level of globalization that we see today. These approaches to identity management often lead to challenges in maintaining the right levels of operational rigor, accountability, and visibility to adequately handle incident response and align key security standards and strategy across the company's estate, or in conjunction with partners and third-party vendors.

In today's architectures and systems, organizations can follow many standards and guidelines when it comes to identity and the correct levels of hygiene that should be applied from a security perspective.

In addition to systems such as workstations that are human operated (corporate endpoints), the number of Internet of Things (IoT) devices in enterprise networks is expected to exceed the number of corporate user endpoints in use by several counts. This change is being heavily fueled by the adoption of smart buildings, which represent additional challenges in the domains of identity, AAA, and profiling of these IoT devices. Based on our industry experience to date, these devices often lack the inherent security capabilities that are customary of corporate systems being used within organizations.

The National Institute of Standards and Technology (NIST) and the Center for Internet Security (CIS) have a rich set of best practices that should be considered in the context of identity while maintaining a zero trust–based architecture. These practices and recommendations transcend beyond simple authentication, including methods for logging and event retention and disabling orphaned or dormant accounts.

Table 9-1 provides an overview of security framework subcategories that are relevant to AAA. This framework is divided into five core functions: Identify, Protect, Detect, Respond, and Recover. Each function contains various categories with subcategories detailing individual outcomes.

DE.AE-3 falls under

- Function: Detect (DE)
- Category: Anomalies and Events (AE)
- Subcategory: DE.AE-3

Subcategory DE.AE-3 states "Event detection information is communicated to appropriate parties."