Using Zero Trust to Defend Against Every Stage of a Ransomware Attack

zscaler™

Introduction

Ransomware has been a threat for more than thirty years, but recently the nature of the battle has changed.[1] Within the last three months, ransomware has evolved more quickly than it's done over the past half-decade. In today's world, global businesses must assume that every ransomware attack is also a data breach, a fact that has profound implications for how we should design security architectures and network defenses.

Far too many organizations continue to rely on legacy perimeter-based cybersecurity strategies. In this model, known as the castle-and-moat approach, defenses—which are typically firewall-based—are centered around the network's perimeter. This leaves resources and applications unprotected when attackers attempt to move laterally across the computing environment—and this sort of lateral movement is the number one contributor to the success of ransomware attacks.

Every successful ransomware attack must include an initial incursion in which attackers establish a foothold within the environment. The next step of a successful attack is facilitating the ability to move laterally within the organization's network. Without lateral movement, the ransomware wouldn't be able to infect more than a single machine, and its devastating effects wouldn't be felt across the entirety of the enterprise. And without lateral movement, the attackers wouldn't have access to the data repositories from which they're exfiltrating invaluable information. Simply put, without lateral movement, there would be no ransomware problem.

Instead, as recent high-profile events have demonstrated, the ransomware threat has grown to crisis proportions. In mid-May of 2021, a ransomware attack forced Colonial Pipeline, one of the largest pipeline operators in the U.S., to shut down operations in the aftermath of an IT systems freeze.[2] This measure brought the transport of roughly 45% of the East Coast's fuel supply to a halt. Less than two months later, the IT solution provider Kaseya, which supplies remote monitoring and management tools to managed service providers (MSPs), reported that a vulnerability in its software had allowed malicious payloads to spread to more than 70 MSPs and their customers. All told, as many as 1,500 end-user organizations were impacted, making it the most widespread ransomware attack to date.

The current wave of ransomware attacks shows significantly more sophistication and innovation than ransomware operators were exhibiting even one year ago. If 2020's events accelerated the pace of digital transformation by three to four years—as numerous analysts claim[3]— it's apparent that ransomware's development has sped up in tandem, with attack tactics evolving more quickly than they had over the entirety of the previous decade.

New strategies that today's ransomware attackers are employing include double extortion, a method in which data is *both* encrypted and exfiltrated to increase leverage and ransom demands, staged attacks including a distributed denial-of-service (DDoS) component, and supply-chain attacks that leverage third-party partners and vendors to gain access to victims' environments.[4]

On the one hand, researchers are observing larger numbers of highly targeted, customized attacks. On the other, they're seeing greater volumes of "spray and pray" style attacks carried out by less sophisticated adversaries relying on commodity malware, or Ransomware-as-a-Service (RaaS) kits that are now readily available on the Dark Web.

[1] Source: https://www.zdnet.com/article/30-years-of-ransomware-how-one-bizarre-attack-laid-the-foundations-for-the-malware-taking-over-the-world

[2] Source: https://www.zdnet.com/article/colonial-pipeline-ransomware-attack-everything-you-need-to-know

[3] Source: https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-covid-19-has-pushed-companies-over-the-technology-tipping-point-and-transformed-business-forever

[4] Source: https://info.zscaler.com/resources-industry-reports-threatlabz-ransomware-review

Ransomware was the third-most common and second-most damaging type of malware attack in 2020.
[World Economic Forum, *Global Risk Report 2020*]

The average ransomware incident costs its victim $1.45 million in total remediation costs.
[Gartner, *6 Ways to Defend Against a Ransomware Attack*, https://www.gartner.com/smarterwithgartner/6-ways-to-defend-against-a-ransomware-attack/]

There has been a sevenfold increase in ransomware's prevalence since the start of 2020.
[Gartner, *Detect, Protect, Recover: How Modern Backup Applications Can Protect You From Ransomware*, https://www.gartner.com/doc/reprints?id=1-25ATIQU6&ct=210222&st=sb]

75% of organizations are expected to face one or more ransomware attacks by the end of 2025.
[Gartner, *Detect, Protect, Recover: How Modern Backup Applications Can Protect You From Ransomware*, https://www.gartner.com/doc/reprints?id=1-25ATIQU6&ct=210222&st=sb]

Ransomware's Evolution

Ransomware is a form of malware that encrypts an organization's most important files, rendering them unreadable and unusable. Criminals then demand a ransom payment to decrypt the files in amounts that are often scaled proportionally to the number of systems infected and the data's perceived value.

