Zero Trust Gateway

WHITEPAPER

Table of Contents

Overview 3

Use Cases 3

Benefits 4

Deployment Options 6

Creating Zero Trust Gateways 6

Centralized Model with AWS Transit Gateway 8

Hybrid Model with TGW Centralized Endpoints & Workload VPC Endpoints 9

Decentralized Model with Isolated Workload VPCs 10

ZTGW Service Architecture 11

ZTGW Secure Design 13

Advantages 14

Ease of Deployment 14

Reduced Operational Burden 14

Automatic Scaling 15

Cost Reduction & Consolidation 15

Resources 19

Overview

The Zscaler Zero Trust Gateway (ZTGW) service allows customers to secure their workload traffic within AWS with minimal effort. Zscaler manages the deployment, management, and maintenance of the security infrastructure in AWS, enabling rapid cloud environment security. This fully managed service eliminates the need for customers to deploy or manage connector VMs, providing a comprehensive, highly available, and scalable solution through native cloud technologies. Customers can then focus on their business without concerns about scalability or infrastructure management.

Use Cases

The ZTGW service is an integral component of Zscaler's Zero Trust Cloud offering. This suite of solutions is dedicated to advancing innovation, technology, and specialized solutions for network security pertaining to workloads and servers within the public cloud environment.

To elaborate on the capabilities, consider the following use cases:

- **Secure Internet Egress:** ZTGW secures egress traffic from AWS VPCs, replacing web proxies and egress firewalls. It connects all egress traffic to the Zero Trust Exchange for comprehensive visibility, control, inspection, threat protection, and data protection for all AWS services and workloads within your VPCs.
- **Secure Ingress Traffic:** ZTGW secures incoming traffic to publicly facing applications. It enables Layer 4 based stateful rules to protect applications from malicious attacks.
- **Private Application Connectivity:** It provides private application connectivity across regions, to other clouds (Azure, GCP, OCI, etc.), and to on-premise data centers without granting network access. This extends Zscaler Private Access (ZPA) to public workloads.
- **Secure East-West Segmentation:** ZTGW facilitates east-west macrosegmentation between VPCs and DirectConnect, allowing local enforcement of Layer 4 rules. This capability eliminates the need for traditional firewalls for traffic control, offering potential cost savings.
- **Secure Inbound Controls:** ZTGW enables organizations to utilize Layer 4 rules to control traffic between on-premise data centers over AWS Direct Connect and AWS, along with reducing the need for Layer 4 firewalls for inbound connectivity to applications access from the internet
- **Secure Private Connectivity:** ZTGW provides controlled & secure network traffic between data centers and multi-cloud environments utilizing AWS DirectConnect and Azure Express Route.
- **Policy-Based Forwarding:** Better control on cloud egress traffic using criteria such as source/destination IPs, domains, and user-defined tags
- **Static IP Addresses:** Leverage static IP addresses for egress traffic from Zscaler that are dedicated to your Zscaler tenant

Benefits

ZTGW offers significant benefits to enterprises, providing a compelling model for an improved overall experience. Zscaler assumes full responsibility for managing the infrastructure and associated costs, further reducing operational overhead for organizations.

SMART CONFIGURATION

Zscaler handles all infrastructure configurations, quickly activating resources and securing public cloud traffic within minutes, without extensive preparation.

RESOURCE LIFECYCLE MANAGEMENT/ WORKFLOW CONTROL

Essential for public clouds, this streamlines resource management from creation to decommissioning, focusing on cost optimization through metric tracking and responsive action.

AUTOMATIC SCALING

Zscaler leverages native public cloud adaptive scaling, eliminating customer concerns about traffic volume, throughput, or bandwidth. This frees customers from managing, configuring, or troubleshooting, unlike resources in their own environments.

