GUIDE TO ZERO TRUST WITH ISTIO (AMBIENT) SERVICE MESH

(Based on SP 800-207: Zero Trust Architecture by NIST and NCCoE and CNCF Istio Project)

Table of Contents

01

Introduction to Zero Trust Network

Zero trust (ZT) is a 'never-trust-always-verify' framework used by IT organizations to increase defenses against external attack to steal sensitive data, private information, and resources. IT and enterprise architects use zero trust architecture (ZTA) based on zero trust principles to design and implement the robust and resilient infrastructure. When zero trust is applied to protect data-in-transit by securing network and communication it is called Zero Trust Network (ZTN).

The idea of zero trust in networks is getting famous among enterprises after 2020, because of the following drivers:

- Increase remote users and workforce with BYOD
- Rise in cloud-based assets
- De-globalization and cyber threats across the world
- Strict focus on compliance and regulations to secure network
- Adoption of Kubernetes container orchestration platform

ZTN can be achieved by removing or reducing trust on users and applications on assets or resources irrespective of the network location. With growing microservices in hybrid cloud, it can be a very daunting task to choose various applications to implement. The best approach to achieve ZTN is using service mesh.

Istio, an open source and widely used service mesh, used to manage network and security for cloud-native applications. In Sept, 2022, Istio project released- Ambient mesh- a modified and side-car less data plane for Istio developed for enterprises that want to deploy mTLS and other security features first and seek to deploy an advanced network later. We will discuss various pillars of zero trust architecture and how to achieve it using Istio (ambient) mesh.

Top 10 pillars of Zero Trust Network

As per SP 800-207 laid out by the US governmental organizations National Institute of Standards and Technology (NIST) and National Cybersecurity Centre of Excellence (NCCoE), and our decades of expertise wrt implementing zero trust at a large enterprise, we propose 10 pillars security and compliance IT managers should consider achieving Zero Trust Network (ZTN).

1. Identity using Authz/Authz
2. Secure channel using mTLS-based communication
3. Certification management
4. RBAC, Multitenancy and Isolation
5. Whitelisting trusted source
6. FIPS and SOC-2 compliance
7. Web-Application Firewall
8. Data loss prevention
9. Vault key/secret management
10. Multicluster visibility

Top 10 pillars of Zero Trust Network

1.Identity using Authn/Authz

Identity means validating the digital identity of a user over their usage of resources such as web applications, APIs, platforms, devices, or databases. A user can be a human- customer, employee, consultant member, participant, or a machine- an application, an API call, hardware devices, etc. The identity with of a user along his permission to use resources can be verified and validated using the authentication and authorization (Authn/AuthZ) mechanisms. Apart from identity management the security requirements of organizations to deal with multiple microservices, would involve granular controls for user and applications, compliance standards, RBAC, etc. In real life, IAM can be really complicated to achieve

2.Secure channel using mTLS-based communication

Mutual Transport Layer Security (mTLS) is a method for authentication between two parties connected over a network. mTLS-based communication is highly secured between two parties (say client and server) as each application authenticates itself first using X.509 certificates and the communication happens based on private keys which also rotate in regular periods. mTLS is considered the successor of the Secure Socket Layer (SSL).

Top 10 pillars of Zero Trust Network

3. Certification management

Securing the connection between the two parties is one part, while certificate management and rotation is an ongoing maintenance part. In cases of regular upgradation of security policies, or security breaches, the old certificate will not be valid. Cloud architects and platform engineers should think of an effective way for administrators and Ops teams to easily rotate SSL or SAML certificates, generate private keys, and distribute them among all the microservices.

4. RBAC, Multitenancy and Isolation

It is an ongoing task to allow or deny users to read/write/delete permission to various resources. Necessary controls should be in place to implement granular policies for role-based access controls (RBAC). Large organizations typically would need to create a dedicated workspace for various projects or platform teams, and there should be provisions to practice multi-tenancy as part of security measures.

5. Whitelisting trusted source

Perhaps the simplest ‘cybersecurity measures’ practiced by security engineers, Whitelisting involves giving administrator-approved IPs and application access to a system. This is particularly very helpful in recent scenarios of BYOD, where certain applications are allowed to login resources in the VPN.

