AN INTRODUCTION TO THE ZERO TRUST FRAMEWORK

Data security breaches and how to rein in the proliferation of unauthorized accesses to our online data are common topics of discussion in the network security profession. Yahoo! admitted that at least 500 million user accounts were compromised by a 2014 cybersecurity breach, and the Department of Justice, the Internal Revenue Service, the US Navy, and Snapchat all suffered material breaches in 2016. The 2015 breaches of the Office of Personnel Management (OPM) in two separate cybersecurity incidents resulted in stolen personnel files for almost 22 million people who had undergone background investigations for federal employment[i].

Despite the increased use of sophisticated data management technologies like firewalls, data encryption, digital identities and strong authentication to help thwart these attacks, data about our institutions and people continues to be compromised. Traditional security controls have not effectively considered users once on the network, or they have been deployed weakly or inconsistently in various stovepipe implementations throughout the enterprise. The rationale is because malicious external actors are presumed kept out and internal users are trusted, but history has shown again and again that not every user, device, or software application that ends up on the network accessing data resources is authorized. This white paper explores the value of an approach known as "Zero Trust" as a potential solution to the ongoing security threat.

What is Zero Trust Framework?

Forrester Research, Inc, a leading market and technology research company[ii], originally coined the term, Zero Trust[iii]. They define the approach as a cybersecurity design philosophy and architecture framework that adds a dimension of network segmentation and lockdown not typically pursued in traditional implementations of network security[iv]. Zero Trust introduces a fundamental shift in the way we think about current "Trust, But Verify" models and approaches to network security by presuming such models to be systemically ineffective.

Under traditional security frameworks, we tend to trust users and/or traffic once it is on our networks, granting generally unfettered access and navigation rights to the segments, applications, data, devices and services hosted on the network. Information security based on the "Trust, But Verify" mantra tends to be ineffective since we trust entities for data access once they are on our networks, and because verification and forensic reconstruction activities are generally performed reactively. Strong edge protections assume malicious individuals from external sources cannot penetrate the complex perimeter defenses in place at network entry points, but experience shows us they are still getting in, and once in can navigate around the network with little resistance, gaining access to compromise, disrupt, or infiltrate a wide range of systems and data resources anywhere on the network.

Zero Trust is a data-centric network design approach that puts micro-perimeters around specific data, applications, network devices and services so that granular rules (policies) for access control can be assigned and enforced. Zero Trust adopts the principle of "Never Trust, And Always Verify" by applying the access control philosophy of "Deny All/Trust By Exception". Any digital identity (person, device, or software application) on the network begins with fully restricted access to or use of network resources. Lateral movement on the network is similarly restricted. Under this configuration approach, access to all resources and navigation paths in the network are by default turned off ("deny all"), and only when rights are granted ("trust by exception") may an entity on the network access a resource or navigate to other network hosted resources in the same domain or authorized other security domains. The result is that the digital

identity would have to be explicitly granted permissions in order to access any network resource or move around in the network.

Zero Trust is engineered to fully integrate into the core of the network because threats originate not only from external sources, but also from malicious insiders who may even be in positions of trust. Zero Trust incorporates a ubiquitous real-time security framework throughout the network that permeates all accesses to network resources and navigation paths, not only at the network's perimeters, but also from within the network and in addition to periodic bulk network scans.

How Does Zero Trust Framework Work?

The Zero Trust Framework (ZTF) is in several key respects better than traditional cybersecurity philosophies. Under this model, all traffic on the network is presumed untrusted until and unless verified as authorized (digital identity is validated) and a confirmation has been made of the access rights explicitly assigned (individual or group/role privilege assignments). The CTO of an information security organization in the Netherlands likens Zero Trust network segmentation to a combination of levees, dams, and floodgates organized and partitioned to defend low-lying areas against storm surges and floods. Even if one levee system is breached, the "breach" is contained to just the specific segmented partition; no other segment is impacted. Using the levee system as a model, the key to ZTF is that all resources, including systems, applications, data repositories, devices and services, are segmented and by default assigned as restricted from access. ZTF employs the "least privilege" strategy for granting data access only to pre-vetted and authorized resource users with a need-to-know basis for mission support.

