Gigamon®

Zero Trust and Importance of Network Visibility

WRITTEN BY:

John Kindervag
Creator of Zero Trust

In the years since I first introduced the concept, Zero Trust has become a global movement that organizations across industries are joining to achieve resilience and protect their IT environments against devastating breaches. The U.S. government has adopted it as the model for securing Federal Agencies' systems and growing numbers of businesses continue to follow suit. Zero Trust is a powerful means of preventing cyberattacks from being successful.

Zero Trust fundamentally changes the incentive structure for security within organizations that adopt it. It provides a robust strategic framework that resonates with leaders, who are empowered to incentivize teams to do the boots-on-the-ground work of stopping bad packets and blocking malicious traffic. Without this incentive structure, security practitioners were often tempted to allow potentially malicious traffic to flow under the guise of maintaining network performance or employee productivity. In this way, Zero Trust is a paradigm shift that makes stronger security possible, but only if the right policies and implemented and enforced in the right places.

To achieve this, you need full visibility across your entire network, regardless of whether assets reside on-premises, are hosted in the cloud or there's a mix of both. Maintaining this high-fidelity visibility is at the heart of successful Zero Trust adoption.

Ever heard the old joke about the police officer who sees a man searching for something under a streetlight? The officer asks him what he lost, and the man says that he's missing his keys. They then proceed to look for the keys together, under the streetlight. After a few minutes, the police officer asks, "Are you sure you lost your keys here?" The man responds, "No, I lost them in the park." The police officer then asks, "Well, why are you searching in this spot if you lost your keys somewhere else?" The man replies, "because this is where the light is."

Gaining complete visibility into the network is like lighting up the whole street. When it comes to implementing Zero Trust, this is the best place to start. Adding streetlights (that is, network visibility) lets you see where the attackers are hiding, since they're not standing under the streetlights where you can see them. And the network is where you can gain confidence by validating that what you've observed elsewhere – whether that's through endpoint detection and response (EDR) or from logs – is really happening. By looking into packets, you achieve the detailed visibility you need to correlate your findings from elsewhere – or see things that would otherwise be invisible. That's what gives you the power to stop threats before they do real harm.

That adopting Zero Trust requires improving network visibility is something I started preaching more than a decade ago, but the

world of cybersecurity still hasn't learned its most important lessons. When the Zero Trust paradigm was still in its infancy, I proposed creating a new type of network, called a "data acquisition network" (DAN),[1] that would be built specifically to enforce Zero Trust.

Unfortunately, DANs have not yet been widely adopted. But now it's time. Cyberattacks are much more sophisticated today than they were in 2012 when I first started writing about the value of DANs. Cybercriminals are much more professional and better funded, and their tools are more powerful than ever. Plus, today's networks are different. Cloud adoption reduces network visibility. These factors combine to make attacks more disruptive and damaging, and they're only getting worse. Yet defenders continue to look for evidence of malicious activity in all the wrong places.

From a Zero Trust perspective, the protect surface is where all the data that's critical or regulated is stored. Defining this is the first step in implementing Zero Trust security. In order to defend the attack surface (which sometimes seems like it's forever expanding), we need to know where to focus our efforts. That clarity of purpose is what defining the protect surface gives us. Typically, in a Zero Trust network, the protect surface is defined on the basis of four things (we can use the acronym DAAS to help us remember them).

- **Data**
  What data needs to be protected?
- **Applications**
  Which applications consume sensitive information?
- **Assets**
  Which assets are most sensitive?
- **Services**
  Which services, such as DNS, DHCP, and Active Directory, could be exploited to disrupt normal IT operations?

These are well-defined, knowable things that we can readily identify. The protect surface is deep inside the network. It's where we need to be looking.

A DAN facilitates the extraction of network data – typically, packets, syslog, or SNMP messages – to a single place where you can then inspect and analyze it in near real time.

A DAN is an attractive concept; anybody who's ever had to troubleshoot networks knows how hard it is to capture packets in a network effectively. Because all traffic traverses the segmentation gateway, which interconnects all MCAPS, data acquisition can be accomplished efficiently.

