illumio

WHITE PAPER

How to Prevent Ransomware From Becoming a Cyber Disaster

Essential steps to reduce ransomware risk with Zero Trust Segmentation

Ransomware is one of the largest cybersecurity threats every year and particularly today. Every year and particularly today. Every week, we get new disclosures that some strategically important company has been breached, paid a ransom or has had their revenues and customers threatened. For example, roughly 40% of the United States population had difficulty purchasing gasoline after a ransomware attack closed the major pipeline supplier for the East Coast[1]. The President of the United States then mandated new security measures[2] for the U.S. Government and, in a separate directive, recommended that all businesses adopt Zero Trust principles as the only solution for defending against ransomware.

These unprecedented actions point to the size and significance of the disruption to lives and the broader economy.

$30B USD

Estimated cost of ransomware remediation in 2023

Cost of Breach is Skyrocketing

Modern ransomware has several key features that make it difficult to remediate an attack:

- Steals credentials to administrative systems
- Encrypts as much as possible, rendering it unusable
- Destroys backups, so restoration is impossible without paying
- Publicly exposes victims to intensify pressure to pay
- Leaks stolen data, potentially violating contracts or compromising competitiveness
- Threatens victim's customers to increase pressure to pay quickly
- Uses time-based ratchets to encourage fast payment "before the price rises"

These characteristics make dealing with a ransomware attack more costly than ever before. Most small businesses pay the ransom, and then still have cleanup and remediation costs, not to mention the loss of revenue and distraction from other projects. Most analysts expect ransomware-related costs to increase substantially as the criminal business model evolves. On a global basis, reports estimate that remediation costs will exceed $30B USD in 2023[3].

[1] Source: https://www.nbcnews.com/politics/national-security/russian-criminal-group-may-be-responsible-colonial-pipeline-ransomware-attack-n1266793

[2] Source: https://fortune.com/2021/05/14/biden-white-house-cybersecurity-order-ransomware-zero-trust/

[3] Source: https://www.infosecurity-magazine.com/news/ransomware-exceed-30bn-dollars-2023/

The Standard Security Model Doesn’t Help Much

Unfortunately, the security architecture we’ve spent the last 15 years building hardly stops ransomware, and in many cases, ensures it has an easy time doing its dirty work. The idea of a hard perimeter, complete with analogies to medieval castle walls and physical controls, has an undeniable appeal. After all, we do put locks and security guards in front of physical buildings to control access.

But just like most corporate headquarters are open once past the front desk, with no further credential checking to roam the halls, most data centers are no different. Once past the perimeter or laptop security suites, the network will take you anywhere you want to go. And for ransomware, that is everywhere and immediately. Most ransomware can infect an entire network in the order of seconds, passing easily from machine to machine on administrative and common use ports that most machines have “on” by default.

Of course, we have augmented the perimeter over the years with scanners, analysis tools, machine learning, predictive tools, and various endpoint technologies. And yet, the news proves that we still haven’t stopped high-profile ransomware attacks. Many smaller firms suffer silently.

Why Do Attacks Get Through?

In many organizations, the perimeter has effectively dissolved into individual laptops, creating many new problems. With most enterprises working remotely, at home, or in alternate locations, there hasn’t been a uniform corporate perimeter for user devices outside of whatever software can be installed remotely on laptops.

But it’s not just users that have moved out of the data center. Applications are now constituted as distributed services. Some have moved to dynamically instantiated containers. Other application services come from SaaS or cloud-hosted solutions. Most organizations have a default multi-cloud strategy. Microsoft enterprise licenses effectively mandate some use of Azure or Microsoft cloud services, and Amazon is a default location for most. So, when a user clicks on a spurious link and unwittingly becomes affected, it only takes a weakness anywhere in the distributed application hosting environment for it to spread from one environment to another. In all too many cases, this happens in seconds — far faster than most detection tools can function reliably.

Protecting Critical Services, Compute Environments, and User Endpoints

The most important controls to combat ransomware are the ones that are in place before any infection. Most CISOs agree that there is an upper limit of effectiveness on end-user training. The more humans in the organization, the more likely that some compromise will occur.