The present ransomware scourge had humble origins: the first attack was largely experimental in nature, involving an auto-installing trojan circulating on floppy disks and mail-in payments. The basic idea that still underlies today's ransomware activities had already been conceived of—namely, that victims would be willing to pay to avoid losing access to digital information—but the three critical factors necessary for large-scale ransomware campaigns' success had not yet come into being.

Before ransomware attack volumes could reach the epidemic proportions we're currently seeing, there needed to be:

- a means of disseminating malware widely, across organizations' internal networks as well as to multiple victims
- strong file-level encryption
- a solid strategy for collecting anonymous payments

With the Bitcoin boom of the 2010s, the stage was set for ransomware operators to dramatically scale their operations. Transactions became much more difficult to trace, making it easier to target businesses and public sector organizations instead of individual consumers, ask for larger ransoms, and collect payments across international borders.

At the same time, ransomware operators greatly improved their capabilities to orchestrate lateral movement across their victims' IT environments. This means they became better at encrypting not just the contents of a single user's device, but all the computers, servers, and backup file systems that housed an organization's information assets.

The Most Infamous Ransomware Strains in History

CryptoLocker

This strain gained prominence in 2013 as the first large-scale ransomware campaign to demand payment in Bitcoin, leveraging encryption at an unprecedented strength.

WannaCry

WannaCry exploited the EternalBlue vulnerability in Windows software in May, 2017, becoming the first ransomware strain to proliferate in organizational settings rather than targeting individuals. WannaCry was more destructive than profit-oriented.

NotPetya

Arriving just one month after WannaCry, NotPetya exploited the same vulnerability,becoming one of the fastest-propagating pieces of malware in history. NotPetya leveraged strong encryption and was typically very destructive.

Ryuk

First disseminated in early 2018, its authors' goal was simple: get the job done, collect as large of a payment as possible, and do no damage if victims paid up. Ryuk was engineered to seek out and compromise network-connected backups as well as in-use files.

The large-scale ransomware attacks that became prominent about four years ago provided a template that today's organized criminal groups have become expert at following. Ransomware is designed to spread easily and quickly across IT environments and is often disseminated via phishing campaigns that are near-impossible to thwart. Attackers strategize to make it cheaper and easier to pay the ransom than repair the widespread destruction that will otherwise result.

In 2020, ransomware operators began employing double extortion tactics on a widespread basis. They're also bombarding companies' websites with malicious traffic at the same time that their networks are down due to a ransomware attack, a tactic called the high-impact "ransomware plus DDoS" strategy. Today's attacks leverage encryption methods of unprecedented sophistication.

2020's Heavy Hitters: Prominent Ransomware Families of the Past Year

Maze

This strain was chief among the ransomware families that popularized the double-extortion trend. Maze operators collected ransoms that were reportedly as high as $15 million before attackers ceased operations late last year.

Sodinokibi/REvil

First thought to have been written by members of a group formerly responsible for as much as 40% of global ransomware attacks, this campaign spreads through PowerShell and JavaScript. REvil got into the double extortion game early in 2020.

RagnarLocker

This type of ransomware spreads mostly through brute-force attacks that take advantage of weak passwords, and is said to have elicited an $11 million ransom from a company whose name has not been published.

DarkSide

First spotted in August of 2020, DarkSide quickly rose to fame with the success of its attack on the Colonial Pipeline stealing data from its victims as well as encrypting it, with the attackers operating their own leak site.

Anatomy of a Ransomware Attack

The majority of ransomware attacks proceed according to a well-defined pattern. Attacks begin with the initial compromise of an environment, which is usually accomplished by tricking a victim into downloading a malicious file, by exploiting remote desktop protocol (RDP), or by running a malicious script in memory to search for privileged credentials that are exposed. Their next stage includes lateral movement and, today, often incorporates data exfiltration as well. Finally, the attack concludes with the installation and detonation of the ransomware itself.