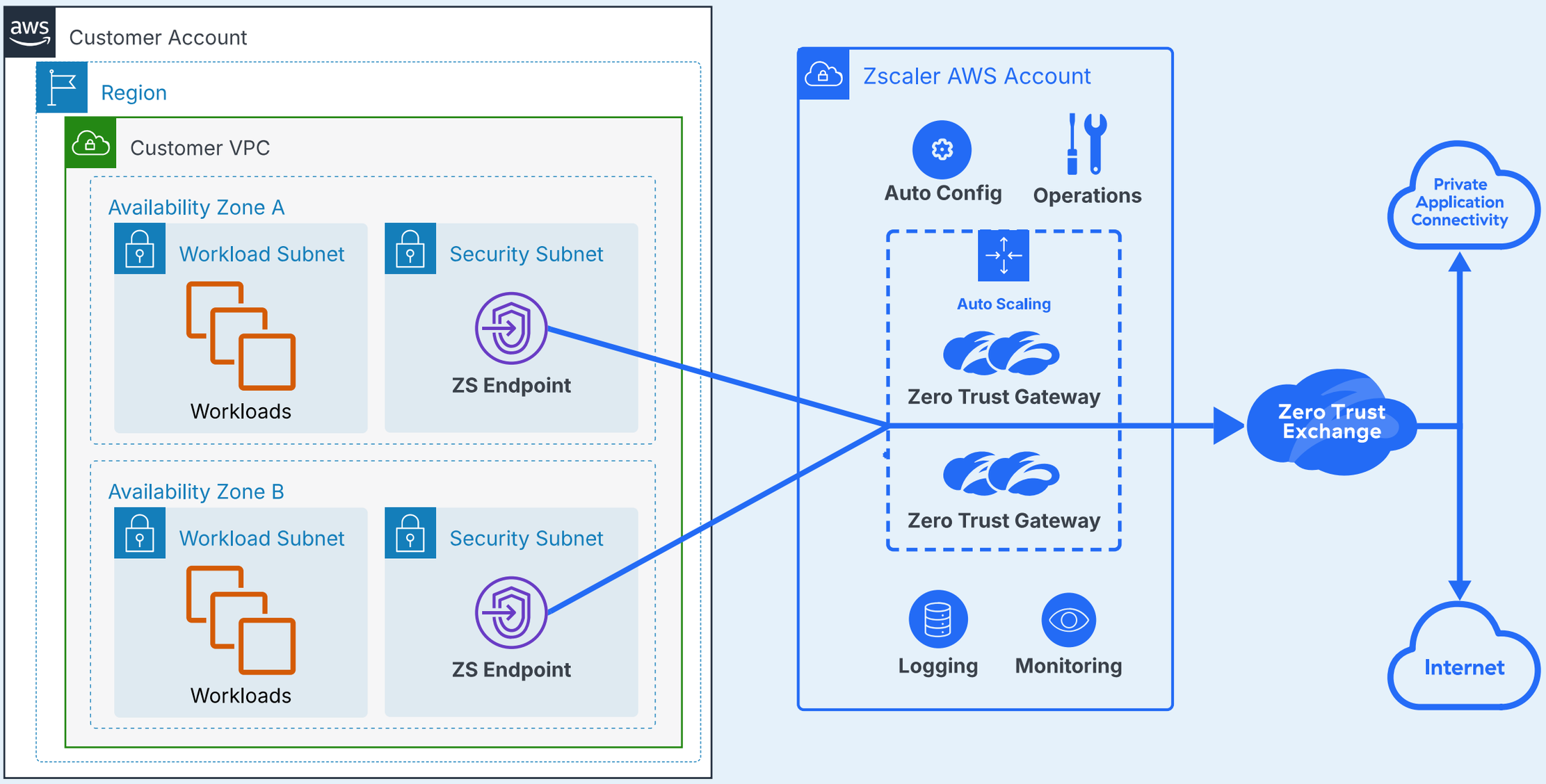
MONITORING/VISIBILITY

While public clouds often have limited default logging and monitoring, Zscaler provides detailed logs for in-depth visibility. It also incorporates native adaptive scaling for events and allows monitoring of every packet, ensuring comprehensive insights and control.

CENTRALIZED MANAGEMENT/LOGGING

The service offers a single portal for centralized multi-cloud logging, simplifying management and enabling real-time monitoring, troubleshooting, and faster incident response. This also streamlines reporting and audit processes across various cloud deployments.

INTRODUCING ZTGW FOR AWS. SIMPLY CONNECT AWS VPC ENDPOINTS



The service is comprehensively managed by Zscaler. Consistent with the broader Zero Trust Exchange, customers are not required to manage the underlying infrastructure. This represents an expansion of Zscaler's demonstrated capability, cultivated over 15+ years, in managing the world's largest security cloud. Our extensive experience and expertise are now being applied to extend this cloud service to facilitate connectivity from AWS via ZTGW.

Regardless of whether your organization has standardized on AWS Transit Gateways, AWS Cloud WAN, Isolated VPCs, or a combination thereof, the ZTGW can facilitate connectivity. The deployment strategy is primarily determined by the following aspects (see next page).

Deployment Options

While this document does not serve as a deployment guide, it explains the concept of a ZTGW from Zscaler's perspective. Subsequent sections will delve into AWS network topologies.

Fortunately, the Zscaler ZTGW does not necessitate alterations to your existing AWS network topology. We accommodate all prevalent topologies and will present illustrative diagrams to visualize these configurations. Given the rapid pace of innovation within public cloud environments, additional options may become available in the future. Should you have any inquiries, please do not hesitate to contact your AWS and Zscaler teams to discuss potential solutions not covered herein.

ZTGW quantity depends on AWS Regions, Availability Zones, and environmental separations. A single region with two AZs, without Prod/Non-Prod VPC distinction, needs two ZTGWs. Separate Prod/Non-Prod VPCs/TGWs may require two distinct ZTGW sets for traffic segregation. VPC Endpoints can be central or in workload VPCs, based on AWS network topology and requirements.

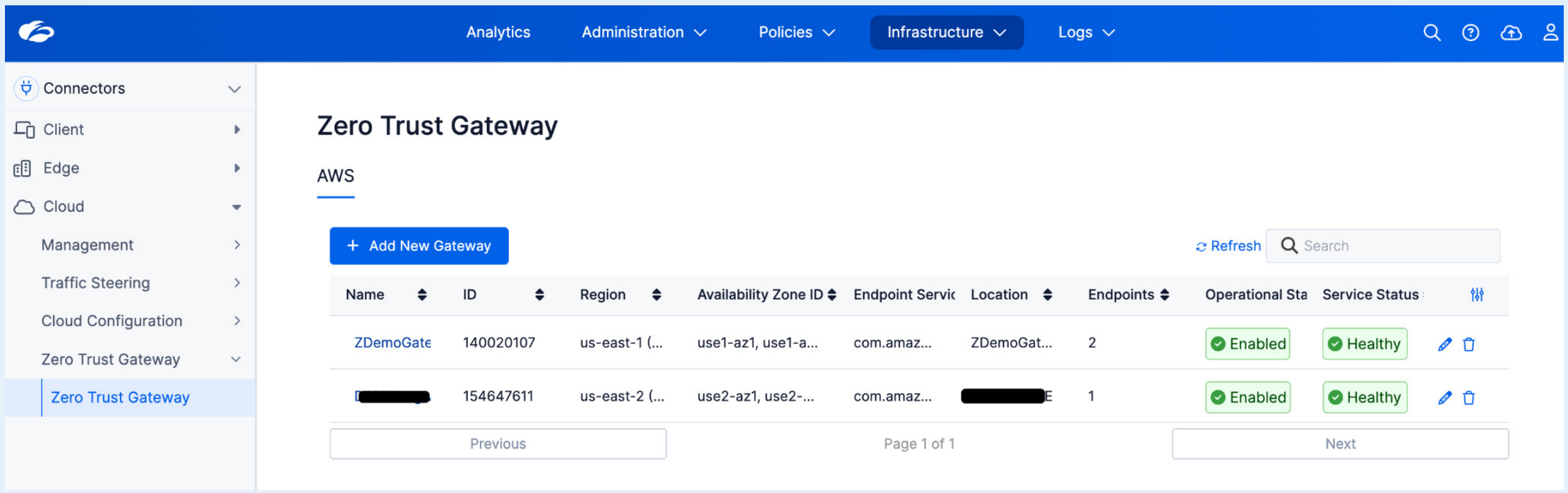
Creating Zero Trust Gateways

A ZTGW operates within a single AWS Region, spanning multiple Availability Zones (AZs). Each ZTGW necessitates a minimum selection of two AZs. Within these selected AZs, managed connectors are deployed and scaled within the Zscaler account, providing a throughput of up to 10Gbps per ZTGW at this time.

Each ZTGW supports connections from AWS Gateway Load Balancer (GWLB) VPC Endpoints within the same AWS Region. Depending on network topology, this could range from two VPC Endpoints in a Security VPC per Region to 50 Isolated VPCs, each with its own VPC Endpoints connected to the Regional ZTGW.