Top 10 pillars of Zero Trust Network

6. FIPS and SOC-2 compliance (for the US-based companies)

The US-based agencies such as NIST and the American Institute of Certified Public Accountants (AICPA) provide guidance and FIPS and SOC regulations for every IT organization. As per the NIST regulation, all non-military, govt agencies and vendors must comply with Federal Information Processing Standards (FIPS) standards. Similarly, System and Organization Controls (SOC) standards specify the way service organizations should handle customer data; it covers 5 major aspects- security, availability, processing integrity, confidentiality, and privacy of customer data. So for any company operating out of North America, complying with compliance such as FIPS and SOC-2 is important.

7. Web-Application Firewall

Web application firewall (WAF) helps protect web applications from attacks such as DOS attacks, SQL injection attacks, cross-forgery, etc. WAF is an L7 protocol that acts as a shield for web applications, and network engineers can create various rules and policies such as traffic filtration to protect against vulnerabilities in the application.

Top 10 pillars of Zero Trust Network

8. Data loss prevention

Data Loss Prevention (DLP) is the practice of detecting and preventing data breaches, exfiltration, or unwanted destruction of private information. DLP is also a part of SOC-2 compliance. The primary aim of the DLP act is to prevent the illegal transfer of data outside organizational boundaries. The network team and cloud engineers should focus on building a system that is robust to threats from malicious insiders or external ransomware.

9. Secret management

In Kubernetes, many services inside or outside the cluster talk to each other using secrets. Developers should practice proper secret management- a practice of storing sensitive data such as passwords, keys, and tokens, in a secure environment such as Vault with strict access controls

.

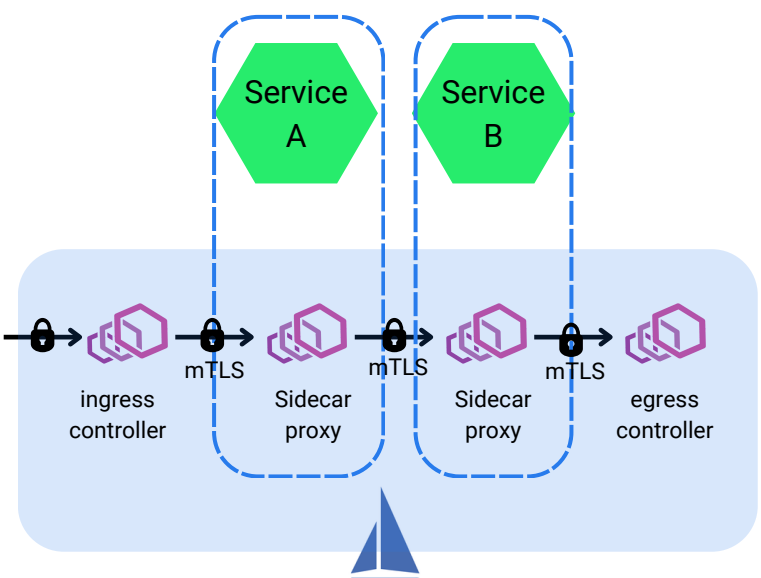
10. Multicluster visibility

IT organizations using cloud and containers profusely should have a central plane for multicluster visibility. They should be able to see workloads, resources, and infrastructure such as ingress and load balancers along with the health and performance status across all the namespace and cluster. SREs should be empowered with real-time logs and metrics aggregation and analysis to reach any situation faster with quick diagnosis.

How Istio ambient mesh can help achieve Zero Trust Network

Istio service mesh is a powerful software to enable zero trust by enabling authentication, authorization, and audit using mTLS and identity controls. Platform teams, and cloud architects of large organizations have implemented security using Istio. To implement security, Istio involves the following components- a certificate authority (CA) for key management, API to distribute Authn/Authz policies to proxies, Policy Enforcement Points (PEPs) implemented using side-cars (Envoy proxies), and extensions to manage telemetry.

Although achieving zero trust using Istio is straightforward, the side-car implementation (refer the image below) of Istio is usually very computationally expensive and hard to maintain; so the project has released a new version called Istio 'ambient' mesh.



