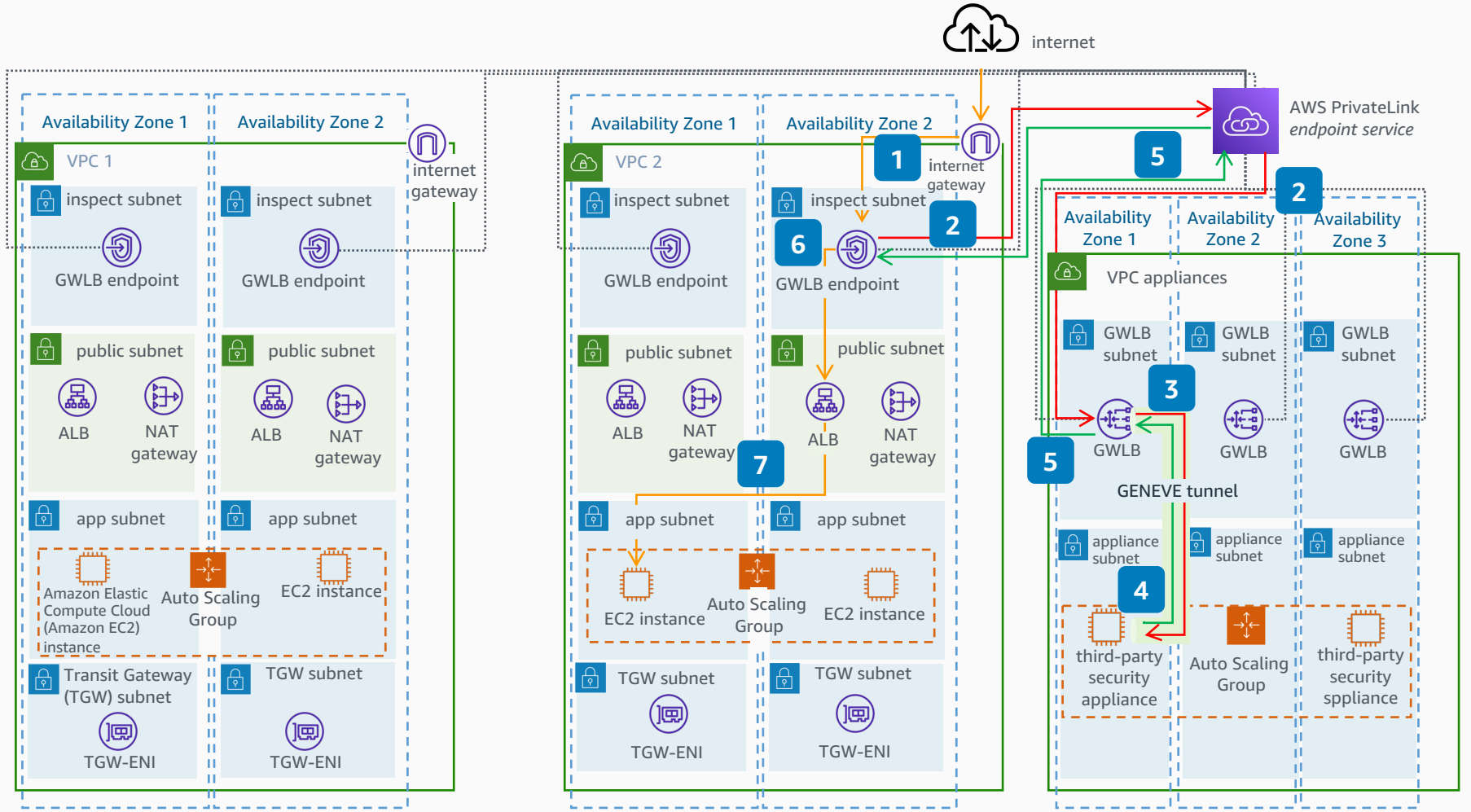
Distributed Inspection Architectures with Gateway Load Balancer

Reviewed for technical accuracy July 21, 2022

North-South inbound distributed inspection with AWS Gateway Load Balancer

Use AWS Gateway Load Balancer to inspect your inbound traffic in a distributed fashion using the same backend security appliances for several virtual private clouds (VPCs).



