

Networking in AWS

This presentation tries to explain AWS networking concepts and differences between traditional Datacentre and Public Cloud networks.

Definitions

- Instance Virtual machine (VM)
- Availability Zone (AZ) set of buildings, Internet uplinks, and power. You can think of it as a data centre, but some Availability Zones contain more than one physical data centre.
- **Region** collection of Availability Zones. You can think of a Region as containing multiple data centres within the same geographic area (US East, Australia, EU West, etc.).

Global Infrastructure



Under the hood - AWS SDN Mapping Service



IP: 10.1.1.11

IP: 10.1.1.12

MAC: 3c:15:c2:d2:96

HYPERVISOR: HOST01 VPC: vpc-abdef IP:10.1.1.10 MAC: 3c:15:c2:d2:aa:12 HYPERVISOR: HOST02 VPC: vpc-abdef IP:10.1.1.11 MAC: 3c:15:c2:d2:93:dc HYPERVISOR: HOST03 VPC: vpc-abdef IP:10.1.1.12 MAC: 3c:15:c2:d2:96:fe HYPERVISOR: HOST04 VPC: vpc-abdef IP:10.1.1.13 MAC: 3c:15:c2:d2:11:a1

arp who-has 10.1.1.13 tell 10.1.1.10

arp reply 10.1.1.13 is-at 3c:15:c2:d2:11:a1

- Similar concepts as in VMWare NSX
- Details are not revealed by AWS
- The mapping service registers every EC2 instance started by a customer
 - When an EC2 instance sends an ARP request to reach any other instance, instead of allowing this broadcast on the network, the hypervisor catches the request, sends it to the mapping service data base, which will reply with the requested MAC address but also with the IP of the hypervisor IP of the host on which the target instance is running.

No OSI Layer 1 and 2

- AWS does not support Layer 2 protocol forwarding
- All frames are intercepted by the hypervisor
- Consequences:

No broadcast domains

No ARP and GARP

No HSRP, VRRP

No CDP, LDP

No LAG (LACP)

Multicast is not supported
 No routing protocols relying on multicast neighbours discovery – OSPF, EIGRP, RIPv2
 Only BGP is supported

Build overlay networks if you need the features above.

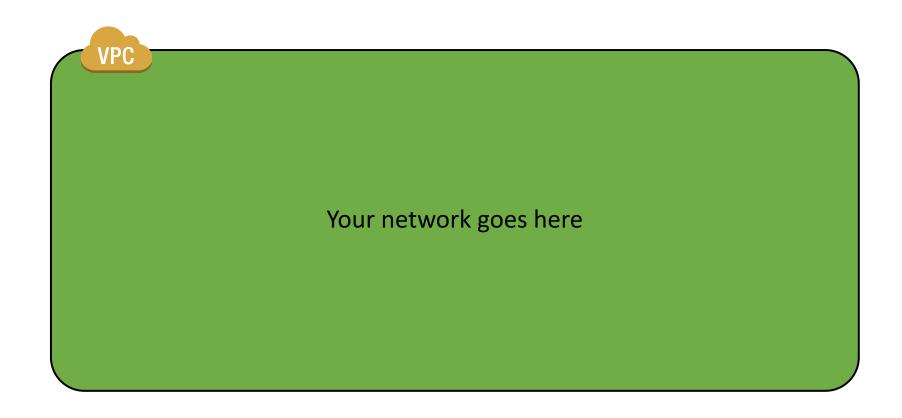
Networking building blocks:

VPCs
Subnets
Routing tables
Internet and NAT Gateways
Elastic Network Interfaces
VPC Endpoints

Networking Building Blocks

Amazon Virtual Private Cloud (VPC)

Essential building block - similar to virtual routing and forwarding (VRF)



VPCs > Create VPC

Create VPC

A VPC is an isolated portion of the AWS cloud populated by AWS objects, such as Amazon EC2 instances. You VPC.