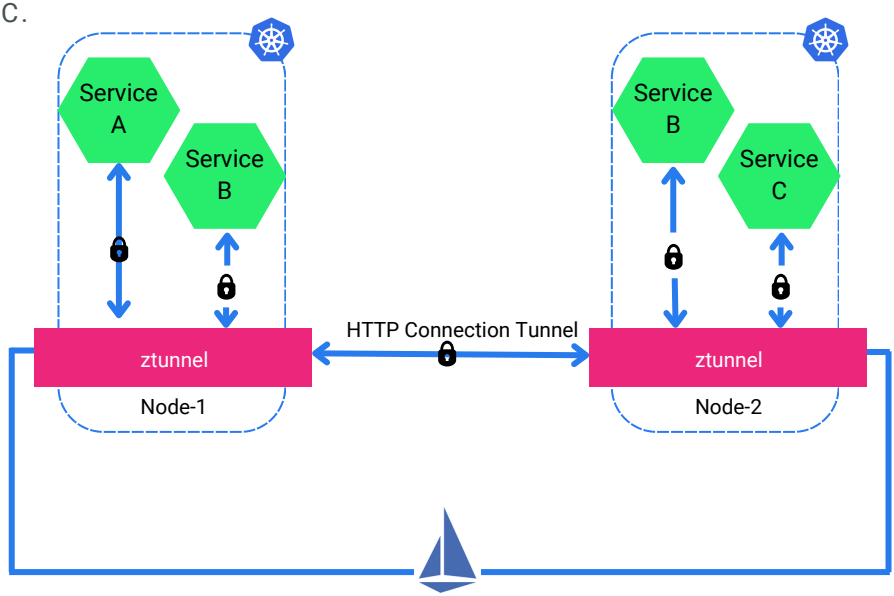
Istio ‘ambient’ mesh provides a lightweight data plane that does not require side-car injection with any microservices. Ambient mesh has distinguished layers in the data plane- **secure overlay layer** and **L7 processing layer** which are designed to implement Istio sequentially in a phase-wise manner and tackle security concerns first.

- **Secure overlay layer** (also known as **zero-trust tunnel tunnelnel**)is an L4 processing layer designed to implement TCP routing and had zero-trust rust security for traffic such as mTLS, Authentication, and Authorization policies.
- **L7 processing layer** (also known as **waypoint proxy**) is designed to handle complex traffic management functionalities such as HTTP routing, circuit breaking, chaos engineering, reties, timeouts, rate limiting, etc, and handle granular Authn/Authz policy implementation.

Ztunnel for secure connection and authentication of services

Ztunnel is an agent, primarily a rust-based proxy, whose responsibility is to securely connect and authenticate elements within the mesh. One can deploy ztunnel as a DaemonSet workload resource on a Kubernetes cluster. Ztunnel is a dedicated L4 technology and is deployed per node in a cluster. The idea is ztunnel will be shared among all the workloads in a node it is deployed to. The ztunnel leverages leverages Kubernetes CNI to establish connections between workloads, secure communication using mTLS, collect HTTP metrics, access logs, etc.

And all the ztunnels are connected with each other using HTTP protocol (refer the image below). If Service A wants to pass data to another Service C in another node, then the ztunnel of node-1 will send HTTP connection requests (over mTLS) to the ztunnel of node-2. Once a TCP connection is established between the ztunnel, the data packets can be transported securely to Service C.



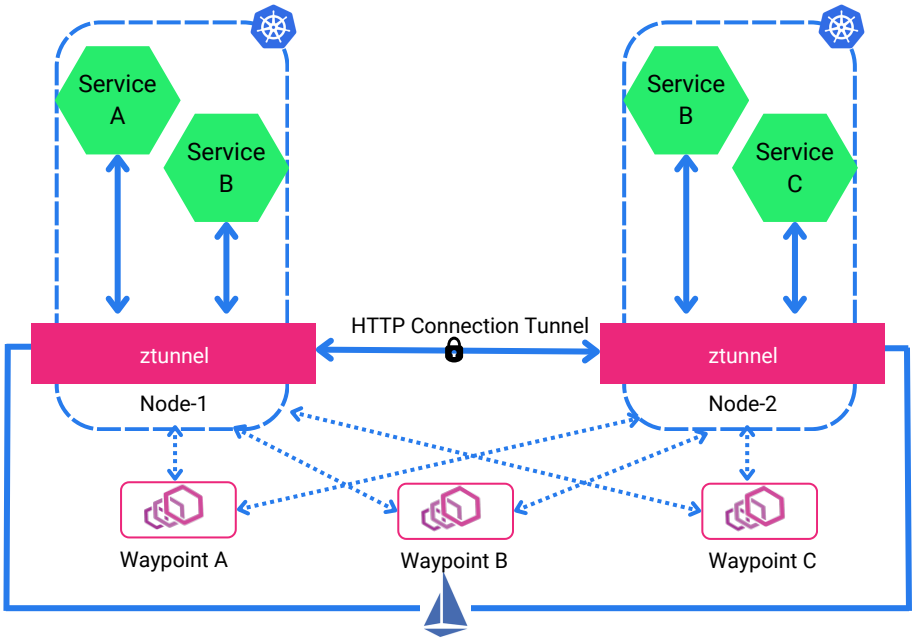
Such connections between ztunnel is referred as HBONE (HTTP-Based Overlay Network Environment). These three important benefits of using ztunnel or a node-level proxy:

