The Why and How of adopting Zero Trust Model in Organizations

Anita Nair, Solutions Architect at EY, India

Keywords: Perimeter Security Model, Zero Trust Model, Threats, DDoS, Ransomware, Malware, Man-in-the-middle, Cross-side scripting, SQL Injection, Phishing, Micro segmentation, Never trust always verify principle, Internet Threat Model, BeyondCorp.

ABSTRACT

As organizations move most of their workloads to public cloud and remote work becomes more prevalent today, enterprise networks become more exposed to threats both from inside and outside the organization. The traditional Perimeter Security Model assumes that threats are always from the outside. It assumes that firewalls, proxies, IDS, IPS and other state-of-the-art infrastructure and software solutions curb most of the cyberattacks. However, there are loopholes in this assumption, which the Zero Trust Model addresses. This paper discusses the Zero Trust Model and its mandates and evaluates the model based on the various implementations by the leading industry players like Google and Microsoft.

Overview of Cybersecurity in Organizations

Most security leaders today are not very confident about the existing security solutions working as expected when most of the organizations' workloads are in the public cloud space. It is therefore critically important that the cybersecurity model deployed, is effective and efficient in meeting the expectations of the most modern working conditions. Threats due to cyberattacks and hacking improvisations should have lesser window of opportunities to obliterate the assets held by organizations. In the following sections of the paper, the currently popular perimeter security model, which most enterprises have deployed today, will be compared with the zero-trust model that the industry is moving towards.

Cyberattacks are various types of malicious attacks that target computer networks, information infrastructures and private computer devices, using various methods to alter, steal and destroy data. The attacks can be active or passive, depending on whether it aims at altering or destroying the system resources or data, or gathering information from the system but not altering or destroying the resource. Cyberattacks bring loss to the organization and may result in loss of business, reputation or monetary loss. The business is ultimately compromised under such circumstances.

Types of Cybersecurity Threats and Attacks

Some of the most common types of cyberattacks and threats that computer networks at organizations are prone to are:

- *Ransomware Software* – This is a type of cyberattack where all files in the system get encrypted and the hacker demands the victim to pay if they want to regain the files. Organizations can protect their network from Ransomware through frequent assessment and good defence systems.
- *Man-In-The-Middle attack* – In this type of attack, the hacker tries to eavesdrop the communication between targets and sniff information while it is being transmitted over the internet. Encryption of data (example: SSL service) is the remedy used to curb such a scenario.
- *Denial of Service attack* – Here the host intentionally sends too many requests to the target server, to crash it. This results in disrupting or preventing legitimate users from accessing websites, applications and other resources. Usually load balancers help in balancing the requests while the

server is flooded with more requests than it can handle. The extended form of this attack is the Distributed Denial of Service attack (DDoS) in which many hosts send requests to the server, each of which tries to crash the system.

- *Cross-side scripting (XSS) attack* – In this attack, malicious scripts are used to exploit a web application. The script checks the sanitization of the web forms and writes malicious code into the website, to get useful information to the hacker, such as the user's cookie details. It is also capable of changing the user's settings and displaying downloadable malware. As a measure to prevent XSS attacks, web forms should always check the submitted values.
- *SQL Injection* – In this type of attack, the hacker places malicious code in the SQL statement via the input fields of the web form or via the URL. The hacker can likewise get the username and password information. Few methods to prevent this type attack are – input validation, parameterized queries, stored procedures, using escape characters in inputs, avoiding administrative privileges and using web application firewalls.
- *Phishing* – This type of attack falls under the category of social engineering. Here, the hacker sends the authenticated user, the webpage that looks the same as the legitimate website and tricks the user to submit their credentials. A thorough understanding of various phishing techniques is required to prevent this kind of attack.

Perimeter Security Model and loopholes

The basic assumption of the perimeter security model is that cyber-attacks always arrive from outside the network. This assumption led to securing the perimeter of the network. Security devices and software methods like firewalls, load balancers, VPNs, DMZ etc formed the basis of perimeter security. In this scenario, the company resources and data existed mostly within the physical walls of the organization. This worked well for protecting against malware, phishing, denial of service and zero-day attacks. Figure 1 represents a traditional network security architecture.

The disadvantages of this model are a) lack of intra-zone traffic inspection b) lack of flexibility in host placement c) single points of failure.

If the network locality requirements are removed in the above-mentioned network, the need for VPN is also removed. A virtual private network allows a user to authenticate to receive an IP address on the remote network. The traffic is then tunnelled from the device to the remote network. All these technologies of VPN, routers, switches, firewalls etc open out advanced capabilities at the network's edge. The core is never suspected and there is no enforcement of any measure in the perimeter model.

The way an external attacker could penetrate the perimeter security via trojan attacks can be explained by the concept of phoning home, which is a term used to refer to the behaviour of security systems that report network location, username or other such data to another computer. NAT is configured to allow internal users to gain network access. While there is a strict control on inbound traffic, outbound traffic through NATing could consume external resources freely. Internal hosts which in this way communicated freely with untrusted internet resources could thereby be abused while attempting the communication. Figure 2 shows how “phoning home” happens in a typical attack where the attack is launched by an internal host.

