Cybersecurity is awash with acronyms, jargon and marketing slogans. There is no dearth of so-called “silver bullets.” Yet, history and headlines demonstrate the ongoing success of adversaries and tell us a different story. Zero trust security offers promise, but it is largely misunderstood and perceived by some as yet another overly hyped buzzword. To properly understand zero trust, it is important to dissect the emergence of this concept, identify what it is (and is not), and determine if it can help advance cybersecurity and how.

Why Zero Trust Is Needed
Zero trust should be examined in the context of some of the biggest breaches in the 21st century (figure 1).

The much-publicized Target breach of 2013, in which 110 million customers had their personal data stolen, including credit and debit card details, does not even make the list of the top 15 breaches of the century. The growing scale of cyberattacks suggests that efforts often fall short against hackers’ tactics. Lack of budget and skilled staff are often cited as some of the persistent problems in securing organizations. Arguably, the security budgets at each one of the enterprises listed in figure 1 exceeds the annual revenues at scores of others, and many were decently equipped (if not well-equipped) with skilled cybersecurity staff and contractors.

But there are more than just numbers to consider when assessing cyberthreats. The growth in interconnectivity and ensuing complexity leaves even the best of enterprises on the defensive, especially when defending against an ever-evolving threat landscape. Some of the factors driving complexity (figure 2) include growing numbers and types of devices connected to the network, increasing demand for application availability and resiliency, consistent quality of service for the work-from-anywhere model, accessibility for strategic business partners and service providers, adaptability for business landscape evolutions and end-user requirements, and audit and regulatory compliance requirements. Security starts when these requirements are satisfied and, historically, has been bolted onto this framework.

As the scale of connections rises and threats evolve, the need for sophistication by attackers goes down (e.g., malware as a service can be purchased for as little as US$150\(^2\)). As business dependency on IT rises and regulatory requirements ramp up, organizations are constantly trying to catch up with attackers and get ahead of breaches. This is akin to having no walls at your home and trying to safeguard your assets by scanning everyone who walks in to ensure that they are doing...
<table>
<thead>
<tr>
<th>Organization</th>
<th>When</th>
<th>Impact</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adobe</td>
<td>October 2013</td>
<td>153M user records</td>
<td>Exposed customer names, IDs, passwords, and debit and credit card information</td>
</tr>
<tr>
<td>Adult Friend Finder</td>
<td>October 2016</td>
<td>412.2M accounts</td>
<td>Names, email addresses and passwords</td>
</tr>
<tr>
<td>Canva</td>
<td>May 2019</td>
<td>137M user accounts</td>
<td>Email addresses, usernames, names, cities of residence, and salted and hashed passwords; OAuth login tokens for Google sign-ins</td>
</tr>
<tr>
<td>Dubsmash</td>
<td>December 2018</td>
<td>162M user accounts</td>
<td>Email addresses, usernames, password hashes and other data such as dates of birth</td>
</tr>
<tr>
<td>eBay</td>
<td>May 2014</td>
<td>145M users</td>
<td>Names, addresses, dates of birth and encrypted passwords. 3 compromised employee credentials allowed complete access for 229 days.</td>
</tr>
<tr>
<td>Equifax</td>
<td>July 2017</td>
<td>147.9M consumers</td>
<td>Social Security numbers, birth dates, addresses, and in some cases driver’s license number and credit card numbers</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Application vulnerability and inadequate segmentation</td>
</tr>
<tr>
<td>Heartland Payment Systems</td>
<td>March 2008 (discovered Jan 2009)</td>
<td>134M credit cards</td>
<td>Standard Query Language (SQL) injection</td>
</tr>
<tr>
<td>LinkedIn</td>
<td>2012 (and 2016)</td>
<td>165M user accounts</td>
<td>Email addresses and passwords; put up for sale for just five bitcoins</td>
</tr>
<tr>
<td>Marriott International</td>
<td>2014-2018</td>
<td>500M customers</td>
<td>Contact information, passport number, Starwood Preferred Guest numbers, travel information, and in some cases credit card numbers and expiration dates</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Started with Starwood and propagated to Marriott after acquisition; attributed to Chinese intelligence group</td>
</tr>
<tr>
<td>MyFitnessPal</td>
<td>February 2018</td>
<td>150M user accounts</td>
<td>Usernames, email addresses, IP addresses and password hashes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Part of total of 617M customer accounts offered for sale on Dream Market</td>
</tr>
<tr>
<td>MySpace</td>
<td>2013</td>
<td>360M user accounts</td>
<td>Email addresses, passwords and usernames; put up for sale for six bitcoins</td>
</tr>
<tr>
<td>NetEase</td>
<td>October 2015</td>
<td>235M user accounts</td>
<td>Email addresses, passwords and usernames; put up for sale for six bitcoins</td>
</tr>
<tr>
<td>Sina Weibo</td>
<td>March 2020</td>
<td>538M accounts</td>
<td>Email addresses and plaintext passwords; denied by NetEaxe</td>
</tr>
<tr>
<td>Yahoo</td>
<td>2013-2014</td>
<td>3B user accounts</td>
<td>Real names, email addresses, dates of birth and telephone numbers</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A second breach compromised 1B names, dates of birth, email addresses and passwords, and security questions and answers.</td>
</tr>
<tr>
<td>Zynga</td>
<td>September 2019</td>
<td>218M user accounts</td>
<td>Email addresses, hashed passwords, phone numbers, and user IDs for Facebook and Zynga</td>
</tr>
</tbody>
</table>
nothing wrong. The general open connectivity philosophy is not far from this scenario, which drives up the cost of security and imposes a constant need for more skilled resources\(^3\)\(^,4\) while also driving up stress levels for security practitioners. And, unfortunately, as shown in figure 3, there is a tremendous skills shortage in cybersecurity. Therefore, it is necessary to rethink the best approach to this problem.