1. Phase-1 implementation of secured communication (using mTLS) for all your microservices will be fast. Simple authentication and authorization policies can be defined at node-level.
2. Maintenance such as version upgrades or CVE patching to node-level proxy will be easier and faster.
3. Whenever teams have implemented phase-1: Security of services, they can implement phase-2: Network management of microservices. In the phase-2 they can create sophisticated traffic and security policies by using L7 proxy or waypoint proxy.

Waypoint proxy for network management and telemetry

Waypoint proxies are basically Envoy proxies, used to implement L7 traffic management capabilities in Istio ambient mesh. Based on the header and credentials, the proxy is capable of applying advanced networking policies such as- circuit breaking, traffic shaping, splitting, retries, fault injection, etc. Waypoint proxy also helps in achieving granular authorization policies for Role-based access control (RBAC) or Attribute-based access control (ABAC).

Waypoint or L7 proxy is deployed into a cluster per identity/workload type. If there are 5 services in a cluster, you can deploy 5 waypoint proxies to handle communication at the application level for each service. One can scale up the proxies as per the load. Coming to the deployment of the waypoint proxy (which is essentially an Envoy proxy) is not installed in a side-car fashion- deployed to each pod of a service. Envoy was installed as a container in each of the pods of a service. And in case of breach of an app (essentially a pod), all the sensitive information such as token, keys, etc could be stolen from the proxy.
However, in case of waypoint proxy, deployed at only service level, breach of an application cannot imply the access to secrets in the proxy.



In the above diagram, waypoint proxies are deployed per service and can be seen as individual gateways or policy enforcement points (PEPs) per service. Note: One can configure Ambient mesh to have multiple services configured to a single waypoint proxy, but to contain security blast radius and improve lifecycle management operations the ideal way of mapping the proxy to service is 1:1.

Istio control plane for managing and observing the network security

Implementation of enterprise-wide security from central control plane

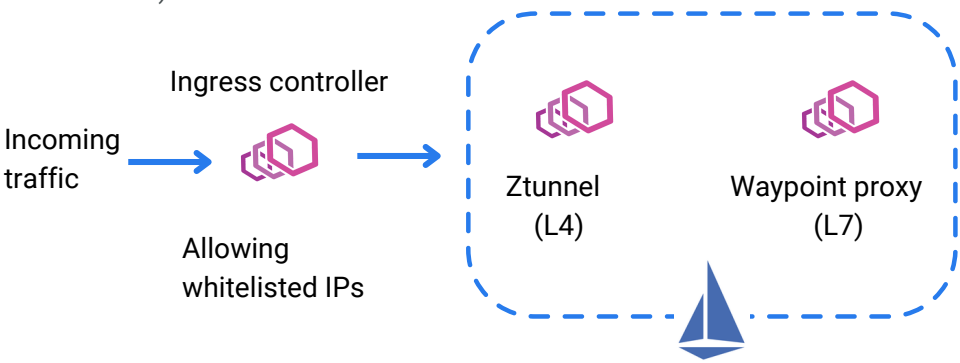
The data plane can be handled from the Istio central plane to push the define and declare security and network policies for each node and each workload through ztunnel and waypoint proxy. Istio is capable of integrating with 3rd party authentication standards such as Okta, LDAP, SAML, SSO providers, etc.

Integrations

Istio control plane also allows the ability to create granular access controls, RBAC and multi-tenancy policies for all the inbound traffic (through waypoint proxies).