A key aspect of ZTGWs is their function as

EXAMPLE TO SHOW MULTIPLE ZTGW DEPLOYED INTO DIFFERENT REGIONS



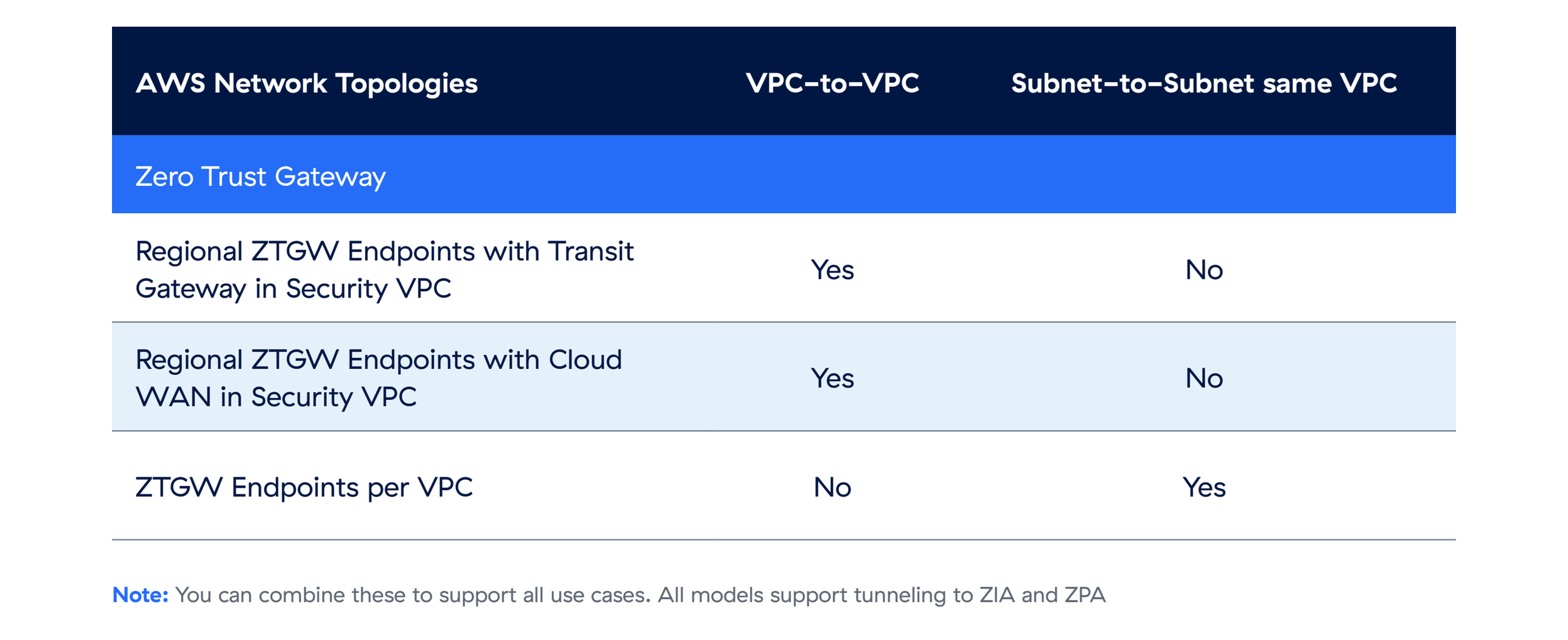
Location objects within the Zscaler ecosystem, familiar to users of the Zscaler Internet Access (ZIA) platform. Similar to other traffic forwarding methods, a ZTGW represents a Location synchronized with both ZIA and ZPA. This enables the use of sub-locations defined by VPC CIDRs, specific attributes via tag discovery, or VPC Endpoint IDs. This capability enhances policy control and reporting flexibility, eliminating the need to deploy a ZTGW for each VPC.

Consider an organization with distinct Prod, Non-Prod, and Test environments in AWS. The separation of traffic is determined by organizational policy. If all these VPCs connect to a single AWS Transit Gateway per Region, a single ZTGW per Region is sufficient. However, if three separate TGWs are used for isolation within a Region, options include connecting all TGWs to the same ZTGW and utilizing sub-locations to differentiate Prod, Non-Prod, and Test environments, or deploying three distinct ZTGWs.

SUPPORTED FLOWS BASED ON AWS NETWORK TOPOLOGY AND WHERE ZTGW ENDPOINTS ARE DEPLOYED

AWS Network Topologies	VPC-to-VPC	Subnet-to-Subnet same VPC
Zero Trust Gateway		
Regional ZTGW Endpoints with Transit Gateway in Security VPC	Yes	No
Regional ZTGW Endpoints with Cloud WAN in Security VPC	Yes	No
ZTGW Endpoints per VPC	No	Yes

Note: You can combine these to support all use cases. All models support tunneling to ZIA and ZPA



Centralized Model with AWS Transit Gateway

Organizations leveraging AWS Transit Gateway can seamlessly integrate ZTGW VPC Endpoints into an existing VPC, such as a Security or Inspection VPC, or establish a new VPC containing the ZTGW VPC Endpoints, subsequently attaching it to the extant Transit Gateway. This configuration enables organizations to either direct the Default Route (0.0.0.0/0) to the ZTGW VPC Endpoints or selectively route traffic for a diverse array of alternative applications.