### What Is Zero Trust?

Although the term may be relatively new, the concept of zero trust has been around for several decades. Critical infrastructures such as military installations, power plants, nuclear facilities, medical facilities and financial institutions have, historically, operated using the concept of air-gapped networks.\(^5\) The idea of air-gapping is to separate the network into high (classified) and low (unclassified) segments, with physical separation between the two to minimize the exposed attack surface. Assets on the high network are highly valuable and the low network is not trusted with access to those assets. Although this model is not user friendly, it delivers stronger security than traditional networks.

The term “zero trust” was coined by Forrester analyst John Kindervag in 2010.\(^6\) It has been an evolving definition, but at the core, zero trust is a philosophy or framework for thinking about cybersecurity in today’s hostile environments.

As Forrester defines it, the traditional “trust, but verify” cybersecurity model offers attackers a broad attack surface that leaves security teams flat-footed and always in crisis management mode.\(^7\) Zero trust implements methods to localize and isolate threats (i.e., “never trust, always verify”) through micro-core, micro-segmentation, and deep visibility to identify threats and limit the impact of any breach.
Although the world is still in the early stages of the zero trust revolution, this thinking has some key benefits including:

- Reduction of exposed attack surface
- Proactive security
- Damage containment
- Relief of pressure on security teams by moving away from alert-driven architectures

**Implementing Zero Trust**

Although zero trust is a framework, it is supported by several different underlying technologies that offer protection on various fronts.

**Microsegmentation**

One of the earliest concepts to be positioned under the zero trust umbrella, microsegmentation was primarily designed to protect against lateral movement to minimize the impact of a breach. Lateral movement to capture information from key assets in the data center is a key component of several attacks. Microsegmentation aims to limit unsanctioned communication between workloads if they have no operative reason to do so, thereby minimizing breach impact (figure 4). Unlike typical north-south protection offered by firewalls, microsegmentation looks at east-west security for workloads.

Several vendors in the market offer products for microsegmentation, including Cisco, VMWare, Illumio and ShieldX. Different microsegmentation solutions tend to have differing models of implementations including agent-based, hypervisor-based, network-based and hybrid deployments.

Agent-based solutions have the advantage of residing on the host and can, therefore, look more easily at protocols and encrypted traffic. Network-based deployments extend a segmentation model that many network administrators are familiar with and minimizes the learning curve. Solutions are usually capable of both on-premise or cloud deployments, but hypervisor-only models may be limited in their capability to look into containers or bare metal while offering benefits in a virtual machine (VM)-heavy environment.

It is also worth highlighting that very basic levels of isolation among different cloud workloads can be implemented using native cloud controls such as security groups, L2/L3 firewalls and subnet level filtering capabilities.
The questions to ask when considering microsegmentation include:

- Is the organization’s applications communication model understood? Modeling network communications is crucial to ensure that segmentation does not break applications.
- Can the solution offer visibility, administration and migration? Policy management can be time-consuming, but it is an important step to ensure true success of the project.
- Does the infrastructure include nontraditional endpoints such as supervisory control and data acquisition (SCADA) systems, Internet of Things (IoT) and mainframes? Many solutions do not cover these nontraditional endpoints.
- Are there compatibility considerations between microsegmentation and other security technologies (e.g., endpoint agents, network intrusion prevention systems [IPSs], next-generation firewalls [NGFWs])?
- Is the organization looking at cloud deployments or primarily on-premise deployments? What level of segmentation is needed for the cloud? Would native controls suffice?

Zero Trust Network Access

Zero trust network access (ZTNA) creates a context-based logical boundary for application access. The goal is to minimize access and abuse by user- and device-context-driven application access. ZTNA challenges the assumption that the location of an entity (e.g., inside a network) should grant trust automatically to a user or device. ZTNA hides all applications on the network and allows access based on attributes such as user identity, device, geolocation and security posture. In other words, users get access to applications they need, but they see nothing else on the network (Figure 5). A ZTNA broker assesses the user’s profile before granting access. The user may be an enterprise employee, a partner or a contractor. ZTNA has its roots in the BeyondCorp architecture and concepts of software-defined perimeters. Sample vendors include Akamai, Netskope, Perimeter81, Pulse Secure and Zscaler.