Name tag	test-vpc11	•
IPv4 CIDR block*	10.2.0.0/16	6
IPv6 CIDR block	 No IPv6 CIDR Block Amazon provided IPv6 CIDR block 	
Tenancy	Default	• 0

* Required

CIDR size from /16 to /28 which then split to subnets

No Layer 2, no broadcast domains, per-interface security rules, so it's OK (and even recommended) to allocated large CIDRs for VPCs and server subnets such as /18.

IPv4 subnets can be any, even non-RFC1918 (but they will stay private to your VPC and why? (20)

Note that IPv6 address ranges are allocated by Amazon as /56 blocks from their own ranges.

Edit CIDRs

Add or remove CIDR blocks for your VPC. Learn more.

VPC ID vpc-04237dc5b3644b54a

VPC IPv6 CIDRs



VPC IPv4 CIDRs



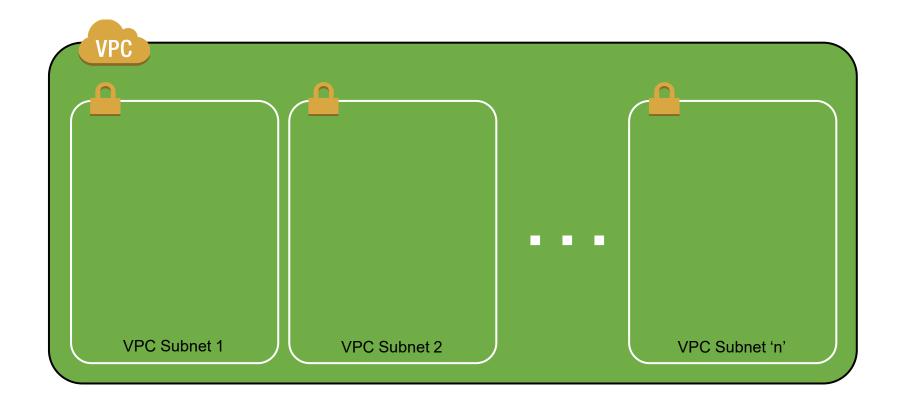
It is possible to add new CIDR ranges, but the address range defined at VPC creation can't be deleted.

Address planning is important as VPCs can have overlapping addresses, but these VPCs can't be peered.

If IPv6 addresses allocated, they are assigned to instances in addition to IPv4, so the number of available IPv4 addresses will limit the number of IPv6 hosts