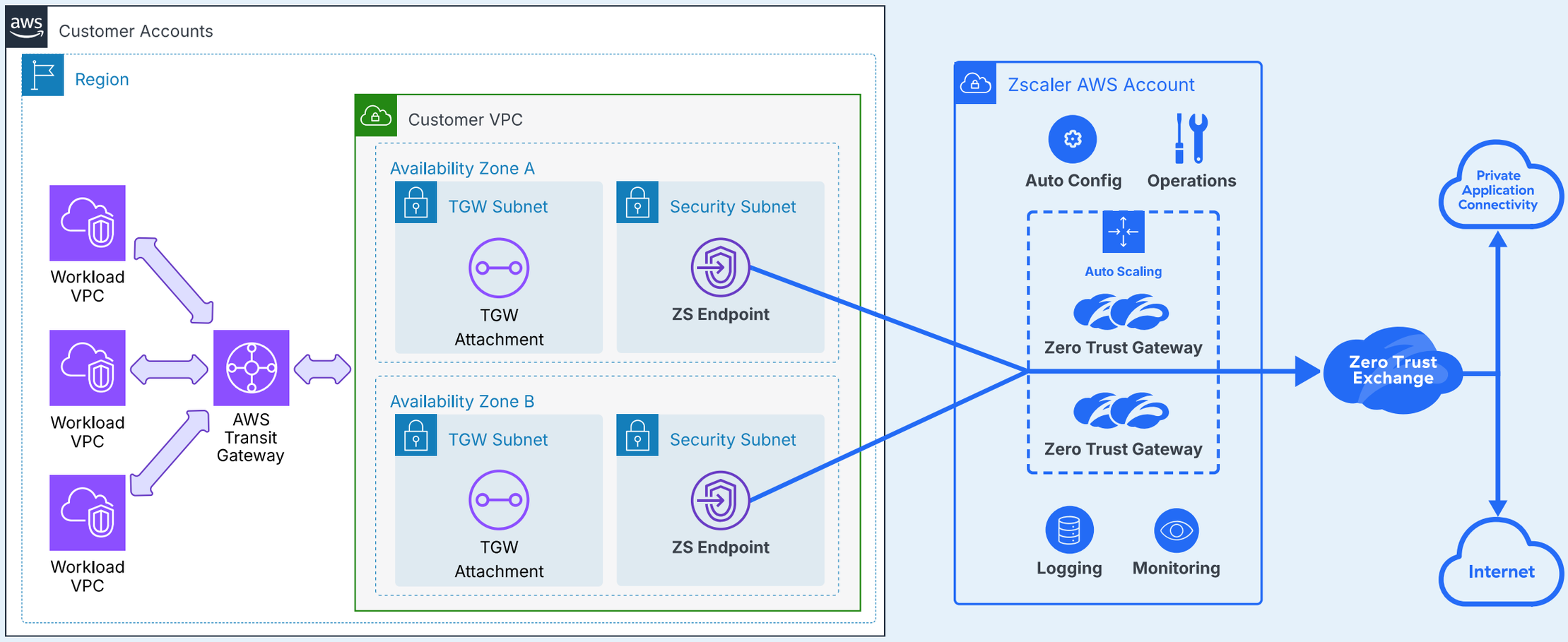
This model enables the connection of multiple Transit Gateways to the same ZTGW within the same Region, or alternatively, the establishment of distinct ZTGW services to segregate various Transit Gateways (e.g., Production, Non-Production, Development).

When ZTGW VPC Endpoints are deployed in a centralized VPC, rather than directly within each workload VPC, the following use cases are supported:

Internet egress security via ZIA

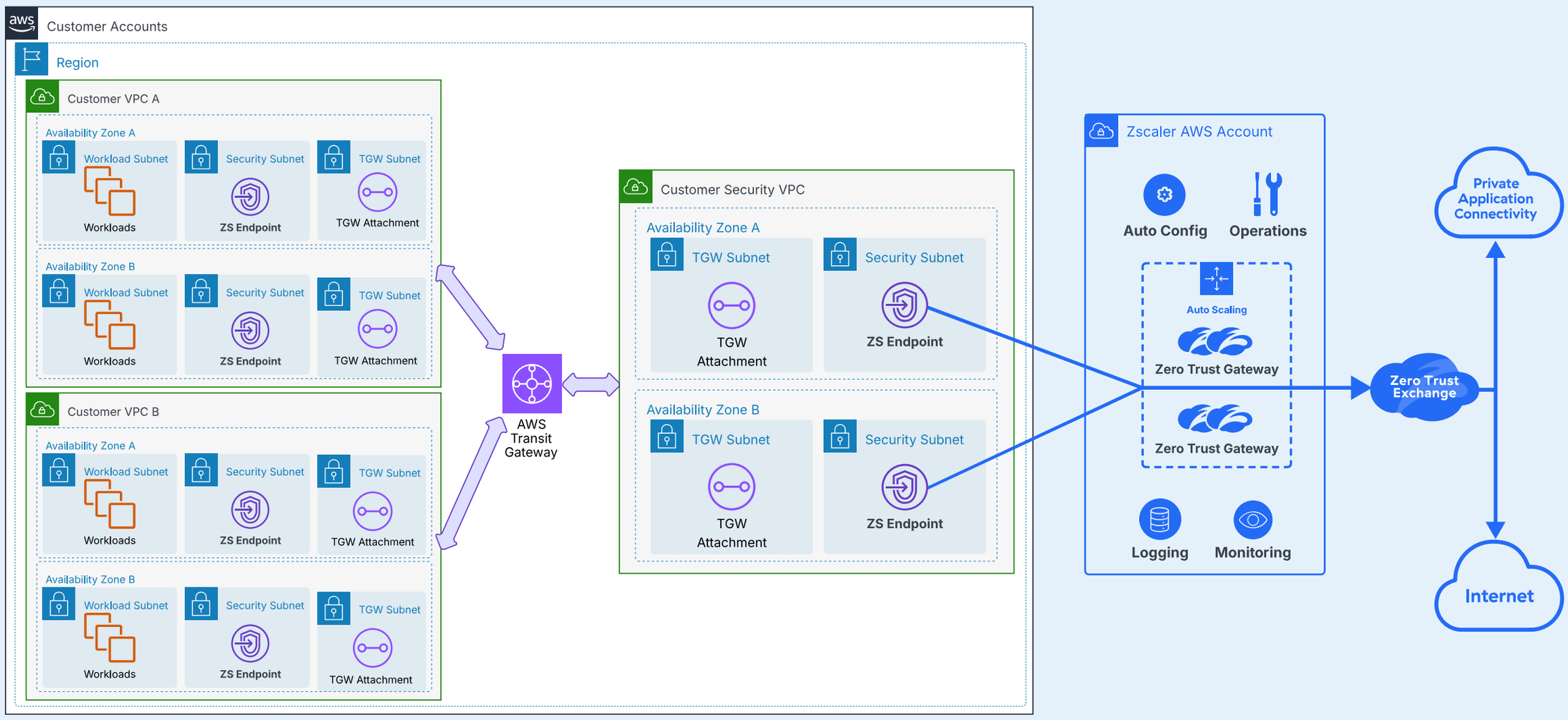
Private application access via ZPA

East-West firewall rules between VPCs within the same TGW



Hybrid Model with TGW Centralized Endpoints & Workload VPC Endpoints

While this particular topology might not be a frequent customer request, its inclusion serves a crucial illustrative purpose. It demonstrates that effectively addressing all the outlined East–West use cases within this specific network architecture mandates the strategic deployment of ZTGW VPC Endpoints. This deployment is required to provide comprehensive visibility of control for not just VPC to VPC traffic, but also traffic within & between subnets in each VPC.