1. Traffic coming from the internet destined for the **Application Load Balancer (ALB)** arrives at the internet gateway and is forwarded to a **Gateway Load Balancer (GWLB)** endpoint using the ingress route table.
2. The **GWLB** endpoint forwards the traffic to the **GWLB** in the appliances VPC using **AWS PrivateLink**.
3. The **GWLB** encapsulates the traffic in Generic Network Virtualization Encapsulation (GENEVE). GENEVE encapsulated traffic is sent for inspection to a security appliance.
4. Once the traffic is inspected, it is sent back to the **GWLB**.
5. This traffic is then returned to the **GWLB** endpoint in the Inspect subnet.
6. The **GWLB** endpoint uses the inspect subnet route table to forward the traffic to the **ALB** in the public subnet.
7. Lastly, the **ALB** forwards the traffic to one of its healthy instances.

*We recommend you follow these best practices when deploying a **Gateway Load Balancer**.*

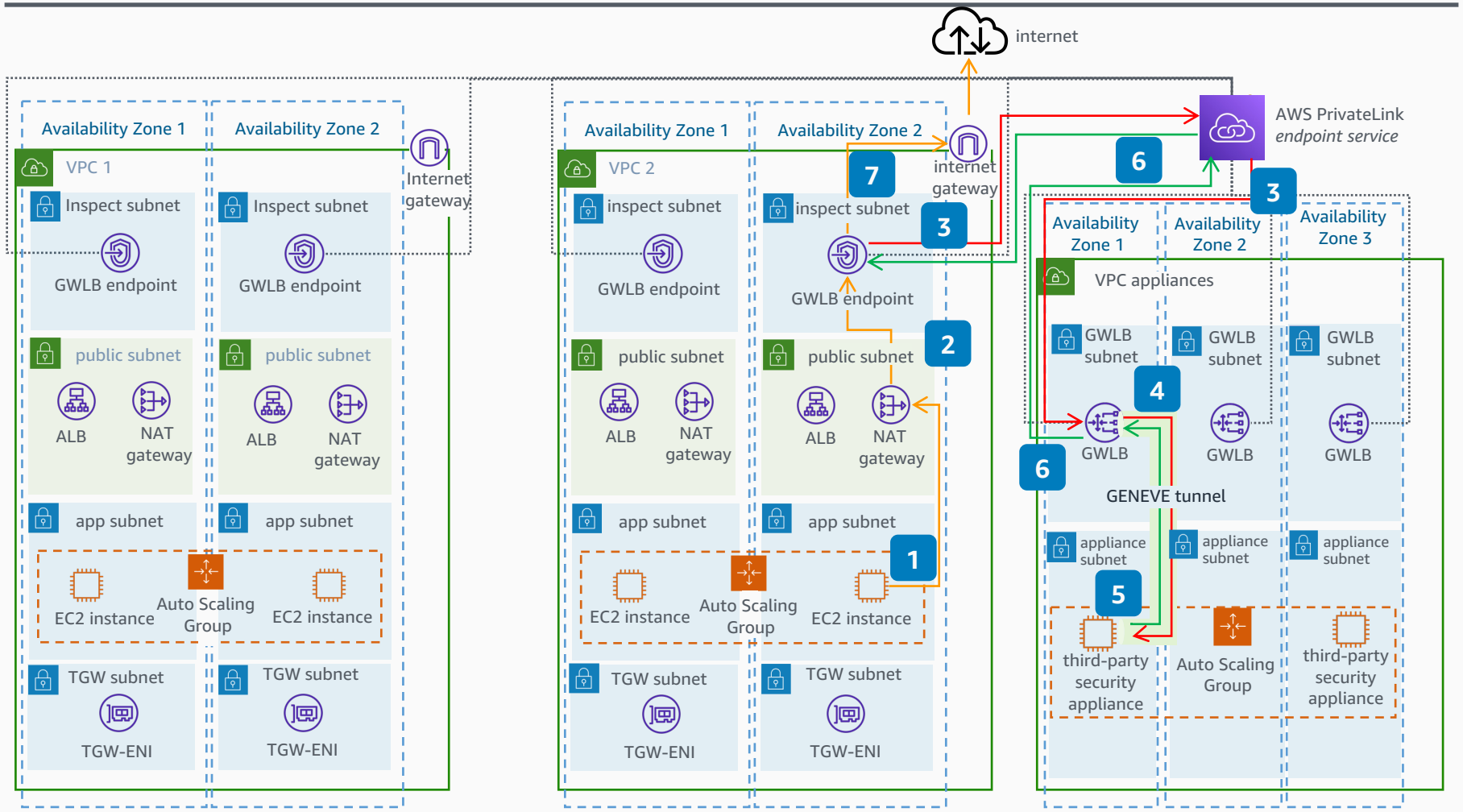
For more information about how to implement a distributed inspection architecture refer to: Scaling network traffic inspection using AWS Gateway Load Balancer.