Networking Building Blocks Subnets

Next step - Create your subnets

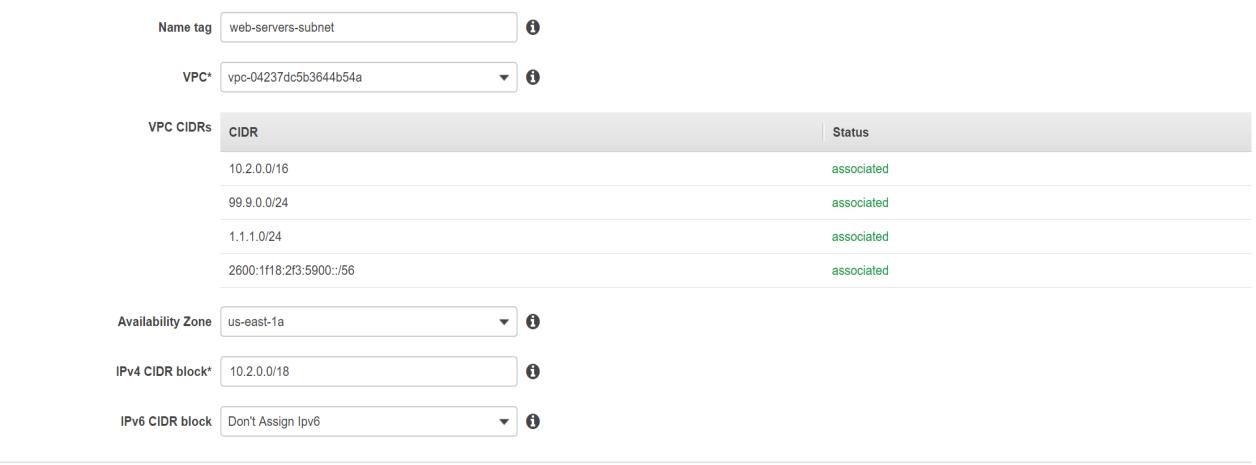


Subnets

- Subnets take a variable portion of the CIDR block assigned to the VPC.
- Subnet can't span multiple AZ
- AWS reserves five addresses in each subnet—the first four and the last address, for example for 10.1.1.0/24:
- > 10.1.1.0: Network address.
- > 10.1.1.1: VPC router (implicit router)
- > 10.1.1.2: DNS server
- ➤ 10.1.1.3: Reserved by AWS for future use.
- ➤ 10.1.1.255: Network broadcast address. Even though AWS does not support broadcast in a VPC, this address is reserved.

Create subnet

Specify your subnet's IP address block in CIDR format; for example, 10.0.0.0/24. IPv4 block sizes must be between a /16 netmask and /28 netmask, and can be the same size as your VPC. An IPv6 CIDR block must be a /64 CIDR block.



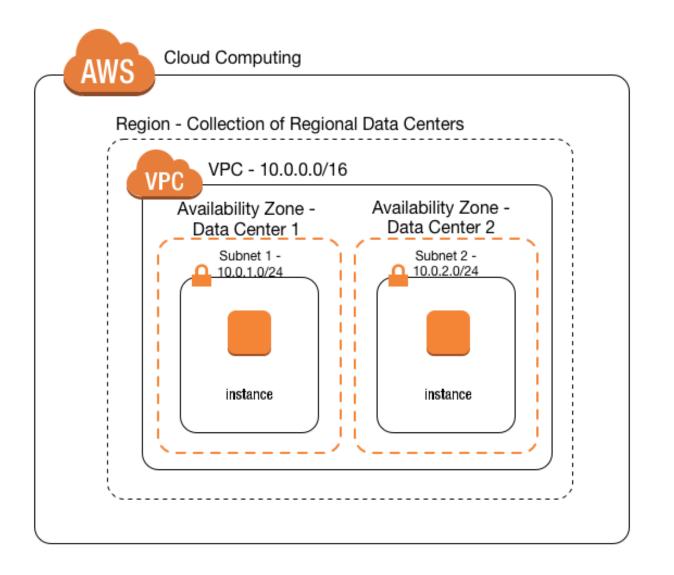
^{*} Required

Watch these (and other) limits

VPC and Subnets

Resource	Default limit	Comments
VPCs per Region	5	The limit for internet gateways per Region is directly correlated to this one. Increasing this limit increases the limit on internet gateways per Region by the same amount. You can have 100s of VPCs per Region for your needs even though the default limit is 5 VPCs per Region. You can request an increase for these limits using the Amazon VPC limits form.
Subnets per VPC	200	-
IPv4 CIDR blocks per VPC	5	This primary CIDR block and all secondary CIDR blocks count toward this limit. This limit can be increased up to a maximum of 50.
IPv6 CIDR blocks per VPC	1	This limit cannot be increased.