Although ZTNA has been touted as a virtual private network (VPN) replacement, not all solutions truly replace VPNs. Also, it is important to note that not all applications lend themselves to ZTNA (e.g., consumer-facing applications).

There are two approaches to implementing ZTNA: endpoint-based or network-based. In the former case, ZTNA initiates at the endpoint, and successful end-user authentication exposes the appropriate

---

**Figure 5—Zero Trust Network Access Minimizes Access and Abuse**

_user A: Can only see Apps A and B_

_user B: Can only see Apps B and C_

_user C: Can only see App D_

Application A
Application B
Application C
Application D
applications. This typically requires an agent to be installed on the endpoint. In the latter case, a connector initiates registration with the ZTNA broker on behalf of the application provider. A user is given access to the service once they pass the appropriate context and identity checks. This model traditionally works well with Hypertext Transfer Protocol (HTTP)/Hypertext Transfer Protocol Secure (HTTPS)-based applications.

The questions to ask when considering ZTNA include:

- Is the organization able to ensure appropriate user-to-application mapping and application-usage modeling? The key to ZTNA success is knowing who needs what access under what circumstances.
- Are there unmanaged devices that need access to applications, and how would endpoint models work in this case? Is it also necessary to consider programmatic access?
- Do partners, contractors or others need third-party authentication services? Does the ZTNA provider offer this capability?
- How scalable and resilient is the ZTNA broker? It could be a single point of failure.
- Where are the points of presence, and what is the expected latency for a cloud provider?

Remote Browser Isolation
Remote browser isolation (RBI) operates on the principle that end users represent the weakest links in any organization and Internet access is one of the largest attack surfaces responsible for breach origination. Therefore, one of the most effective ways to shrink the endpoint/end user compromise in a cloud-first world is to isolate web access to eliminate browser exploits and threats such as ransomware, malvertising and phishing.

Building on the concept of air-gapped networks, RBI creates a virtual airgap to ensure security while also addressing the user experience. Instead of trying to detect all potential dangers and training people to avert them, the RBI model isolates the risk. When an employee clicks on a link, RBI remotely renders the page to ensure that the malware does not even reach the endpoint (figure 6).

Enterprises offering browser isolation include Cyberinc, Ericom, Menlo Security and Symantec (Broadcom). Browser isolation can be delivered at the endpoint or in the network. The former necessitates an agent installation with appropriate virtualization support on the endpoint but is deployed closer to the point of breach. The network-based RBI deployments are more popular, can be on-premise or in the cloud and usually require minimal changes at the endpoint.
The questions to ask when considering RBI include:

- Does the user experience remain relatively intact with RBI? Some solutions require more user training, which is counterproductive.
- Does the solution cover different security risk scenarios such as web threats, file-based threats, email-based threats and credential theft attacks? Some solutions do not remotely render, which may open the endpoint to risk from media- or CSS-based attacks.
- Do the deployment models align with the enterprise vision and goals? If using the cloud, does the solution have the right points of presence to minimize latency? What edge deployment options does the provider support?
- What level of administrative visibility and control does the solution offer? Do more components need to be purchased to gain visibility and necessary security?
- Does the solution interoperate with security tools such as secure web gateway (SWG), firewalls, and security information and event management (SIEM)?

**Conclusion**

Zero trust is not a panacea and no vendor should give that impression. The goal of zero trust is to minimize the risk by assuming that breaches are a constant and adapting deployment strategies to localize, isolate and limit a breach's impact.

Revisiting the Target breach through the lens of zero trust shows that there are several points where it could have played a role (and perhaps has in other instances). It is well-documented that the breach started with a phishing attack on a third-party heating, ventilation and air conditioning (HVAC) contractor. This led to a compromise of credentials, which, in turn, allowed access to internal systems, eventually allowing the jump to the point-of-sale (POS) systems across stores. Although some security solutions did generate an alert, they were lost among several thousand other alerts security teams already had to address. A zero trust model would have stopped the initial compromise using technologies such as RBI; credential abuse would likely have been blocked by ZTNA and the jump to the POS segment could have been prevented by microsegmentation.

Most zero trust technologies coexist, complement and interoperate with traditional technologies. However, it is important to understand and evaluate them to suit each organization.

The best return on investment (ROI) usually comes from assessing risk and then understanding feasibility. For instance, if controlled access to cloud services is the biggest risk, ZTNA might be a good place to start, but if end users are a weakness, RBI would be a smart investment. The best place to start the zero trust journey is to assess where risk is the highest.

**Endnotes**

10 BeyondCorp, https://beyondcorp.com/