Reviewed for technical accuracy July 25, 2022

AWS Reference Architecture

North-South outbound distributed inspection with AWS Gateway Load Balancer

Use AWS Gateway Load Balancer to inspect your outbound traffic in a distributed fashion using the same backend security appliances for several VPCs.



1. Traffic coming from an instance destined to the internet arrives at the **NAT gateway**, which translates the source IP of the packets.
2. The **NAT gateway** forwards the translated packets to a **GWLB** endpoint using the public subnet route table.
3. The **GWLB** endpoint forwards the traffic to the **GWLB** in the appliances VPC using **AWS PrivateLink**.
4. The **GWLB** encapsulates the traffic in GENEVE. GENEVE encapsulated traffic is sent for inspection to a security appliance.
5. Once the traffic is inspected, it is sent back to the **GWLB**.
6. This traffic is then returned to the **GWLB** endpoint in the inspect subnet.
7. The **GWLB** endpoint uses the inspect subnet route table to forward the traffic to the internet gateway.

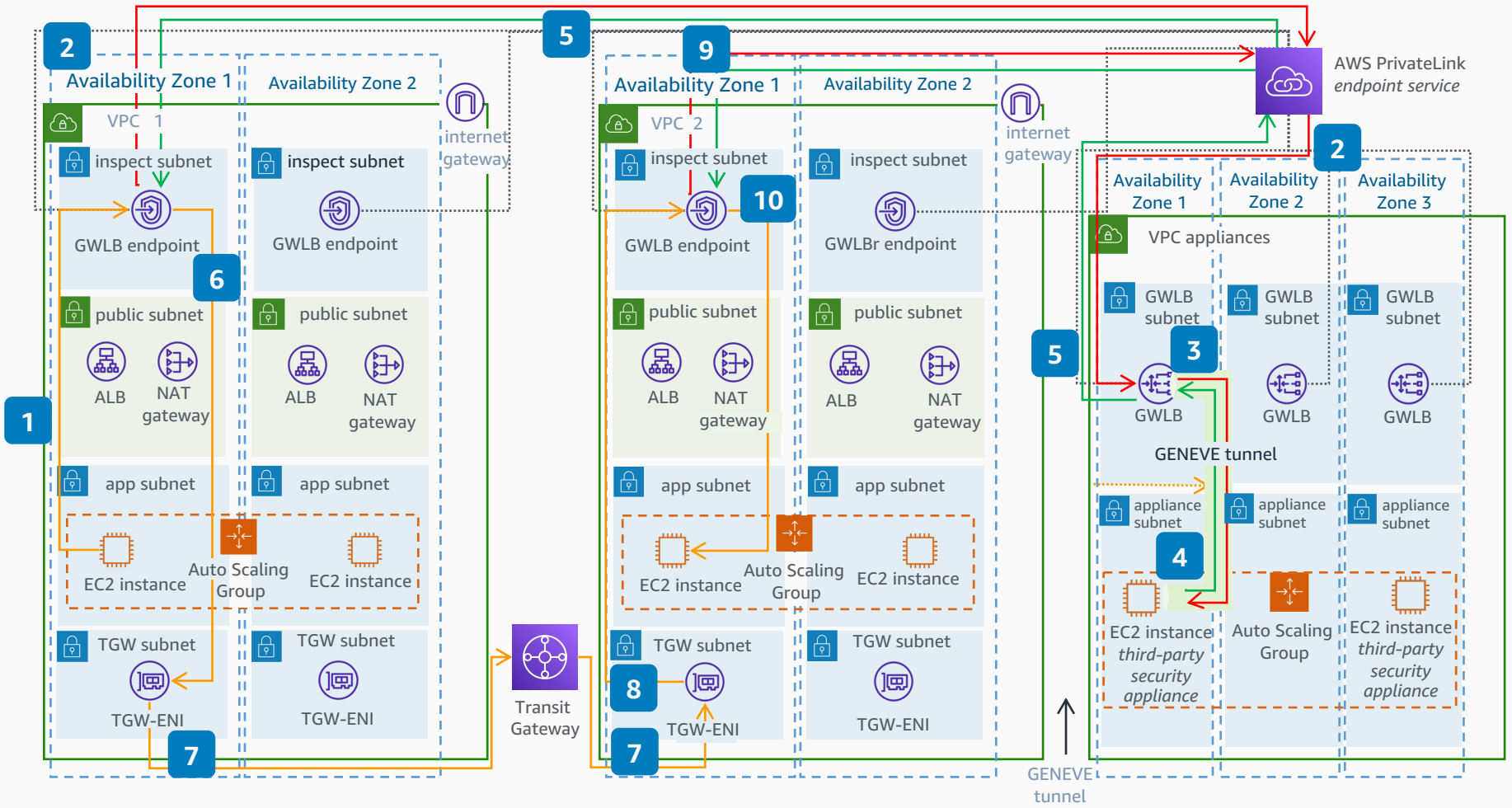
*We recommend you follow these best practices when deploying a **Gateway Load Balancer**.*