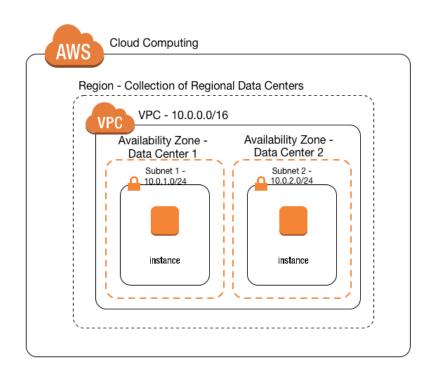
VPC and Subnets - Summary

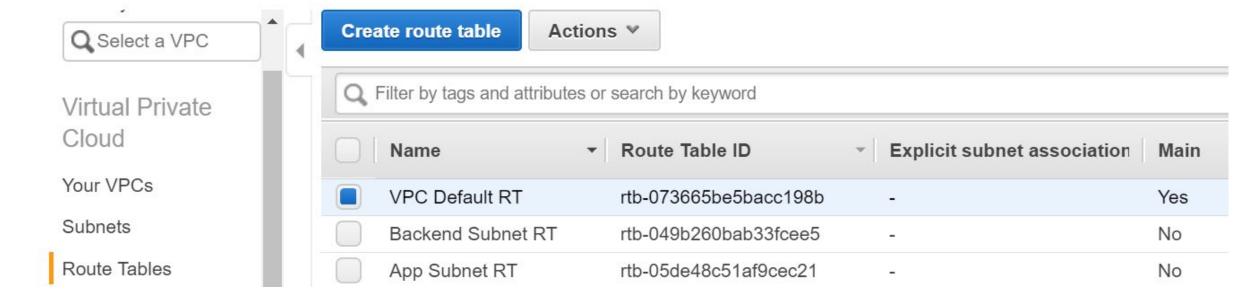


Networking Building Blocks

Routing Tables

- Every subnet has a route table, and a single route table can be associated with multiple subnets.
- Route tables act like a source-based policy-based routing (PBR) rule. In other words, you can choose which direction packets should go based on the subnet the instance is in.
- Create a routing table and then associate to a subnet
- If no routing table is associated to a subnet, then per-VPC default RT is used (Main)

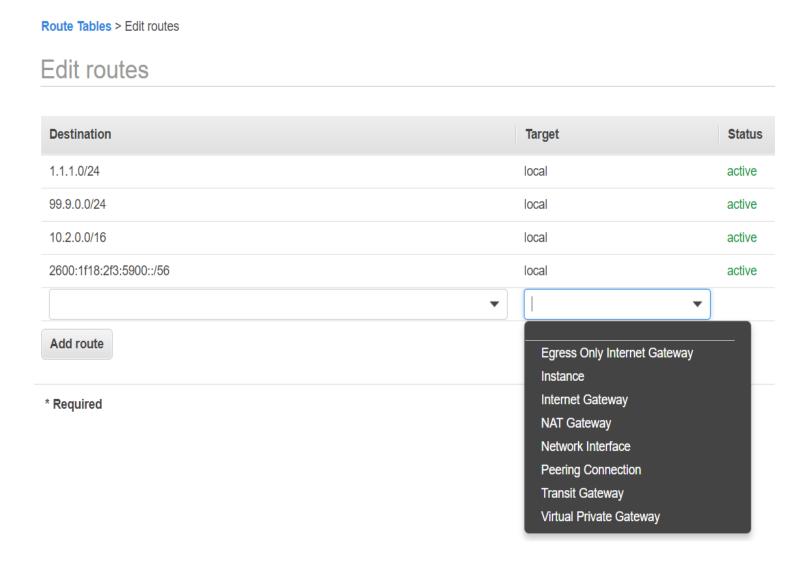




Add route

Like in a traditional network - specify network in CIDR format and target

CIDR ranges for VPC can't be changed, with target **Local** – this is Implicit AWS VPC router



Instance - Virtual machine (VM) — Firewall or NAT Instance, as an example

Network interface - "floating" NIC, can be assigned to a running instance, so think failover between two active/passive instances, also for VPC endpoints

Internet Gateway – AWS-managed construct providing access to Internet from a Public subnet (explained later)

NAT Gateway – AWS-managed construct providing access to Internet from a Private subnet

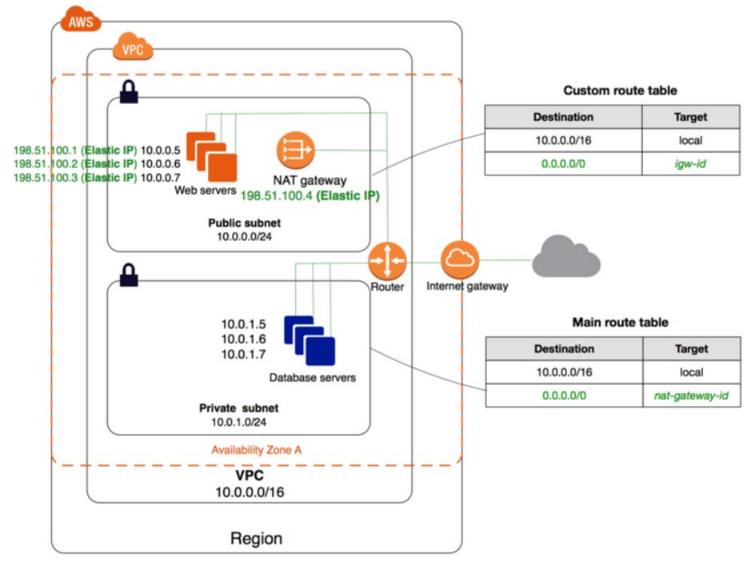
Egress Only GW – access from private network for IPv6 (as no NAT in IPv6)