All of this traffic can be mirrored and forwarded to a DAN MCAP where security information management (SIM) and network analysis and visibility (NAV) tools centrally capture, analyze, and log all traffic transversing the network. NAV, along with traditional SIM tools, provides a type of network omniscience that is imperative in today's threat environment.

— **FORRESTER RESEARCH,** *"BUILD SECURITY INTO YOUR NETWORK'S DNA: THE ZERO TRUST NETWORK ARCHITECTURE,"* **NOVEMBER 2012**

Logs Are Not Enough

A prime example of a "wrong place" where cybersecurity teams often look is within logs. Collecting logs became popular after monitoring network activity was mandated within the Payment Card Industry Data Security Standard (PCI DSS). At the time of the earliest credit card data breaches, logging wasn't commonplace. So the payment card industry required organizations to maintain logs of activity within their cardholder data environments and review them at least once daily. This made logging – and log management – ubiquitous, which was good news for the security information and event management (SIEM) vendors. What it didn't really do was improve visibility.

THE FRAGILITY OF LOGS

Logs are fragile. From a network administrator's perspective, it's more important to keep the packets flowing than it is to have visibility. This means that at times when the network's performance is strained (such as, for instance, during a cyberattack) logging will be halted to maintain uptime. This is a fundamental flaw in the entire model: you'll go dark during an attack, which is exactly when visibility is needed most.

Let's take NetFlow sampling as an example. Network devices like firewalls and switches generate NetFlow data automatically. But this is metadata, not actual packets, so it provides only limited information. Plus, it's sampled, meaning that instead of collecting every packet that transits the device, NetFlow collects one in every n packets. Because of this, NetFlow data is incomplete. It's like a still surveillance camera that takes a picture every five seconds: there's still plenty of time for a crime to happen in between the snapshots.

Moreover, when the networking equipment that's collecting the NetFlow data is under load, it'll sample less frequently to maintain performance. This makes NetFlow particularly unreliable as a means of maintaining network visibility when there's lots of traffic.

SPAN PORTS ARE SO LAST CENTURY

The same problem surfaces when you're relying on switched port analyzer (SPAN) ports for packet capture. SPAN ports (sometimes also called mirror ports) can create copies of selected packets passing through a switch or router and send them to a designated location. Administrators can select the type of monitoring they want to do by configuring the port. But because the primary purpose of the switch or router is to forward production packets, collecting SPAN data is given a lower priority on the device. This creates a serious issue: when there's too much traffic, SPANs will drop packets. If you're relying on SPAN ports for visibility and traffic levels spike during an attack, that's exactly when you'll lose the ability to see what's happening.

YOU DON'T GET ENOUGH OF THE GOOD STUFF

An over-reliance on logging leaves you in the dark. Normalization is problematic because it limits what logs can tell you. Syslog is the most popular protocol used to send event data to a central place for storage or analysis. But syslog isn't standardized. If you have multiple different devices from different manufacturers, the syslog event messages that they generate may have different fields. To handle these discrepancies, the system that stores these logs (usually a security information and event management (SIEM) platform) will strip off the fields that are outside of normal syslog parameters or that don't correspond to predefined fields within its database.

The issue is that evidence of an attack can and often does fall outside of the parameters of a normal syslog message. For purposes of detection, this data may be the most important information of all, yet it's exactly what gets dropped.

EDRs Are Not Enough

Another "wrong place" where today's security teams are spending too much time and effort looking is at the endpoint. Endpoint detection and response (EDR) is the latest trend in

cybersecurity marketing. It's one of the fastest-growing areas in our industry. This doesn't mean that it provides the best ROI or the biggest bang for the technology buyer's buck.[2]

In part, EDR is popular because it's a natural evolution of the first cybersecurity technology to be marketed and sold at scale – anti-virus software. The problem is that in today's world, the most important data no longer resides on the endpoint. Organizations with mature cybersecurity programs don't let employees store sensitive or business-critical data on their laptops anymore. Thus, endpoints aren't typically a protect surface. Instead, the most important data resides on an internal server, where it can be pulled and accessed, or it resides in the cloud, where it's easier to protect. So EDR's focus isn't where the attackers are interested in going. All this emphasis on the endpoint is a red herring.