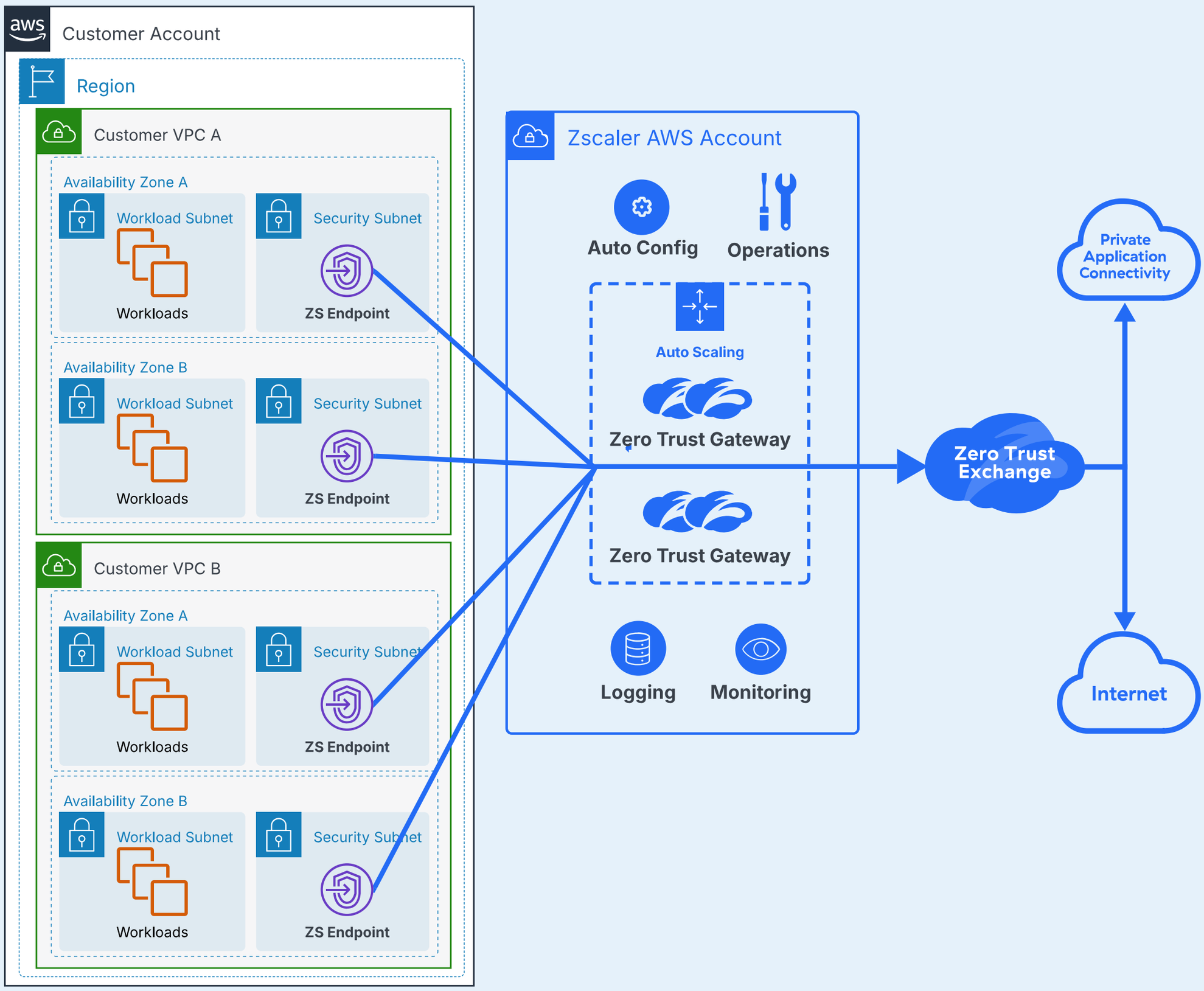
Decentralized Model with Isolated Workload VPCs

When ZTGW VPC Endpoints are deployed into a workload VPCs, and not in a centralized security VPC connected to Transit Gateway, the following use cases are supported:

Internet egress security via ZIA

Private application access via ZPA

East–West firewall rules between VPCs within the same TGW



ZTGW Service Architecture

ZTGW service is broken into 2 primary services:

Control Plane

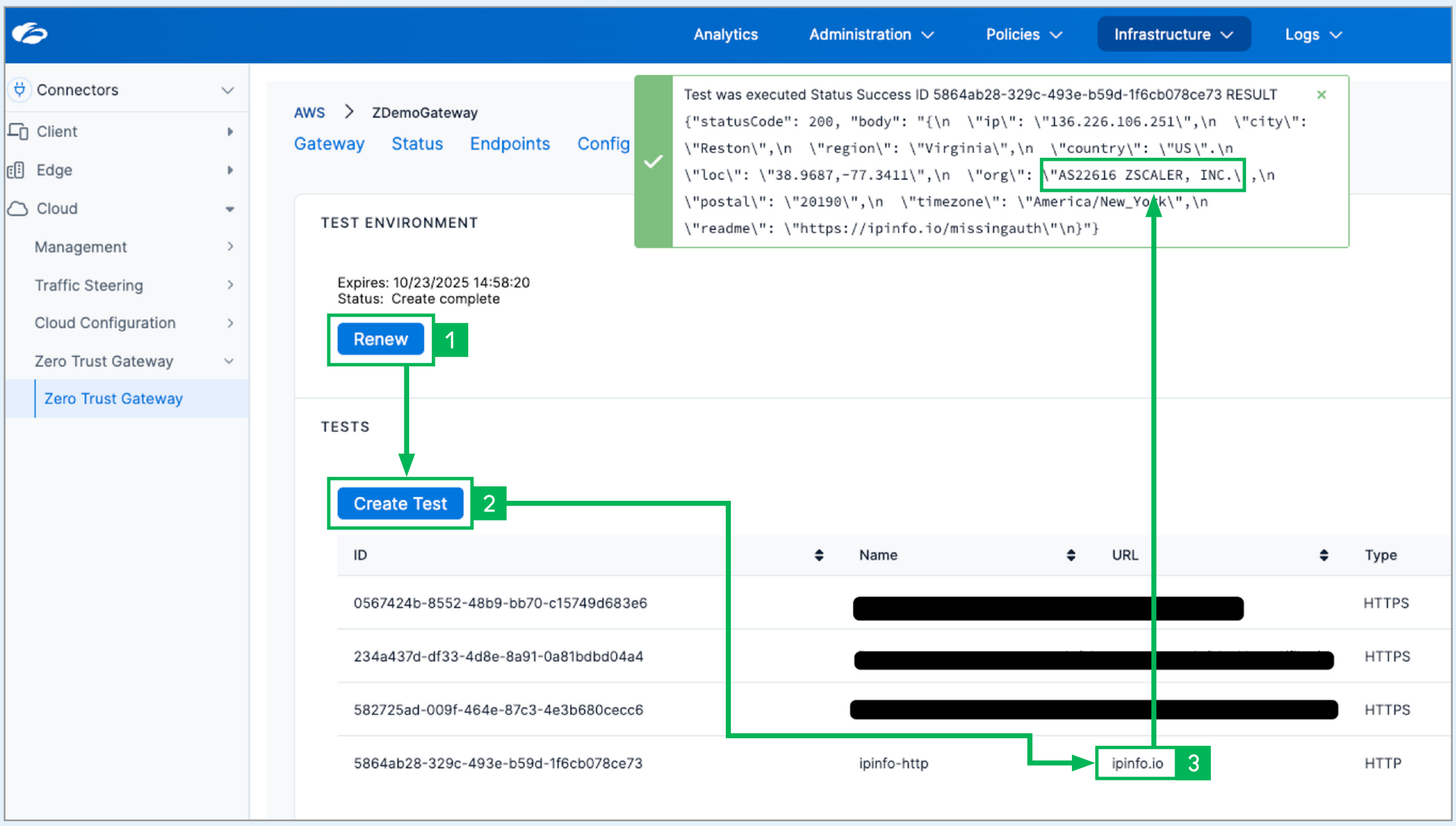
This is for configuration & orchestration of the services