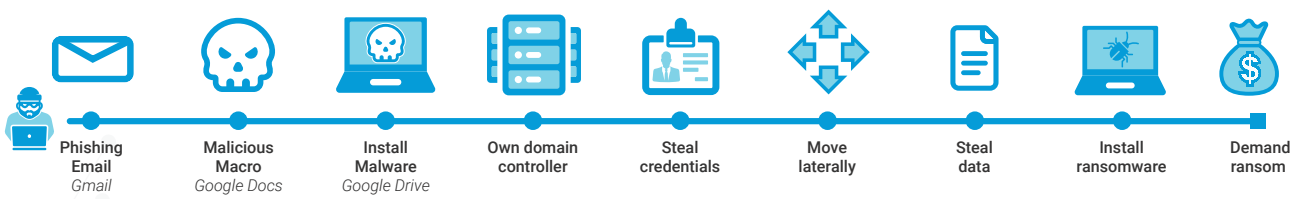

Fig. 1. Sample ransomware attack lifecycle

Ransomware Attack Vectors: Phishing, Drive-by Downloads, Brute Force Attacks, and Stolen Credentials

All ransomware attacks begin with efforts to gain an initial foothold in the victim's environment. This could occur through the delivery of a malicious file, the execution of a malicious script, or the compromise of a user's account credentials. From early 2018 until mid 2020, the percentage of ransomware delivered via an email attachment had been decreasing, but it's now again on the rise.[5] Among the new ransomware families that have proliferated in the past year are several that rely on phishing campaigns to spread the malware. Drive-by downloads remain among the most popular attack vectors, though growing numbers of ransomware attacks are also exploiting known software vulnerabilities. Other attacks leverage new Zero Day exploits.

Ransomware operators are increasingly making use of Secure Sockets Layer (SSL) encryption when delivering their payloads. More than 90% of all internet traffic is now SSL-encrypted, and ransomware attackers want to blend in. In fact, there's been a 500% increase in SSL-encrypted ransomware attacks over the past year.[6]

Lateral Movement: The Key to Enterprise-Wide Devastation

Once attackers have found their initial entry point into an environment, they'll need to identify where their victim's most valuable data is stored and processed. They also must ensure that they're able to compromise all or most of their target's business-critical systems. This is the key step that enables attacks to spread widely within an enterprise computing environment.

Attackers often gather information about victims' infrastructures while terminating processes related to security and monitoring. This way, they're able to maximize the number of files that the ransomware will encrypt and can do so without having their presence detected.

An additional step that's taking place in a significant number of this year's attacks is data exfiltration. This often requires more extensive reconnaissance, as well as the ability to compress and extract large numbers of files.

Stealing Data, Installing Ransomware, Demanding Payment

By the time the victim becomes aware of the incident, it's likely that the bulk of the damage has already been done. Today's ransomware operators are not only encrypting primary data stores, backups, and archives, but they're also wiping out configuration settings and administrative process documentation to make recovery that much harder.

In addition, attackers are more likely to re-target organizations that pay ransoms, but also repeat attacks against those that don't. As many as 80% of companies that admitted to paying up in a recent anonymous survey experienced a second or even multiple subsequent attacks.[7] Organizations that don't pay can be subject to revenge targeting, while those that do are known to be willing to pay large sums of money to criminals to avoid a crisis.

[5] Source: https://www.proofpoint.com/us/blog/security-briefs/ransomware-initial-payload-reemerges-avaddon-philadelphia-mr-robot-and-more
[6] Source: https://info.zscaler.com/resources-industry-reports-state-of-encrypted-attacks
[7] Source: https://searchsecurity.techtarget.com/news/252502519/Repeat-ransomware-attacks-hit-80-of-victims-who-paid-ransoms

Why a Zero Trust Approach is the Answer

The rising numbers and mounting severity of today's attacks provide ample evidence that the prevailing strategies simply aren't working. It's become all-too-common to hear security leaders argue that robust backup and data protection comprise the best defense against ransomware attacks. In fact, this is no defense at all. Instead, it's a reactive approach that assumes that infection is all but inevitable.

Simply planning to restore from backups in case of ransomware infection fails to take into account the high costs associated with the downtime that's fundamental to this process. Ransomware operators select victims that are especially intolerant of downtime, whether this is because its per-minute costs are excessive (as is the case with lost production capacity in manufacturing) or because being inoperative is inherently dangerous (consider the consequences of downtime in healthcare, for instance).

This approach also fails to account for the fact that the ransomware strains currently in circulation are better at finding, targeting, and destroying backups as part of the attack lifecycle than they've ever been in the past. And it ignores the fact that ransomware operators now threaten to publish sensitive data, wage distributed denial of service (DDoS) attacks against critical systems, and use other layered tactics that data restoration does not address.