Of course, endpoints should be protected, but their importance is secondary to that of the DAAS elements within the protect surface.

It is true that EDR logs can add additional value if they're correlated with the data you collect from a DAN. Correlating those logs with telemetry from other network and security technologies can give you a deeper level of visibility. But EDR, in and of itself, is not enough.

Eradicate the Blind Spots

The reality is that organizations continue to have enormous blind spots, particularly when it comes to visibility into network traffic. All too often, folks assume that because the network is running — and because it's performing well — attackers can't be present. But without packet-level visibility across the entire organization, you can't know this for sure. A federal law enforcement official once told me of a breach in which the attackers actually optimized network performance in order to accelerate data exfiltration — and no one within the target organization even suspected that their network had become faster and more reliable because it had been compromised.

Companies generally focus on ingress, looking at what's coming into the network (like malware), instead of what's leaving the network. But what's most important in identifying data breaches is understanding whether confidential or sensitive data has been exfiltrated from your network or systems. We need to pay attention exactly where our most valuable data is and think about how it might come into the hands of a bad actor. Building a data acquisition network gives you comprehensive and granular visibility. This makes it possible to observe exfiltration as it's happening. That's what's needed to stop it.

Want Real Visibility? Build a DAN

Legacy networks were designed and built from the outside in. Zero Trust requires new ways of thinking about network design, where you aren't trying to achieve security by overlaying sets of controls over existing infrastructure, but instead are re-envisioning the network's fundamental design to enforce security throughout. One essential concept in Zero Trust is that you must inspect and log all traffic that traverses the network. A DAN makes this possible by mirroring all traffic and forwarding it to a central location for monitoring and analysis.

Zero Trust isn't a product. There's no single vendor-supplied solution you can buy that will give you Zero Trust. From a technological perspective, Zero Trust is implemented as a system. And there are products that will work well within that system. Network packet brokers are a good example. A network packet broker supports the aggregation of high-speed network traffic access points (TAPs), which can copy all packets — of any size — saving the duplicates for use in security monitoring. This creates a visibility fabric that mirrors all the traffic flowing across the network, giving security teams the data they need to generate insights that lead to action. TAPs allow for unidirectional traffic monitoring, which means that monitoring can be done covertly, even if the entire switching infrastructure is compromised. Such capabilities are what enable security teams to operationalize visibility.

Lastly, NIST also sees the need for a Data Acquisition Network and NIST 800-207 states: "The enterprise records packets seen on the data plane, even if it is not able to perform application

layer inspection (i.e., OSI layer 7) on all packets. The enterprise filters out metadata about the connection (e.g., destination, time, device identity) to dynamically update policies and inform the PE as it evaluates access requests."

PACKET CAPTURE: NOT JUST FOR TROUBLESHOOTING ANYMORE

Traditionally, packet capture was used solely for IT troubleshooting. The visibility that it affords was harnessed for purposes of maintaining and optimizing network performance, not defending against cyberattacks.

One of the primary reasons for this is that IT and security departments often don't cooperate as well as they should. Packet capture was done by network performance monitoring solutions, and the budget for these tools fell within the IT budget, not the security budget.

Second of all, packet capture required a great deal of storage, and in the past, that storage was expensive. It also requires monitoring, for which — given the fact that security teams never have enough people — isn't always easy to find the resources.

Today, packet capture should rightfully be a central, enabling technology that makes achieving Zero Trust security possible. It does this by delivering full, true and granular visibility across the protect surface. This gives attackers nowhere to hide, and makes lateral movement impossible.

Packet capture myths and realities

One common misconception about direct packet capture (PCAP) is that it necessitates full packet storage. This simply isn't the case. Once the network traffic has been captured, security teams have several different options for what to do with it. Which one you should choose depends on your organization's unique needs, the business criticality of the workloads involved, and the circumstances.