System owners, in collaboration with security engineers and architects, assign access rights to known digital identities, groups, or roles using a Role- or Attribute Based Access Control (RBAC/ABAC) methodology[v]. Access is given only to the network segments and resources needed to do a job or function. All other resources on the network remain restricted and off-limits until access rights are explicitly granted.

So, digital identities asserted on the network are granted access to resources based on the validity of the identity and the access rights explicitly granted to the identity. Personal Identity Verification (PIV) Card assertions, Security Assertion Markup Language (SAML), and Kerberos network authentication protocols are three of the more common protocols used to assert valid digital identities on networks today. Digital identities that find their way onto the network are restricted under Zero Trust to only the resources with access rights explicitly assigned. No other data access or navigation on the network is allowed unless authorized by predefined rights and assigned privileges. Identities on the network with no access rights assigned or seeking access to unauthorized resources are quarantined, inspected, and logged with real-time notifications issued to security engineers for immediate attention and investigation.

Under ZTF, network security is built from the inside out to complement strong perimeter controls. Rather than simply logging the traffic on the network for later inspection, Zero Trust first inspects the traffic on or coming onto the network in real-time in a non-disruptive manner for validity. If the traffic is suspect, notifications are sent, and then it is logged. This model replaces the traditional network design with multiple tools and control mechanisms, such as intrusion detection systems (IDS), traffic monitoring tools, continuous diagnostics, public key infrastructure, and smart cards, deployed in an attempt to foster a more secure environment.

Key Considerations to Building a Zero Trust Network Architecture

Moving to a Zero Trust network architecture is a significant challenge given the current investment and sunk costs in security tools. The earlier analogy of segmented flood protections in the Netherlands closely parallels the constructs of a Zero Trust Framework in information technology (IT) networks. Zero Trust segmentation requires a security solution that provides visibility into applications, data content, devices and services on the network by specific segment partitions, and that enforces access controls for digital identities seeking to gain access to these segments and to the resources hosted within the segment partitions. A key requirement of a Zero Trust approach is that it can be transparently integrated into the network without impacting existing routing and switching protocols. Zero Trust advocates for a segmented network (discrete security zones) with security built into the core of the architecture, rather that jury rigged and bolted on as an after-thought. This is typically achieved by implementing security appliances that can provide transparent integration, thereby reducing or eliminating compatibility issues and configuration collisions with the installed base of other adjacent networking devices.[vi]

To build Zero Trust into a network architecture, you must first identify critical systems, applications, data repositories, devices and services to be segmented and protected, and then map the transaction flows for these network assets.

- Critical data resources include, but are not limited to, sensitive but unclassified and mission proprietary applications and the data repositories that serve them containing, for example, personally identifiable information (PII), health related information (HIPAA), proprietary financial data, network management data and other intellectual property.
- Critical devices include network infrastructure (backbone) and peripheral devices on or connected to the network, for example, routers, switches, firewall devices, servers, personal computers, mobile (handheld) devices.

Next, IT and security professionals can plan and deploy Zero Trust segmentation gateways strategically in the network with access policies, creating trust boundaries (subnets) to host those applications, data repositories, devices and services with the access controls suitable to sensitivity levels of the hosted assets. A combination of network discovery tools for asset management and visibility, flow data analysis tools to analyze traffic patterns and user behavior, packet capture and analysis tools for malware determination, and network forensics tools to assist with incident response and criminal investigations are needed to gain the required level of situational awareness and notification. Examples of trust boundaries include specific domains for proprietary financial and other mission essential applications and data repositories, remote access services and device domains, guest access domains, B2B extranet access domains, infrastructure and peripheral device and management services domains, and personal computer and mobile device domains.

Finally, default access privileges for all trust domains and the individual assets hosted in those domains are set to Deny All/Trust-By-Exception.