Securing traffic at the edge with whitelisting.

Security policies can be implemented at the edge and to the ingress traffic. Large enterprises can use the Istio control plane to whitelist IPs for using certain services. This is very useful while implementing network access controls, or remote access enablement, or beta testing (allowing a limited set of tester of beta-customers to use your services and provide feedback).



Automated certificate management with Istio agents

Security managers or platform engineers can provision stronger identities to every workload with X.509 certificate. The Istio control plane acts as the certificate authority and issues certificates to the Istio agents (running alongside each proxy). The control plane automates the key and the certificate management at scale with the help of the Istio agents. The idea is whenever a Kubernetes workload would start, Envoy proxy would seek the certificate and key from Istio agent in the same container.

Istio agent will be responsible for monitoring the expiration of the certificate and accordingly, would rotate with the new certificate from the control plane. Note that the root certificate can be kept in Vault rather than stored in the same local PC.

Enabling WAF for Multicloud apps with Istio

Cloud providers such as firewalls to protect web applications from common exploits at the edge. A few common ones are AWS WAF, Azure WAF and Google Cloud Armor to provide defense against SQL injection, DDOS attacks, and cross-site scripting (XSS), at the edge. All the firewall can be applied in front of Istio ingress gateway before traffic enters the mesh.

Data loss prevention using Istio network policies

You can avoid any sensitive information such as username, tokens, and financial transaction data that are not logged or leaked using Istio network policy rules. The data loss prevention (DLP) rules can be defined in HTTP listeners or virtual service objects of Istio. The DLP rules can be applied to mask access logs, or any output data with potentially sensitive information.

Gain multicluster visibility

Similarly, Istio allows security managers to monitor and measure the integrity and security posture of all microservices. Istio generates runtime telemetry to help network administrators, SREs, and DevOps to constantly track the behavior and performance of services across the organizations. The control plane emits metrics, traffic flows and service dependencies (using Kiali), which is essential to understanding, and reacting to security incidents.

Another noteworthy point is that the Istio ‘ambient’ mode will have side-car patterns as well. In case you want to have implemented the Istio ‘side-car’ pattern and now want to implement the ambient mesh, the control plane will support both versions. This is the best case for organizations that want to gain multicluster and multilcoud visibility in a single plane.

Next Step

The idea of attaining zero trust network (ZTN) is to secure a network with verification of services and users in each transaction, attain 360 degree visibility for faster reaction in case of security breaches, and have a fault tolerant system for more resilience.

With ZTN, security teams can eliminate the risk of stealing the data or resources from the network. On the other hand, ZTN enables end-customers to get a consistent and secured experience from anywhere, anytime and any device.

Istio service mesh is essential and go-to software to achieve zero trust network and secure data-in-transit. The new side-care-less dataplane of Istio ‘ambient’ mode makes it a more compelling, hassle-free and computation inexpensive option to implement a service mesh.

With Istio ‘ambient’ mesh you can check almost all top pillars for ZTN such as- mTLS, Authn/Authz policy, securing channel, certificate management, RBAC and multitenancy, whitelisting, Web-app firewall, data loss prevention, secret management, and multicluster visibility.

Authors

Ravi Verma, CTO, IMESH

Ravi, a technology visionary, brings 12+ years of experience in software development and cloud architecture in enterprise software. He has led R&D divisions at Samsung and GE Healthcare and architected high-performance, secure and scalable systems for Baxter and Aricent. His passion and interest lie in network and security. Ravi frequently discusses open-source technologies such as Kubernetes, Istio, and Envoy Proxy from the CNCF landscape.

Debasree Panda, CEO, IMESH

Debasree understands customer pain points in cloud and microservice architecture. Previously, he led product marketing and market research teams at Digitate and OpsMx, where he had created a multi-million dollar sales pipeline. He has helped open-source solution providers- Tetrate, OtterTune, and Devtron- design GTM from scratch and achieve product-led growth. He firmly believes serendipity happens to diligent and righteous people.

About IMESH

IMESH offers Kubernetes-native application network and security platform to manage multi-cloud and hybrid cloud environments. The IMESH platform is built on top of Istio service mesh and Envoy API gateway and helps cloud, platform and security teams to make Kubernetes application more secure, manageable, and reliable.

Visit: https://imesh.ai/
email: contact@imesh.ai