For more information about how to implement a distributed inspection architecture refer to: Scaling network traffic inspection using AWS Gateway Load Balancer.

AWS Reference Architecture

East-West distributed inspection with AWS Gateway Load Balancer

Distributed East-West inspection architecture with AWS Gateway Load Balancer.



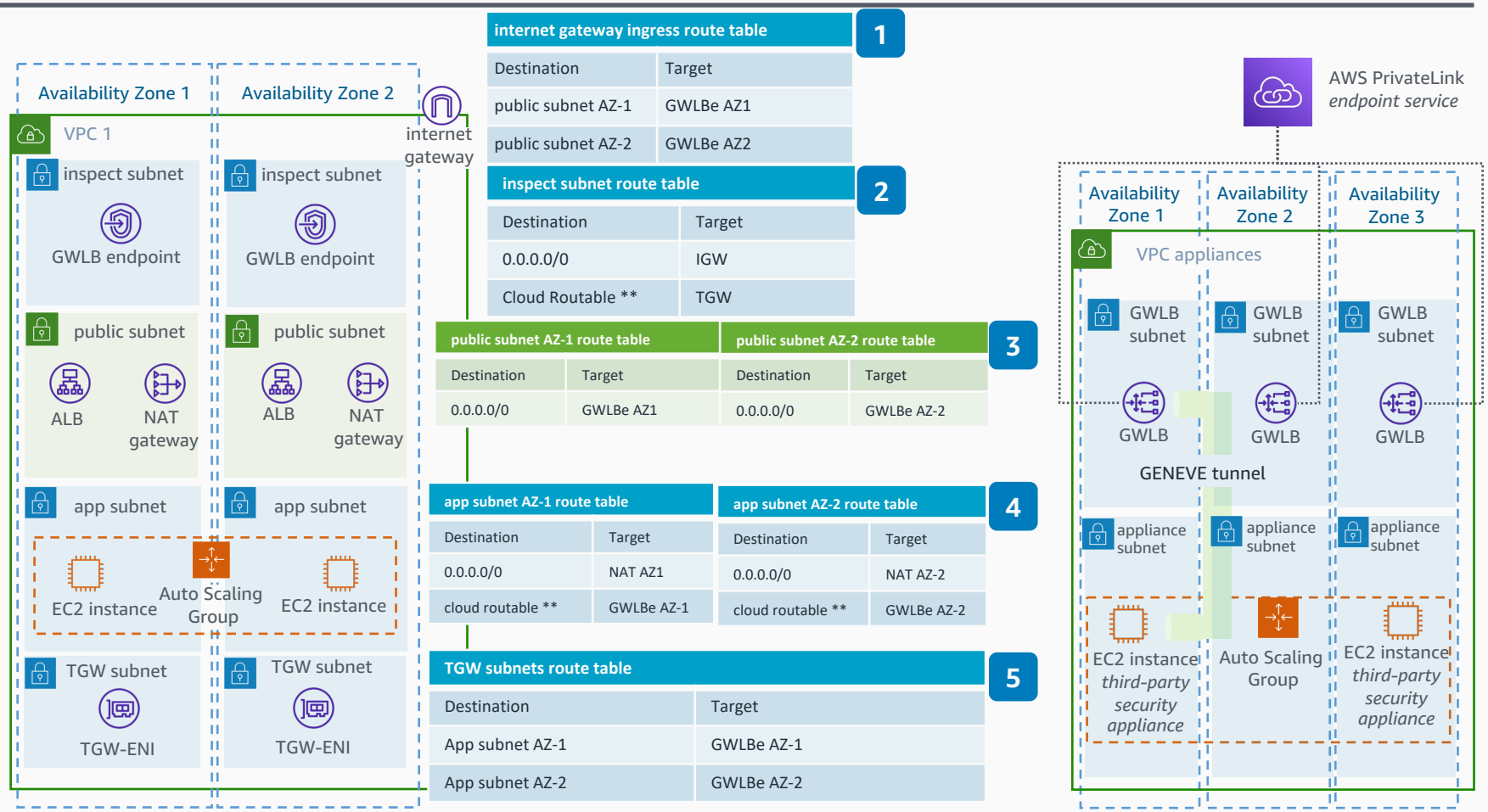
1. Traffic coming from an instance in VPC 1 destined for an instance in VPC 2 is forwarded to a **GWLB** endpoint.
2. The **GWLB** endpoint forwards the traffic to the **GWLB** in the appliances VPC using **AWS PrivateLink**.
3. The **GWLB** encapsulates the traffic in GENEVE. GENEVE encapsulated traffic is sent for inspection to a security appliance.
4. Once the traffic is inspected, it is sent back to the **GWLB**.
5. This traffic is then returned to the **GWLB** endpoint in the Inspect subnet.
6. The **GWLB** endpoint uses the inspect subnet route table to forward the traffic to the **TGW** endpoint in the TGW subnet.
7. The traffic is forwarded in accordance to the **TGW** route table associated, and arrives in VPC 2.
8. In VPC 2, traffic is forwarded to the **GWLB** endpoint.
9. The traffic is re-inspected, following the same flow as previously [2,3,4,5].
10. The traffic is forwarded by the **GWLB** endpoint to the instance in the app subnet.

We recommend you use [Transit Gateway appliance mode](#) in the Inspection VPC Transit Gateway attachment to maintain flow symmetry.

AWS Reference Architecture

Distributed inspection with AWS Gateway Load Balancer route tables