As a result, many organizations have moved to an “assume breach” posture or started using terms like “post-intrusion security.” After all, even if a laptop is infected, it must have a pathway to transmit its code to another user or server. When those paths are not available and alarms are triggered for policy violations, existing detection tools and teams have the critical seconds, minutes, and hours they need to identify and quarantine the threat.

Breaches may be inevitable, but total infection is not.

Industry attention has focused on Zero Trust frameworks to tighten controls around networks, users, data, devices, and more using the principle of "least privilege." Dating back to the early 1970s, least privilege is the simple assertion that access to networks, systems, accounts, or data should only be the minimum necessary. Zero Trust ensures that only authorized traffic flows are permitted. The President of the United States recently mandated a Zero Trust architecture for most U.S. Government agencies and recommended it for all business and civic organizations.

Zero Trust Segmentation eliminates

90-99%

of connection paths.

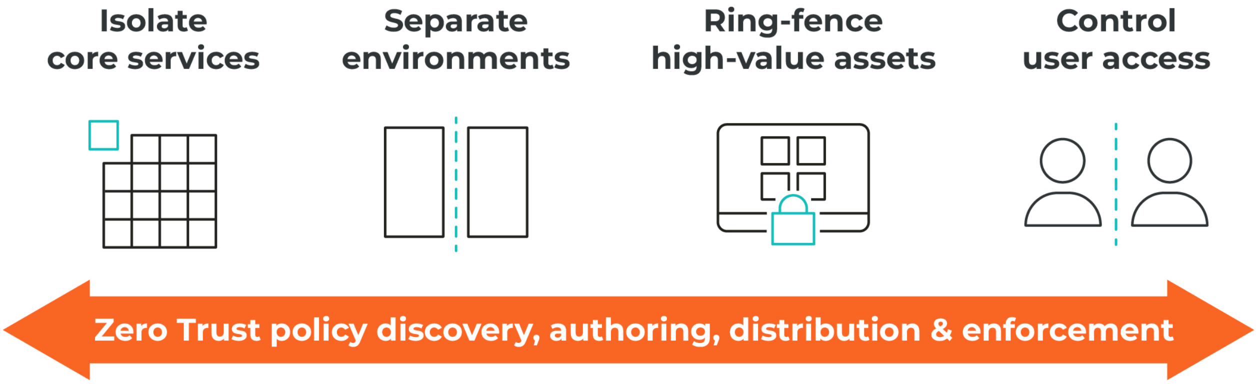
When applied to the network, the wide-open gigabit-speed networks of the past become heavily gated communities where motion across the data center, cloud, or container environments is constrained. Zero Trust Segmentation reduces connectivity to only what is required and typically eliminates 90-99% of the connection paths in the data center. When implemented in a distributed model where every server, OS instance, and container host becomes an enforcement point, every server becomes a sensor. Any violation of the policy or a scan to find open ports immediately raises an alarm, even as access is denied.

It turns out that most ransomware follows the path of least resistance. And over 70% of ransomware attacks use common protocols like Remote Desktop Protocol (RDP) or file sharing technologies like Server Message Block (SMB). These protocols are "on by default" in standard operating systems installations. But interestingly, most users don't use either of them. So, why are these ports even open?

Under the principles of Zero Trust Segmentation, unused ports should not send or receive data. Eliminate the path for infection to spread and it simply cannot. Reducing open ports to the bare minimum necessary results in very few paths out of the infected laptop exist, most of which are not vulnerable to malware exploits. This provides a hint at the best place to start reducing ransomware risk.

A Fast Path to Reducing Ransomware Risk

The fastest way to reduce the risk of ransomware spread before an infection occurs is to remove the paths it uses to propagate. From there, additional controls can be put in place to ensure that other types of attacks also fail.