- **Full packets** can be stored in cases when you want to have direct PCAPs available for offline analysis. This might make sense in case of a security alert or incident, when investigators need access to the granular details that will allow them to figure out exactly what was compromised, and how this was done. Storing full packets may also be appropriate for highly sensitive information (such as military intelligence) or in highly regulated industries. Storing full packets means you'll quickly accumulate large quantities of data, which can be labor-intensive to search manually. One way to reduce the expense and resources required is to store full packets for a limited amount of time (such as three days).
- **Metadata** that's been extracted from the PCAP data stream can be stored instead. Many different forms and types of metadata can be extracted for analysis, and can be stored using a fraction of the space that would be needed to store full packets. It's possible to extract specific information (such as source and destination IPs, affected ports, or protocols use) that's most relevant to a particular investigation. Metadata storage allows for the historical tracking of device and network activities over time, and it's quicker and easier to correlate metadata over longer periods than it is to parse more information-dense PCAP datasets.

Once you're capturing full packets, it's not necessary to choose between storing all of them all the time, some of them sometimes, or just metadata. Instead, it's a question of where to best use which approach, and for what purposes. The goal should be to use resources efficiently while meeting your alert and incident investigation needs.

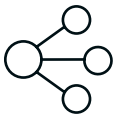
Five Steps to Building a Zero Trust Network

As stated in the President's National Security Telecommunications Advisory Committee's Report to the President on Zero Trust and Trusted Identity Management,[3] which has become an authoritative source defining Zero Trust and outlining how to implement it, there's a five-step process for Zero Trust implementation. These steps are relevant to all implementations, whether commercial enterprise or government networks.

1. DEFINE YOUR PROTECT SURFACE

What data, applications, assets and services are important to your organization? These are known as the DAAS elements, which we discussed earlier in this paper. As defenders, our goal is to protect something, and the DAAS elements within the protect surface help use remember what that "something" is. Only by figuring out what matters can you understand where to focus your efforts.

By defining a protect surface, you've created something that's orders of magnitude smaller than the overall attack surface – something that's manageable, controllable, and well-defined. We can then move our security controls as close as possible to the protect surface to define a micro-perimeter around it.



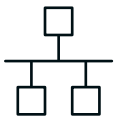
2. MAP THE TRANSACTION FLOWS

Understand how the systems work together as a system. This will show you where controls need to be placed. Before the advent of Zero Trust, security architects designed controls upfront on the basis of generalized ideas of how networks are designed, but reference architectures don't work for everyone. Zero Trust entails creating a security environment that's bespoke — one that's tailor-made for the individual organization.

Mapping transaction flows entails tracing:

- Which applications have access to which critical data
- Which users have access to those applications
- Which users and applications have access to which infrastructure

With this information in hand, you can determine which technologies to implement.



3. ARCHITECT THE ENVIRONMENT

You transaction flow map will show you where to put the elements in your security architecture. All of the controls — whether they've next-generation firewalls (NGFWs) or cloud container security controls — should be designed on the basis of an understanding of how the system works as a whole. Controls should be as close as possible to the protect surface, and the entire environment should be designed to make it possible to consistently enforce policies that safeguard the protect surface.

By visualizing the environment as a whole, you can see where it makes the most sense to put the controls.

4. CREATE ZERO TRUST POLICIES

You can start to build out Zero Trust policies by answering six questions defined by Rudyard Kipling in a 1902 poem in relation to your network traffic:

- **Who** should have access to a resource? This defines your authentication and identity policies.
- **What** applications should they be allowed to access?
- **When** in what circumstances should access be allowed?
- **Where** are these resources located? On-prem? In the cloud? If so, is it a cloud services that's located in a particular region or country?
- **Why** this contextual information can help with data classification.
- **How** we verify policies are actually being enforced.

Ultimately, each policy statement will come down to a simple "allow" or "deny," but you can have massive amounts of data — and very complex criteria — that you use to make that decision. Today, it's crucial to enforce these policies at Layer 7. All modern attackers know how to bypass Layer 3 controls, which is problematic in the cloud, since cloud-native controls operate at Layer 3.

5. MONITOR AND MAINTAIN THE ENVIRONMENT

The final stage in the process is to gather telemetry, perform analytics and automate policy updates in response. The idea is that you'll build a feedback loop that will enable the entire system to become stronger over time. This can only be achieved if you have enough data about the stresses that are being applied to your environment to understand how it can be adapted in response. After all, modern IT environments are ever-changing, so Zero Trust policies have to constantly adapt, too.