Distributed inspection architecture route tables using AWS Gateway Load Balancer in multiple Availability Zones.



* There's always a local route for VPC CIDRs in every VPC route table.
** Cloud Routable CIDRs are private CIDRs that are reachable without using the internet.

1. **Internet gateway ingress route table** is applied to traffic coming from the internet to the public subnets. It forwards traffic to the **GWLB** endpoint in the destination Availability Zone to keep flow symmetry.

2. **Inspect subnets route table** is applied to traffic already inspected. This route table defines what traffic will be sent to the internet and to the **TGW.**

3. **Public subnets route tables** are applied to public subnets. All the traffic is forwarded to the **GWLB** endpoints in the source Availability Zone.

4. **Application subnets route tables** forward traffic differently, depending on whether the destination IP is public or private, to preserve the source IP of internal traffic. All the traffic is forwarded to the **GWLB** endpoints in the same Availability Zone to keep symmetry.

5. **Transit gateway subnets route table** is applied to traffic coming from the **Transit Gateway**. To keep symmetry, these route tables send traffic to the **GWLB** endpoint in the destination Availability Zone.

For more information about how to implement a distributed inspection architecture refer to: [Scaling network traffic inspection using AWS Gateway Load Balancer](#).

Reviewed for technical accuracy July 25, 2022

AWS Reference Architecture