Data Plane

This is where customer workload traffic flows through

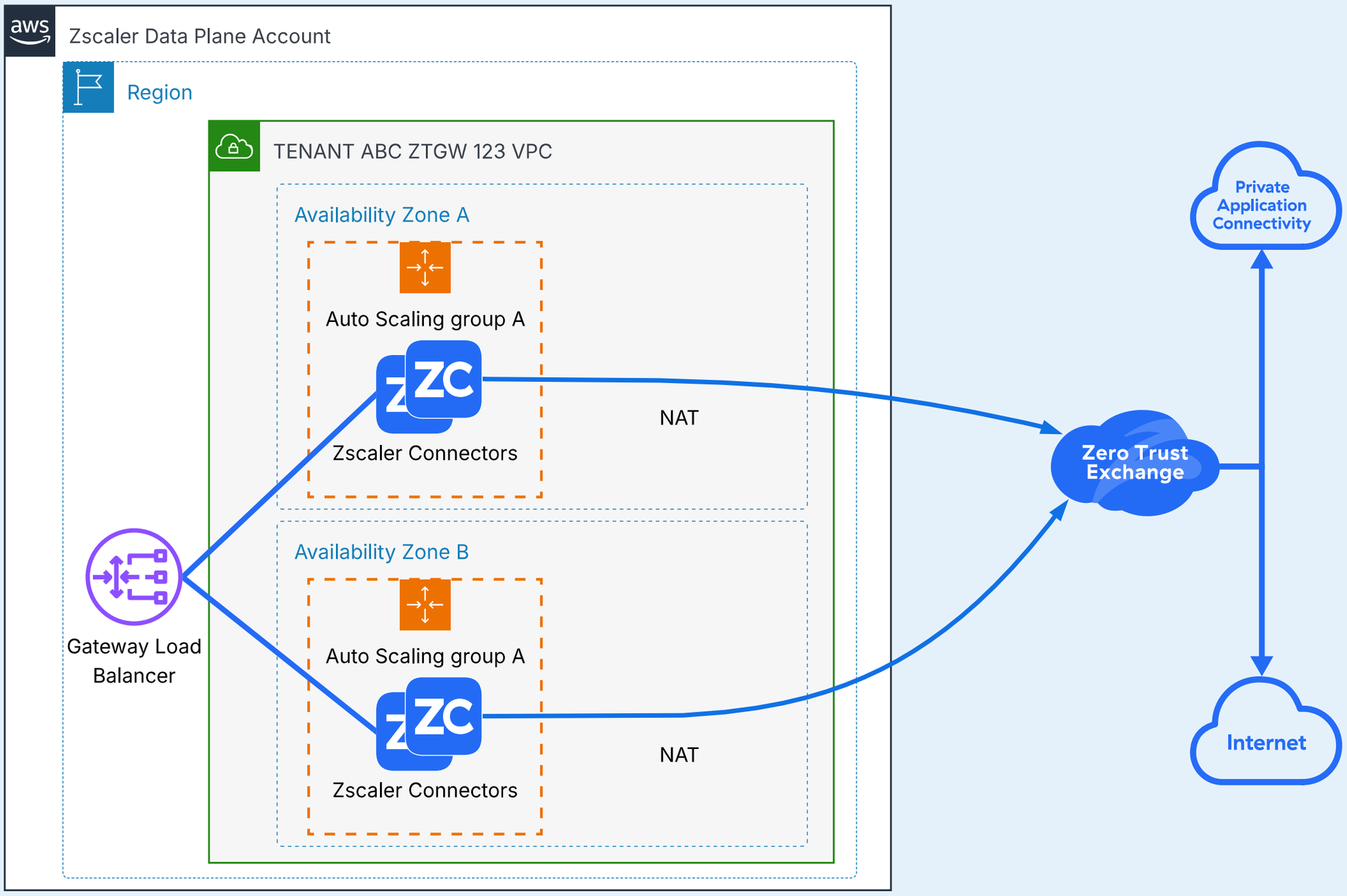
The Control Plane is a multi-tenant service built using modern AWS native services in dedicated Zscaler owned AWS accounts. Various operations such as logging, metrics, maintenance, initiating new ZTGW deployments, software updates, and rule syncs occur via the control plane. The control plane includes historical configurations, metrics, and analytics. No customer traffic flows through the control plane, but a traffic test can optionally be performed which is managed through the control plane.

EXAMPLE TO SHOW MULTIPLE ZTGW DEPLOYED INTO DIFFERENT REGIONS



The Data Plane employs a single-tenant architecture with VPC isolation. All data plane components and services are dedicated, ensuring traffic isolation and preventing sharing with other Zscaler customers. Each ZTGW service currently has dedicated AWS components, such as:

- VPC
- NAT services with Public IPs per AZ
- Auto Scaling Groups per AZ
- Gateway Load Balancer Service
- ZTGW Connectors Compute Engine



Zscaler is constantly improving and innovating to make the services resilient and cost-effective. Today the ZTGW Connectors can securely tunnel to Zscaler Internet Access (ZIA) or Zscaler Private Access (ZPA), send traffic Direct out of the Connectors, or enforce East-West local traffic policies.