Zero Trust and the Antifragile Network

Operationalizing visibility supports all of the steps in Zero Trust implementation, but it's especially essential for defining your protect surface, since you can't protect what you can't see. A network packet broker can supply fine-grained visibility across all of your networking devices, making it possible to see across all traffic, all the way through layer seven, the application layer.

It's also impossible to achieve the fifth step, monitoring and maintaining the environment, without that network visibility. The telemetry from a packet mirroring solution not only helps prevent significant security events, but also will yield valuable insights that can be used to improve your processes over time.
This concept at the core of the idea of antifragility.[4] Fundamentally, Zero Trust enables the building of an anti-fragile network, one that improves its resiliency to shocks, stresses, mistakes, attacks and failures over time. Rather than degrading its performance as technology evolves, a true Zero Trust network will evolve to work better and better.

Few of today's networks are truly antifragile. Building a Zero Trust network requires inverting many of the paradigms that legacy network designers held near and dear. But the result – achieving the capacity to making sure cyberattacks can't succeed – is well worth the effort.

It's ten o' clock. Do you know where your data is?

About the Author

John Kindervag is the Senior Vice President Cybersecurity Strategy and Global Fellow at ON2IT. He spent the previous four years at Palo Alto Networks as Field CTO. Before Palo Alto Networks, John spent eight and one-half years at Forrester Research as a Vice President and Principal Analyst on the Security and Risk Team. John is considered one of the world's foremost cybersecurity experts. He is best known for creating the revolutionary Zero Trust Model of Cybersecurity.

In 2021, John was named to the President's NSTAC Zero Trust Sub-Committee and was a primary author of the NSATC Zero Trust report that is being delivered to the President of the United States. Additionally, John was named CISO Magazine's 2021 Cybersecurity Person of the Year.

John has a practitioner background, having served as a security consultant, penetration tester, and security architect. He has been interviewed and published in numerous publications, including The Wall Street Journal, Forbes, Bloomberg, and The New York Times. He has also appeared on television networks such as CNBC, Fox News, PBS, and Bloomberg discussing information security topics. John has spoken at many security conferences and events, including RSA, SXSW, ToorCon, ShmoCon, InfoSec Europe, and InfoSec World.

John has a Bachelor of Arts degree in communications from the University of Iowa and lives in Dallas, TX.

About Gigamon

Gigamon offers a deep observability pipeline that harnesses actionable network-level intelligence to amplify the power of observability tools. This powerful combination enables IT organizations to assure security and compliance governance, speed root-cause analysis of performance bottlenecks, and lower operational overhead associated with managing hybrid and multi-cloud IT infrastructures. The result: modern enterprises realize the full transformational promise of the cloud. Gigamon serves more than 4,000 customers worldwide, including over 80 percent of Fortune 100 enterprises, nine of the 10 largest mobile network providers, and hundreds of governments and educational organizations worldwide.

To learn more, please visit gigamon.com.

[1] John Kindervag, Build Security into Your Network's DNA: The Zero Trust Network Architecture, November 2012. Available at: https://www.virtualstarmedia.com/downloads/Forrester_zero_trust_DNA.pdf.

[2] For additional context, see John Kindervag, "No More Sucking Chest Wounds," On2IT. Available at: https://on2it.net/no-more-sucking-chest-wounds%EF%BF%BC/.

[3] The President's National Security Telecommunications Advisory Committee, Draft Report to the President: Zero Trust and Identity Management. Accessed at: https://www.cisa.gov/sites/default/files/publications/Final%20Draft%20NSTAC%20Report%20to%20the%20President%20on%20Zero%20Trust%20and%20Trusted%20Identity%20Management.pdf.

[4] This concept is derived from Nassim Nicholas Taleb's book *Antifragile: Things That Gain from Disorder* (Random House, 2012).

Gigamon®

Worldwide Headquarters
3300 Olcott Street, Santa Clara, CA 95054 USA
+1 (408) 831-4000 | gigamon.com

02.23_02