Isolate core services
Separate environments
Ring-fence high-value assets
Control user access
Zero Trust policy discovery, authoring, distribution & enforcement

The most used "core" services typically connect to every laptop or server instance. Active Directory, monitoring, performance management, and security applications often connect systems to a central brain. It would be unsurprising that these protocols are the first target of most malicious actors. After all, compromising a broadly used protocol or central server with access to every other machine in the environment accelerates the infection across the entire environment.

But, setting the core services aside, other forms of segmentation also help restrict malicious activity of any kind. Dividing systems and users into groups, environments, and locations all help to reduce the possible connections. However, if there are many systems in a Production environment, for example, it makes sense to further reduce the segmentation so that applications are isolated and ring-fenced from each other. If one application is compromised but can't talk to any other application, the potential for spread is radically reduced.

So, if you want a fast start, how quickly could you reduce ransomware risk? There are several important steps to ensure success.

And with the right segmentation technology, it is even possible to apply Zero Trust to every flow — essentially removing the potential for unexpected traffic to find an open port to exploit.

Isolate Core and Management Services

You can reduce risk as quickly as you isolate the core and management services to only their essential use. A group of IT administrators may need to use RDP to administer servers, but the broad user population should not, and the servers should not accept RDP connections from them as a simple example.

Let’s consider the ports that are the greatest risk:

	Example	Risk
Highly connected ports	Core services Management ports Polling/Reporting systems	Broad to total environment exposure
Peer-to-peer ports	RDP, WinRM SMB, RPC, WMI, DCOM Social media	Arbitrary machine-to-machine traffic is normal
Application teams	DB Core services Microsoft applications Common Linux utilities	Many published vulnerabilities

- **Highly connected ports** are the most effective ports to spread an infection, just like a highly connected freeway system is the best way to move goods around a country. Management and core services typically connect to most of the systems in a compute environment. But while every server may need to talk to the core service host system, no server should talk to another server on that core service port. Eliminating lateral movement on the core services can be as simple as closing these ports to any destination besides the legitimate host of that service.
- **Peer-to-peer protocols** intend to make it possible for any machine to contact any other machine in a well-defined way for the purpose of simple data exchange. That would sound like an ideal target if one was writing malicious code or ransomware. If every server or laptop is expecting to talk to any other server, then any server can infect any other — the worst possible situation. If traffic from ‘anywhere to anywhere’ is normal, how would one detect anomalies? But again, in the real world, these protocols should be heavily restricted.

  Many organizations require administrators to first connect to a secure system — a so-called “jump box” — and then to use that as the only acceptable source for Remote Desktop Connections. When reinforced by Zero Trust Segmentation technology, a wide-open peer-to-peer protocol becomes restricted to a single system, radically reducing the potential for lateral spread.
- **Well-known ports** have both the advantage and the disadvantage of being widely familiar. Because these ports have existed for many years or even decades, they can form a dependable way to exchange information. But this staying power also means that malicious actors have had years to study them and find vulnerabilities and exploits. Given that no patching regimen eliminates old software, hackers can play the odds and usually find a well-known system vulnerable to attack.

  When Zero Trust Segmentation closes these ports or limits them to only specific, intentional paths, it is no longer possible for malicious intent to use them.

Taken together, reducing the connection possibilities for highly connected ports, peer-to-peer protocols, and well-known ports radically minimizes the potential for malware spread.

Most attacks use known vulnerabilities and known ports to move around, and Zero Trust Segmentation eliminates these paths without breaking existing application traffic. Because these ports are common and typically involve a handful of legitimate destinations in the data center, policy development moves quickly. Often, a few minutes is all that it takes to put a Zero Trust policy in place.

Tighten Environmental Controls

After taking control of core and management services, tightening environmental controls should be next on the list. After all, most organizations already have some distinction between Development, Test, and Production environments. But beyond reinforcing these simple controls, it is often possible to identify additional easy boundaries to put in place and to tighten IT/OT controls.