The reality is that most of today's network defenses are overly focused on the perimeter, but far too lenient when it comes to internal traffic. As a result, they tend to be overly permissive when allowing systems within the corporate environment to communicate with one another. This is a remnant of the legacy assumption that a gateway firewall could keep all malicious traffic out of the network. Firewalls do a good job of monitoring inbound and outbound traffic, but are blind to internal traffic, so they can't detect the lateral movement of attackers. This is an outdated approach that doesn't meet the needs of the modern distributed enterprise, where users, devices and applications are everywhere, not just situated within the confines of the data center.

Next-generation firewalls with individual packet-inspection capabilities can do some of the necessary work, but it's relatively easy for malware to bypass their defenses. What's more, firewalls tend to be slow at identifying malicious traffic: often the payload has already been delivered by the time they flag the connection as malicious.

What is Zero Trust?

Zero trust is a holistic approach to securing modern organizations that assumes that no user or application should be inherently or automatically trusted. Instead, the zero trust mindset begins with the assumption that everything is hostile, and that trust can only be established on the basis of the individual user's identity and the context. Policy is to be enforced everywhere, serving as a gatekeeper across the entirety of the environment.

While digital transformation is boosting business agility and improving access to valuable data, it's also expanding access to the attack surface. Zero trust requires migrating to a model in which security is delivered through the cloud, closer to where business assets and user activities are now centered.

Zero trust is not simply about a single technology like identity or application segmentation. Instead, Zero trust is a strategy, a foundation upon which to build an entire security ecosystem that securely connects users to applications using business policies over the internet.

The Foundation of a Zero Trust Platform

At the heart of a zero trust architecture are four key tenets:

1. **Complete inspection:** Terminate every connection for deep inspection of all traffic for threats and data loss, including SSL.
2. **Least privilege access control:** Determine access policy with granular identity, authentication, and full context on the user, device, app, and content.
3. **Minimize the attack surface:** Establish inside out connections to make apps invisible to the internet and impossible to attack.
4. **Eliminate lateral movement:** Connect to an application, not a network, preventing the lateral spread of adversaries.

How Zero Trust is Changing the Game in the War Against Ransomware

Every ransomware attack involves a complex and intricate sequence of steps. For the attack as a whole to succeed, the attackers must succeed at each one of these points.

If an organization is able to bring its environment into alignment with the principles of zero trust, they'll be able to halt attackers at every stage of the attack lifecycle or kill chain. Should your defenses fail at any one point, you'll still have multiple additional layered protections that will keep the malware from spreading or the attackers from exploring your environment.

Thus, we advocate for a platform-based approach that ensures that policies are enforced whenever and wherever business users connect to technology assets within your ecosystem. There are four primary facets to ransomware prevention:

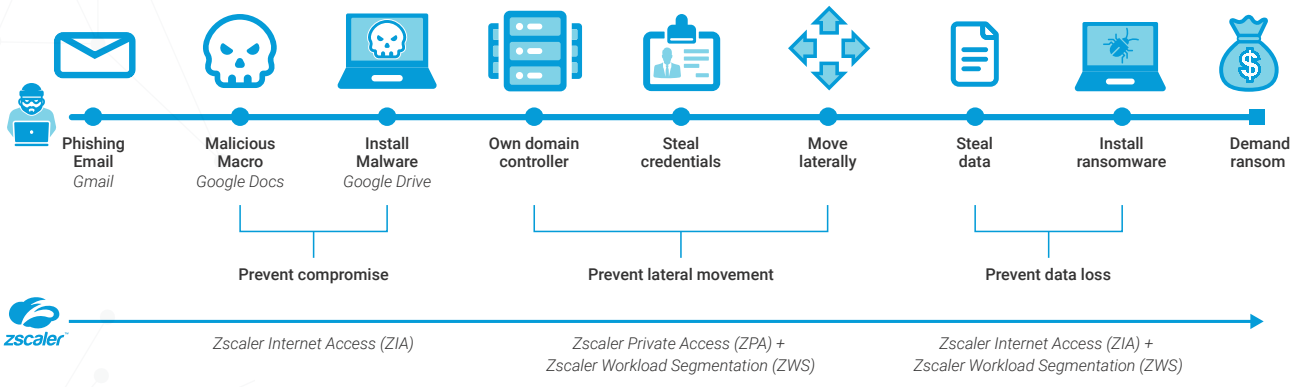

Fig. 2. Defending against ransomware

1. Eliminate the attack surface:

- Applications, user, and device identities should be hidden from the outside world. They should never be visible or discoverable from the public internet.
- Never permit inbound connections to Exchange or web servers located within a network DMZ. If these servers are visible to the outside world, they're open to attackers' attempts at reconnaissance. Once attackers can see them, they'll be able to find and exploit vulnerabilities on those machines, and they'll immediately have a strong foothold within the network.
- Instead, leverage a Zero Trust Network Access service like Zscaler Private Access that enables applications to connect to users via inside-out connectivity. Rather than permitting users to connect to the network, grant individual authorized users access to specific private apps.
- Extend this approach across both managed and unmanaged devices and apply it to all applications, not just web apps.

2. Prevent initial compromise:

- Full visibility is essential, including packet inspection for all traffic. This will stop malicious downloads, block email-borne infections, and halt malware arriving across other channels. You can no longer blindly trust SSL-encrypted traffic.
- Application access should be secured through an exchange. Application access should never occur through the open internet, because this makes it easy for attackers to discover what's in your environment. When applications are accessible only through an exchange, they're invisible to external attackers.
- Leverage sandboxing and AI-powered detection that can discover previously unknown ransomware variants through behavior analysis. With AI-based sandboxing, files are held in quarantine until it's been ascertained that they're safe. This eliminates the threat of zero day-based infections. Without these capabilities, never-before-seen threats will be free to encrypt files on your systems.
- Enforce consistent least-privilege access controls, and constantly monitor for policy, entitlement and configuration gaps.

3. Block lateral movement:

- Use a proxy architecture to connect users and workloads to applications and resources rather than trusting traffic from an internal network or a subnet. This is the cornerstone of the microsegmentation approach, which limits access to applications and resources so that attackers that breach a single one can't cause damage beyond its confines.[15] It's also far less complex than implementing microsegmentation within a traditional network architecture.
- Rely on application segmentation rather than network segmentation. Network segmentation is both inadequately secure and overly complicated to implement and manage.

- Ensure that all connections to your resources are taking place through a platform built on a zero trust architecture such as the Zscaler Zero Trust Exchange, which ensures that all traffic, regardless of user, location, or device type, adheres to all corporate policies when accessing corporate resources, regardless of whether these resources are internal, external, or third-party SaaS.
- Employ active defenses like deception technologies that act as lures to draw attackers away from valuable data assets and serve as tripwires that link your security team to the attackers' presence.

4. Stop data theft:

- Include measures to prevent data exfiltration among your ransomware prevention strategies. This is a necessity because today's "double-extortion" ransomware attacks involve the theft and publication of sensitive data along with its encryption.[17]
- Implement data loss prevention (DLP) controls that can identify and block data leakage or theft by inspecting all northbound traffic (including SSL-encrypted traffic) in real-time.
- Establish "default deny" policies that allow communication only with known-good destinations.
- Use a cloud access security broker (CASB) to enforce granular controls over your cloud applications. An added benefit is that you'll have visibility into what's known as "shadow IT," or unsanctioned cloud app usage.
- Guard against and prevent cloud misconfigurations.

Conclusion: Tackling Ransomware Requires a Comprehensive, Zero Trust Approach

The current ransomware crisis won't abate anytime soon. In fact, if history has anything to teach us, it's that attackers are endlessly innovative, and will likely discover new techniques for bypassing legacy defenses faster than ever.

The Zscaler Zero Trust Exchange provides a comprehensive foundation for your ransomware prevention strategy. Every user connection, to every resource in your environment, is secured individually with full SSL inspection to uncover novel attack tactics as well as familiar ones. An AI-driven sandbox quarantine ensures that unknown threats won't gain access to your environment, and off-net users can simply add a lightweight agent, the Zscaler Client Connector, to their devices to get the same security, policy enforcement and access controls that they'd have if they were working at your headquarters. And, as the Zscaler platform is built on a groundbreaking zero trust architecture, it is impossible for ransomware to spread laterally across the enterprise to infect additional systems or repositories of sensitive information to siphon out for double extortion. Zscaler has helped thousands of customers prevent ransomware attacks from slowing their businesses down while providing superb scalability and exceptional user experiences.

See how Zscaler uses zero trust to deliver unmatched protection for your organization.

Zscaler, Inc.
120 Holger Way
San Jose, CA 95134
+1 408.533.0288
www.zscaler.com