- Authorized digital identities in the form of person entities (PEs), non-person entities (NPEs) such as devices and software applications, and/or group/role members are explicitly assigned access rights to trust domains and their hosted assets by least privilege policy settings.
- Digital identities, groups, roles, and any other “PE” and/or “NPE” entities on the network not explicitly granted access rights are quarantined, inspected, logged, and denied all accesses to assets or other trust domains based on the default Deny All/Trust By

Exception access policy; automated notification is rendered to security engineer staffs for all transactions that are suspect.

To improve implementation outcomes and avoid adverse impacts to the embedded base of network assets and the existing architecture, Zero Trust segmentation gateways and access policies should be implemented in planned stages:

1. Deploy segmentation gateways creating preliminary trust domains with open access policies.
2. Assign network assets to their predetermined trust domains based on sensitivity levels of the assets.
3. Coordinate staged assignments (lock down) of access rights to trust domains, network assets, and digital identities.
4. Engage traffic inspection, logging, and notification tools for the enabled trust domains.

A ZTF deployment may result in additional processing overhead on the network and nominal performance considerations for network administrators. Test plans should include before and after test cases that address possible performance degradations and/or denial of service implications of individual lock-down configurations. Considerations during testing should include the following to be performed by test engineers and reviewed by program managers and resource owners:

- Ping between hosts with large payloads to detect performance bottlenecks and any packet loss,
- Measure throughput using ttcp or iperf to detect throughput degradation,
- Monitor switch, segmentation gateway, and other device logs for performance changes and any dropped packets, and
- Watch syslog to see if workstations or servers experience similar changes.

Conclusion

Many in IT leadership positions envision re-architecting network environments to integrate Zero Trust to be a daunting, very complex and costly undertaking. It can be difficult to determine where and how to start such an endeavor and, if you do start, how to incorporate the architecture changes needed for Zero Trust without completely disrupting existing plans and frameworks to prevent adversely impacting ongoing mission operations. Some believe enough security is already built in as a result of existing investments in perimeter controls, continuous diagnostics, antivirus, firewalls, PKI and smartcards, VLANs and switch access control lists (ACL). On the other hand, Zero Trust is not a totally new concept when viewed in the context of its individual elements.

Zero Trust introduces a subtle change in the *focus* of our security endeavors from one that trusts the traffic once it is on our networks, granting almost unfettered access to the hosted resources, to one that "never trusts, and always verifies". The key to Zero Trust is resource segmentation and the use of very tightly enforced least privilege and trust-by-exception access control policies on every segment and asset in the network. While such controls may present additional runtime overhead and nominal performance considerations, suitable tuning and testing for possible performance bottlenecks, dropped packets, denial of service and/or other error conditions will produce optimal performance outcomes. With this architecture and access philosophy refinement, it is doubtful that the OPM or Yahoo! security breaches could have succeeded.

[i] John Kindervag, "'Zero Trust': The Way Forward in Cybersecurity"; InformationWeek 'DarkReading'; Jan 10, 2017 Downloaded from http://www.darkreading.com/attacks-breaches/zero-trust-the-way-forward-in-cybersecurity/a/d-id/1327827

[ii] Forrester Research, Inc.; 60 Acorn Park Drive, Cambridge, MA 02140; www.forrester.com

[iii] Kindervag, "Way Forward"

[iv] Forrester Research, Inc., "Developing a Framework to Improve Critical Infrastructure Cybersecurity"; Downloaded from http://csrc.nist.gov/cyberframework/rfi_comments/040813_forrester_research.pdf

[v] "RBAC" Role-based Access Control is a method of regulating access to computer or network resources based on roles of individual users. "ABAC" Attribute-based Access Control similar to RBAC defines an access control paradigm whereby access rights are granted to users through use of policies driven by digital identities and related person, role, or environment attributes. Access is the ability of a user to perform authorized tasks in the network such as "view", "create", "modify", "navigate", SearchSecurity

[vi] Au, "Steps"