The hacker may send emails to all employees of a company whose addresses can be found in the internet by masquerading as a discount offer from a restaurant near the office. If an employee out of curiosity clicks the link, a malware is installed that phones home and provides the attacker with a session on the employee's machine. In this fashion, the attacker first compromises a low-security zone host and moves through the network towards high-security zones where the host has access to. The weak points in the perimeter security are usually places where firewall exceptions are made for various reasons such as, a web developer needing SSH access to production web servers or an HR representative needing access to an HR software's database to perform audits.

Once a privileged workstation is located by the attacker, a keylogger maybe installed and the developer password could be stolen, which could be used to elevate privileges on the production application host. Database credentials could be stolen from the application and the database contents could be exfiltrated.

Zero Trust Model and it's mandates

Zero Trust Model is where there is no trusted perimeter. Everything is primarily untrusted. A device, user and an application would, by default, receive the least privileged access to the architecture even after authentication and authorization. The mandates of zero trust are: a) never trust b) always verify c) enforce least privilege.

The concept of zero trust was first introduced by Forrester research and is implemented by enterprises that need to secure highly sensitive data from cyber threats. The purpose of zero trust architecture is to

address lateral threat movement within a network by leveraging micro-segmentation and granular perimeter enforcement, based on data, user and location. This is also known as the "never trust always verify" principle. The point of infiltration mostly is not the attacker's target location. The way you define movement or access depends on the user and their interactions and behaviour. For example, a user from the marketing department often has no access to sensitive financial files but would have access to CRM systems and marketing content. Hence, identifying who the users are, and verifying if their actions during a session are appropriate, is very important. It is important to make a note of which applications the users are trying to access and if it fits their roles and responsibilities.

Architecting Zero Trust Networks

Figure 3 shows the zero-trust architecture at a high level. The supporting system of the architecture is the control pane. All other components come under the data plane, which the control pane coordinates and configures. Requests for access to the protected resources are made through the control pane, where both the user and the device should be authenticated and authorized. Fine grained policies are to be applied to this layer, based on the role in the organization, time of the day or type of the device. The more secure the resource, the tighter would be the authentication.

Once the control pane decides that the user or device request is allowed, it dynamically configures the data plane to accept traffic from that client alone. It can also coordinate the details of an encrypted tunnel between the requestor and the resource. In summary, a trusted third party is granted the ability to authenticate, authorize and coordinate access in real time.

In zero trust, one must assume that the attacker can use any arbitrary IP address. Hence protecting resources using IP addresses no longer works. Hosts, even if they share "trusted zones" must provide proper identification. Since attackers can employ a

passive method and sniff traffic, host identification is not enough, strong encryption is also needed.

The three components of zero trust networks are a) user/application authentication b) device authentication c) trust. Apart from the user or application, device authentication is just as important. A trust score is computed and the application, device and the score are bonded to form an agent. Policy is then formed against the agent in order to authorize the request. With the authentication/authorization components and the aide of control panel in coordinating encrypted channels, we can be sure that every single flow on the network is authenticated. Unlike the perimeter security model, where security ends as soon as the traffic reaches the VPN concentrator, in this model, security is ingrained through out the network.

Implementing zero trust brings about several benefits to the business. Foremost among it is that it reduces the threat surface. It also provides increased visibility to all user activities.

The Internet Threat Model is defined in RFC 3552, which is also the model used by zero trust networks to plan their security stance. Zero trust networks expand on the Internet Threat Model by considering compromises on the endpoints. The response to these threats is to harden the systems proactively against compromised peers, and to facilitate detection of those compromises. Detection is done by scanning those devices and by the behavioural analysis of the activity from each device. Frequent upgrades to software on the devices, frequent and automated credential rotation and in some cases frequent rotation of the devices themselves is employed to mitigate compromises at the endpoint.

All zero trust networks use Public Key Infrastructure (PKI) that defines the set of roles and responsibilities that are used to securely distribute and validate public keys in untrusted networks. Entities like devices, users and applications are authenticated using digital certificates and this is done via automation. Because the public PKI system relies on publicly trusted authorities to validate digital certificates, that are costly, less flexible and not fully trustworthy. Hence, zero trust networks prefer private PKI.

Zero Trust Implementations

In this section we discuss two implementations of zero trust model - one from Microsoft and the other from Google.