ZTGW Secure Design

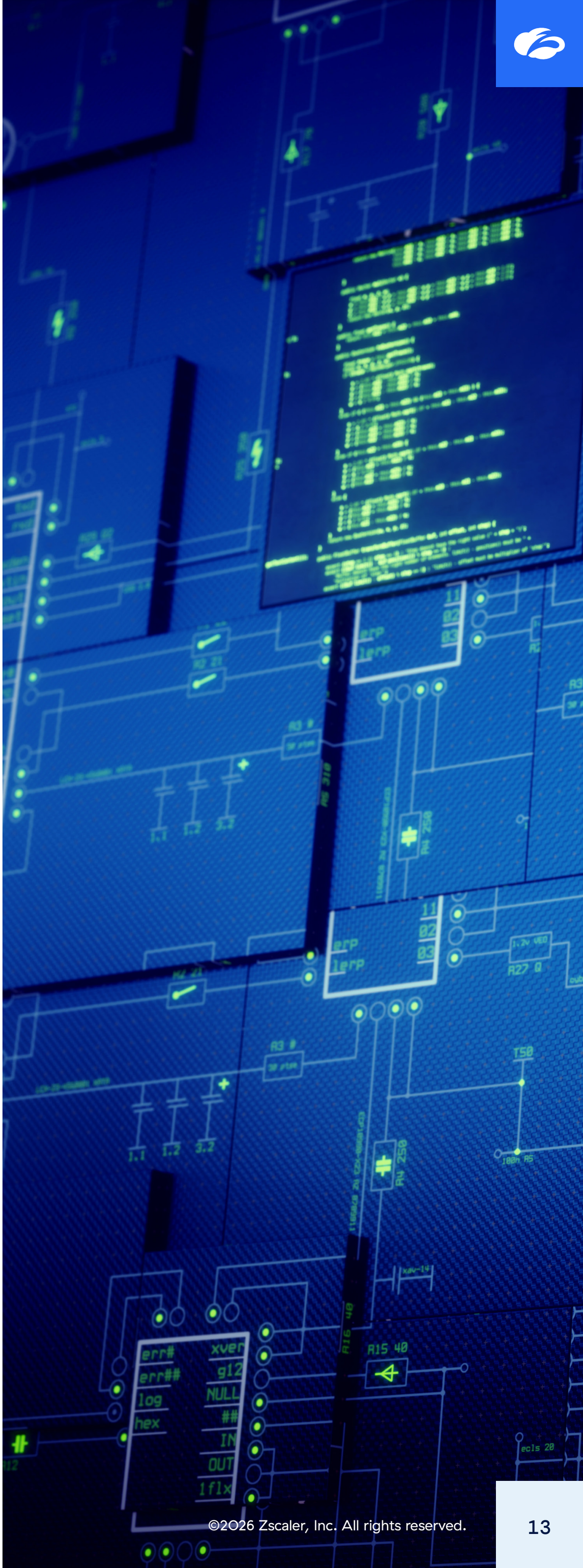
The ZTGW service control and data plane are engineered with multiple security layers and best practices to ensure the integrity of customer data and overall service access are restricted to authorized requests. Customers requiring more in-depth details on the service's design are encouraged to contact their Zscaler account team for a comprehensive discussion.

- Gateway Load Balancer (GWLB) Service that powers the ZTGW will only accept VPC Endpoint connection requests from AWS accounts allowed by the customer within the Zscaler portal — defined by either AWS Account IDs or by Accounts integrated with the Tag Discovery service. No other AWS account can discover or connect to your ZTGW services.

- Zscaler's security team rigorously reviews services pre-deployment and continuously to ensure adherence to best practices, including least privilege and robust IAM policies. Cloud Security Posture Management (CSPM) tools are utilized to proactively detect and remediate any misconfigurations or security drifts.

- Each ZTGW is single tenant, meaning there is no mixing of workload traffic going to the same ZTGW Connectors from different tenants or customers. This ensures complete network isolation within the ZTGW Service.

- ZTGW Control Plane and Data Plane services are in separate AWS accounts.



Advantages

Implementing the Zscaler ZTGW for AWS enables organizations to prioritize policy, visibility, and security policy, rather than being hindered by operationalization challenges. The following are the top five benefits:

Ease of Deployment

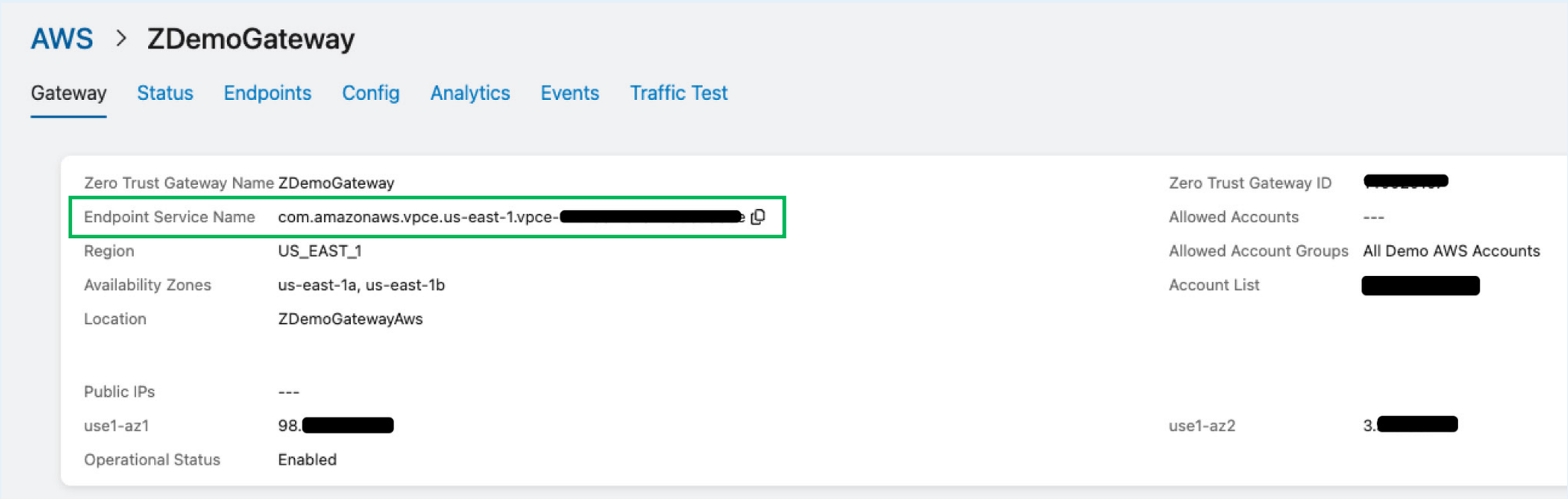
Deploying a ZTGW takes under five minutes. After providing the location name, region, availability zones, and trusted AWS accounts, a new ZTGW service is deployed in the Zscaler AWS account. The Zscaler console will then display the Service Name.