Despite having a firewall between DEV and PROD environments, most organizations could tighten their default controls significantly. Over time, rulesets drift, and often broad rules are implemented for a particular use case and then forgotten or not later narrowed. Because IT is the primary user in the DEV environment and also the administrator of the PROD environment, it is common for access to open over time for ease-of-use considerations. Using a quality Zero Trust Segmentation solution, the visualizations will quickly categorize and display all DEV-to-PROD flows for analysis. Instead of combing through a large firewall rule base, it becomes easy to identify the truly legitimate flows and simply block everything else. Eliminating the excess pathways cleans up administrative access, but also reduces the possibility of infection spreading from one environment to another.

But even if you confined your interest to the PROD environment or a portion of a cloud deployment, you could likely identify several other easy boundaries to place. Database traffic usually sits behind most applications and yet only flows in narrow paths — a single port or group of ports. Placing boundaries so that database clusters are tightly controlled for administrative and data access can happen quickly, and help protect one of the most valuable targets. Similar efforts can be made for “crown jewel applications,” whose communication patterns are broad but well defined. Without knowing all the details, it is usually possible to isolate these applications from the broader data center population.

Additionally, tightening controls around OT and embedded systems doesn’t have to be difficult or cumbersome. There are many embedded systems in a typical data center — VPN concentrators, storage arrays, and appliances of various kinds. Placing simple boundaries that restrict these devices to their necessary ports and that restrict client machines to only listening for the expected appliance removes significant exposure. IP cameras, door badge readers, printers, and the whole world of OT devices should likewise be placed behind a boundary such that they communicate only to their control plane server and lack the ability to communicate in any way with the broader data center population.

Control Endpoint Connectivity to Limit Infection

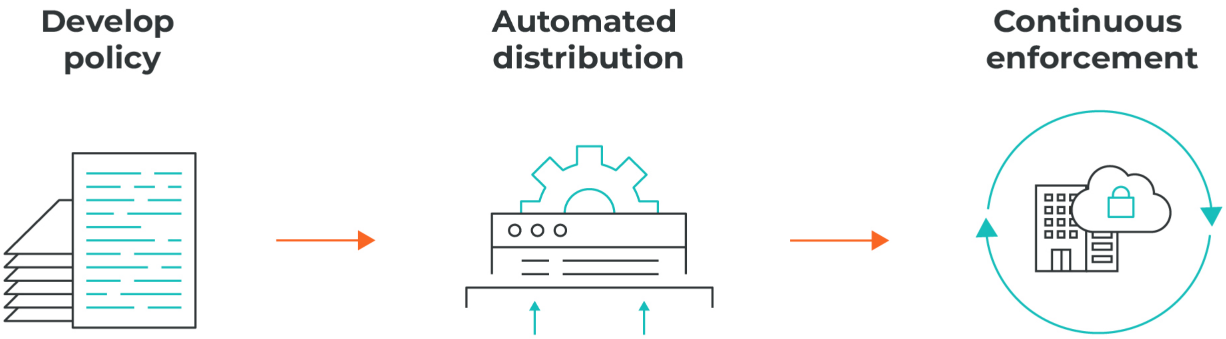
While the data center is the ultimate target of most ransomware, infections normally take hold on user endpoints first.

There are several key strategies to reducing the risk of ransomware spread from the start:

- Close all unused ports: Laptop operating systems typically have a very open profile on any domain joined interface. And yet most users use a very narrow set of applications and services. Closing the unused ports eliminates them as a potential vector for spread.
- Eliminate peer-to-peer spread: Controlling peer-to-peer protocols at the network edge eliminates many common ransomware propagation strategies. RDP, SMB and other similar protocols can be blocked inbound and outbound.
- Control access to cloud and data center assets: Like inside the data center, once the basic services are secured and the bulk of open ports are closed, the next logical step is to ensure that users can only send and receive traffic to authorized servers and that servers only accept traffic from authorized users. Zero Trust Segmentation can coordinate policy between endpoints and servers to make sure the policy can be enforced universally.
- Tie Zero Trust policies to identity: User Identity provides the best foundation for controlling user access into data center or cloud locations. Prior to login, there is no reason for servers and applications to be open to a given laptop. Upon login, access can be opened to only the necessary resources. When the user logs out, access permissions can be revoked. In contrast to the "always on, always open" default approach, identity-based segmentation ensures that ports are only open when needed.