At Microsoft, they realized that with growing cloud-based services and mobile computing, the technology landscape for enterprises would have higher need for the zero-trust access architecture. Figure 4 shows the different steps espoused by Microsoft to mature an organization's approach to security

Figure 4

- a) Follow least privilege access principles for identities, whether they are people, services or IoT devices.
- b) Once identity has been granted, data flows from a variety of endpoints – from IoT devices to smartphones, BYOD to partner managed devices and on-premise devices to cloud infrastructure. It is important to monitor and enforce device health and compliance for secure access.
- c) Apply controls to applications and API that provide interface by which data is consumed.
- d) It is important to classify, label and encrypt data and to restrict access to it.
- e) Whether one uses on-premise infrastructure, cloud infrastructure, container-based solution or microservices, the medium represents a critical threat vector. It becomes important to use telemetry to detect and flag risky behaviours.
- f) Segmentation of networks (micro segmentation) and deployment of real-time threat protection, end-end encryption, monitoring and analytics help secure networks.
- g) With increased visibility, an integrated capability is needed to manage the influx of data.

Microsoft identifies four scenarios to achieve zero trust: a) employees can enrol their devices to the device management to gain access to the company resources b) device health-checks per application or service can be enforced c) when not using a managed device, employees or business guests can use a secure method to access corporate resources d) employees can have user interface options (portals, desktop apps) to discover and launch applications and resources.

Microsoft's structured approach to implementing the various zero trust stages is shown in Figure 5. Figure 6 shows the reference architecture used by Microsoft using its own services to implement zero trust.

Figure 5

Figure 6

Google's implementation of zero trust is called BeyondCorp. It began as an internal google initiative to allow every employee to work remotely without VPN. BeyondCorp allows for single sign-on, access control policies, access proxy, and user and device-based authentication and authorization.

The fundamental components of BeyondCorp system include, Trust Inferer, Device Inventory Service, Access Control Engine, Access Policy, Gateways and Resources. The block architecture diagram is shown in Figure 7.

Figure 7

The various components of the system are:

- Resources** - an enumeration of all applications, services and infrastructure that are subject to access control.
- Trust inferer** - a system that continuously analyses and annotates the device states.
- Access policy** - a programmatic representation of the resources.

- Access Control Engine** - a centralized policy enforcement service referenced by each gateway.
- Device Inventory Service** - a service that continuously collects, processes and publishes changes about the state of known devices.
- Gateways** - a medium by which resources are accessed, such as SSH servers, Web proxies etc.

Evaluation of Zero Trust and Conclusion

Although zero-trust makes the designed network highly secure, it can still be compromised by attackers in some unique cases. The following are some scenarios and pitfalls.

- Identity Theft** - All decisions in zero trust networks are made based on authenticated identities. If one's identity is stolen, an attacker can masquerade their way through a zero-trust network. The identity which is linked to a secret should be therefore protected in different ways. Since zero trust networks need both the device and user/application to authenticate, it raises the bar compared to ordinary networks.
- DDoS attacks** - While zero trust networks are concerned with authentication, authorization and confidentiality, it does not provide good mitigation against DDoS attacks.
- Endpoint Enumeration** - It is easy for an adversary to observe which systems talk to which end points in a zero-trust network. Zero trust networks guarantee confidentiality but not privacy. Tunnelling the traffic through site-site tunnels makes it more difficult to see which individual hosts are communicating. However, it undermines the concept of zero trust.
- Untrusted Computing Platform** - The underlying computing platform of the zero-trust architecture should be trustworthy. If not, it is not possible to defend it.
- Social Engineering** - Whether it is phishing attacks or social engineering via face-face communication, zero trust networks can do only so much to defend against such attacks. Behavioural analysis coupled with user training is the only way this can be mitigated.
- Physical Coercion** - The best we can do under such a circumstance is to keep least sensitive information and systems vulnerable to compromise of a single

individual. Physical attacks against individuals are best mitigated by a consistent process of cycling both devices and credentials.

- g) *Invalidation* – It applies to long-running actions that were previously authorized but are no longer. An action could be an application-level request or a network session. How quickly and effectively ongoing actions can be invalidated, deeply affects the security response. The way to mitigate it, is to perform more granular authorizations on actions that are short-lived. Another approach is to periodically reset network sessions. The best approach is to have enforcement components to track ongoing actions and take ownership of the reset.
- h) *Control Plane Security* - It is possible to completely thwart the zero-trust architecture, if the control plane security is compromised. For sensitive systems like the policy engine, rigorous controls should be applied from the beginning. Group authentication and authorizations should be considered, changes to control plane should be made infrequently and should be broadly visible. Another good practice is to keep the control plane systems isolated from an administrative point of view, which means, they are kept in dedicated cloud provider networks or datacentres with rigorous access control.

While zero trust systems introduce new consideration points for network security, it resolves many other security issues. By applying automation and tried-and-tested security primitives and protocols, zero trust models will be able to replace the perimeter model as a more effective, secure and scalable solution.

References:

- [1] Zero Trust Networks – By Evan Gilman and Doug Barth
- [2] Microsoft Documentation
- [3] Google whitepapers on Beyondcorp

About the Author:

The author is a Software and Technology enthusiast who works at EY, India. Her areas of interests include Cloud Computing, Cybersecurity, Software Design Patterns, DevOps, Computer Networking and Infrastructure, Emerging Technologies and Research. She is a senior member of IEEE and a life member of CSI. She maintains a blog: <https://mindovertechnologyblog.wordpress.com/>