Whether the service name is provided to business units for connecting VPC Endpoints from their VPCs or automated via AWS CloudFormation or Terraform, the process transitions from deployment to connectivity operations. AWS route tables are updated to direct traffic, such as the Default Route (0.0.0.0/0), to the ZTGW VPC Endpoints. Security policies are then managed through Zscaler.

Reduced Operational Burden

Unlike EC2 instance–based solutions, which involve a shared responsibility model, the ZTGW service is fully managed by Zscaler. Customers are solely responsible for ensuring VPC Endpoints connected to the ZTGW Service are within authorized AWS Accounts and that routing is properly configured. Operating across multiple Availability Zones with Auto Scaling in Zscaler's AWS account, the ZTGW service functions as a native AWS offering. This eliminates the need for key rotation, SSH access for troubleshooting, or updating configurations. While Zscaler manages updates for its EC2 instances like Cloud Connectors, ZTGW requires no customer intervention, even for major changes. Zscaler's extensive operational insights, including logging and metrics, further enhances the benefits provided to customers through this solution. Zscaler manages and operates the world's largest security cloud, now extended to the ZTGW connectivity solution.

EXAMPLE ZTGW DETAILS WITH ENDPOINT SERVICE NAME TO CONNECT TO FROM AWS



Automatic Scaling

The Zscaler ZTGW addresses throughput limitations by leveraging AWS Auto Scaling. This design allows the ZTGW to support sustained and burst traffic demands while optimizing costs. Currently, each regional ZTGW supports up to 10Gbps, with plans to increase throughput over time. Aggregate throughput requirements exceeding 10Gbps are accommodated by distributing ZTGW instances across multiple regions and environments.

Cost Reduction & Consolidation

A cloud-native solution utilizing AWS services optimizes costs with ZTGW. Due to variable factors, a universal cost calculation is not feasible. However, increased traffic (GB/month) correlates with greater cost savings, which can be categorized into three primary areas.

COST SAVINGS VARY; TABLE SHOWS MINIMUM SAVINGS, EXCLUDING EXISTING SECURITY SERVICES

ZTGW Cost Reduction		
Area	**Cost**	**Services**
EC2 Compute	AWS cost	EC2 Instance Hourly, EC2 Amazon Elastic Block Storage (EBS)
VPC Networking	AWS cost	Data Transfer Out (DTO) to Internet, NAT (Network Address Translation) Gateway, Internet Gateway, Public IP addresses, GWLB service GWLB cross-zone data transfer,
Operational	AWS & Soft cost	Amazon CloudWatch Time and expertise to deploy and manage the solution is significantly reduced

This document will now detail how ZTGW contributes to the reduction of these costs:

COMPUTE

The ZTGW service operates natively within AWS, eliminating the need to deploy and manage EC2 instances for Zscaler connectivity. Previously, securing AWS VPC network traffic often required numerous EC2 instances, depending on factors like region count, VPC count, and throughput. This could incur a minimum cost of $100/ month for a small, highly available pair securing a single VPC or a limited number of workloads in a security VPC utilizing Transit Gateway.

NETWORKING

Zscaler ZTGW for AWS enhances security and offers significant network cost savings compared to traditional EC2-based firewall solutions. Traditional methods often incur substantial Data Transfer Out (DTO) and NAT Gateway costs, whether deployed with public IP addresses or behind NAT Gateways.

Three primary networking cost drivers are:

- **DTO Costs:** Based on monthly per-GB processed volume, these tiered costs can range from hundreds to tens of thousands of dollars for high data egress.
- **NAT Gateway Runtime Costs:** Even without traffic, running NAT Gateways can cost hundreds of dollars monthly.
- **NAT Gateway Per-GB Processed Volume:** Similar to DTO, this volume-based cost can add hundreds of dollars monthly for processing 10 TB or more.

ZTGW mitigates these costs by hosting its services within Zscaler's AWS account. The only customer-incurred networking costs are the hourly and data processing charges for GWLB VPC Endpoints connecting to the ZTGW service. Critically, DTO and NAT Gateway costs are entirely eliminated for customers because traffic exits AWS from Zscaler, not the customer's AWS account, via AWS PrivateLink.

While a ZTGW subscription is required as part of the Zero Trust Cloud platform (based on the number of ZTGWs needed and monthly data volume), the Total Cost of Ownership (TCO) is reduced due to the service's architecture within AWS, allowing for cost savings to be passed to customers.

For a detailed understanding of potential cost savings in your AWS environment, contact your Zscaler account team. You can also baseline current costs using the AWS billing & usage console and the AWS pricing calculator: https://calculator.aws.

OPERATIONAL

ZTGW eliminates the need to deploy, manage, and troubleshoot EC2 instance-based security appliances within AWS, simplifying daily operations and planning. This offers several key advantages:

- **Simplified Sizing:** Zscaler manages ZTGW VPC Endpoint sizing, eliminating the need for internal sizing exercises for regional egress points.

- **Reduced Configuration Complexity:** ZTG simplifies deployment by requiring only the connection of ZTGW VPC Endpoints and route table updates, removing the need for complex Terraform modifications for security appliances, load balancing, and auto-scaling.

- **Unified Visibility and Control:** All traffic routed through Zscaler provides end-to-end visibility and control, consolidating security management and eliminating the need for multiple vendors or tools.

In summary, the Zscaler ZTGW offers significant benefits in terms of simplicity and cost savings, by enabling the consumption of security as a service. This allows your organization to concentrate on refining security policies. For further exploration of how the ZTGW can benefit your organization, please refer to the resources in the next section for a comprehensive next step.

Resources

To learn more about the service make sure to read this blog:
https://www.zscaler.com/blogs/product-insights/zscaler-zero-trust-cloud-zero-trust-gateway

Further information, including videos, reading materials, and an interactive click-through demonstration, can be found at the Zero Trust Gateway Hub:
https://app.storylane.io/hub/dlvieowha6az

For details on Zero Trust Cloud Deployment Options, please refer to:
https://www.zscaler.com/resources/solution-briefs/deployment-models-for-zero-trust-cloud.pdf

Additional ZTGW configuration information is available in the Zscaler Help Portal documentation:
https://help.zscaler.com/cloud-branch-connector/what-zero-trust-gateways

+1 408.533.0288 **Zscaler, Inc. (HQ) • 120 Holger Way • San Jose, CA 95134** **zscaler.com**

zscaler™

Act Fast.
Stay Secure.