Automating and Scaling Zero Trust Protection

Ransomware doesn't sleep, and neither should your Zero Trust infrastructure. Policy automation provides continuous enforcement of defined policies. Once an organization develops policy for some or all of the scenarios mentioned above, that policy needs to be distributed across the environment to every server, virtual machine, endpoint, container and cloud system. Automation can ensure that all systems continuously enforce the defined policy — even when IP addresses change, new application instances instantiate, or are removed.



Look for a segmentation solution that has the following key capabilities to deliver automation and scale:

Build Zero Trust protection into every workload

The best Zero Trust Segmentation accompanies the workload throughout its entire life. Build Zero Trust into base or "golden" images and every machine built from the template will be ready to receive Zero Trust policy. Update build automation to ensure that Zero Trust Segmentation is part of every automation run and even the most volatile and dynamic systems will conform to policy at all times.

Provide "Segmentation as a Service" to DevOps

Zero Trust Segmentation policy engines typically provide full API access and full abstraction from the underlying infrastructure. This means that segmentation works just like any other cloud application service. DevOps code can instantiate systems, request segmentation services and instantly conform to published policy — all in seconds. When Zero Trust Segmentation is just another simple service, the barrier to adoption falls and the overall posture of the organization improves.

Reduce policy management overhead

Ransomware moves much too fast for manually adjusted firewall rules to have any effect. Zero Trust Segmentation policy must be constantly and continuously up-to-date. An automated Zero Trust policy doesn't require the manual overhead of a traditional firewall policy, making it far easier to administer.

Eliminate network dependencies and manual workflows

Automation only works as well as its targets are abstracted from the physical infrastructure. When abstracted from network addresses and constructs, segmentation can be fully automated — just like server builds. Zero Trust Segmentation uses labels taken from the names that an organization already uses, making it fast to tie into existing automation. Additionally, by using firewalls already present in operating systems and container hosts, there's no dependence on network architecture or technology, so any arbitrary segmentation is easy without changing a single network device or configuration.

Developing Confidence in the Face of Threats

It's easy to come up with all kinds of complicated multi-vendor strategies to combat different aspects of a ransomware attack. But focusing on the fundamentals always delivers solid results — malware can't spread where there isn't a network path it understands.

These results don't have to take a long time or a lot of effort to accomplish. Some successes using Zero Trust Segmentation include:

- 11,000 systems under full Zero Trust policy in three months to pass an audit
- 40,000 systems secured under full DevOps automation, including policy and enforcement
- Isolate $1T/day of financial transactions at a single bank
- Secure PII of every mortgage holder in the United States
- 145,000 systems secured at global enterprise

Ransomware is one of the most visible threats facing infrastructure and security teams today. But it is possible to take immediate action to stop an infection from turning into a big event. Users may click on links, but when their laptops are well isolated, core and management ports are locked down, and the environmental controls are tightened, ransomware simply can't move around the data center.

Best of all, Zero Trust Segmentation doesn't work only for ransomware — it eliminates lateral movement for any human or software acting with malicious intent. Zero Trust Segmentation doesn't rely on detection or analysis capability. It's a proactive control that works all the time, kept continuously active and accurate with automation.

How to Get Started

The Illumio Zero Trust Segmentation Platform is the industry's first platform for breach containment. Delivering real-time visibility, a radically simple policy creation engine, and automated segmentation and enforcement in minutes, Illumio stops ransomware and breaches from spreading across the hybrid attack surface.

Visit www.illumio.com/products for more information or try The Illumio Experience.

About Illumio

Illumio, the most comprehensive Zero Trust solution for ransomware and breach containment, protects organizations from cyber disasters and enables operational resilience without complexity. By visualizing traffic flows and automatically setting segmentation policies, the Illumio Zero Trust Segmentation Platform reduces unnecessary lateral movement across the multi-cloud and hybrid infrastructure, protecting critical resources and preventing the spread of cyberattacks.