Peering connection/Transit Gateway – for VPC interconnects

Virtual Private Gateway – for On-Prem interconnects, or for VPC interconnects in some cases

Networking Building Blocks

Internet and NAT Gateways



You may see references to *public* and *private* subnets. There is no difference between them except the presence of Internet GW. IP Addresses assigned to Instances interfaces are always private.

Public subnet has Internet GW

which is AWS logical construct providing access to Internet. AWS builds **one-to-one NAT** to associate Instance interface IP to Public IP addresses (aka Elastic IPs)

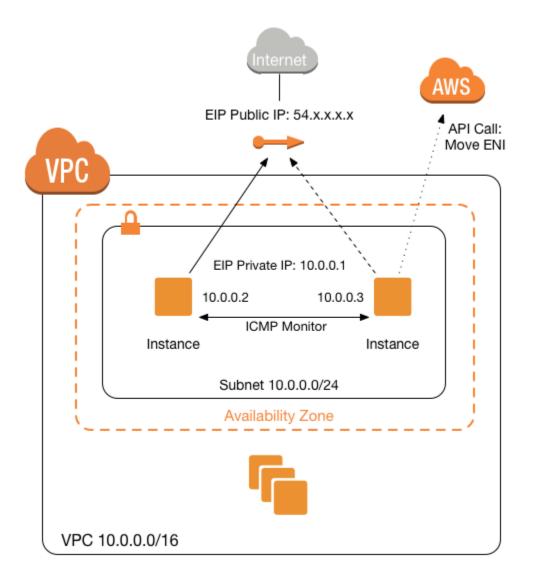
Private subnets have no Internet GW, so there is no 1-1 NAT and direct access to/from Internet. Instances have to use NAT gateway (AWS Managed) or any Instance providing NAT or proxy services (firewall running as a VM, for example).

Note these NAT instances should have EIP, so be in located a Public subnet.

Note: NAT Gateway is billed per hour and not included in free tier.

Networking Building Blocks

Elastic Network Interfaces



Elastic network interface (ENI) that moves between two or more instances is similar to VIP. This is also called a *floating ENI*.

Elastic network interfaces can be used to provide failover between FW instances.

From Palo Alto documentation:

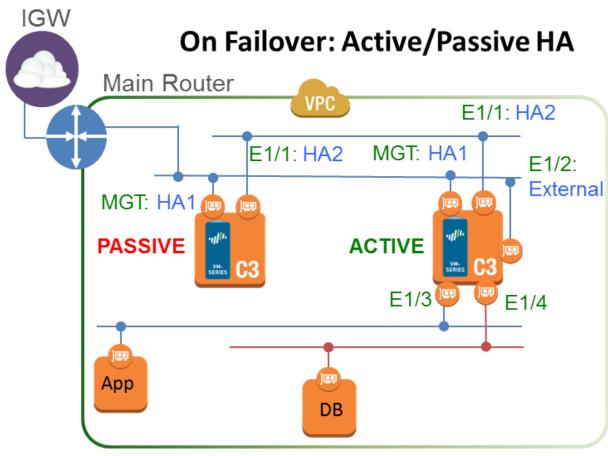
When the passive peer detects this failure it becomes active and triggers API calls to the AWS infrastructure to move all the dataplane interfaces (ENIs) from the failed peer to itself.

Cloudwatch can also trigger Lambda functions to move ENIs between instances.

Also consider alternative designs (such as Load balancers)

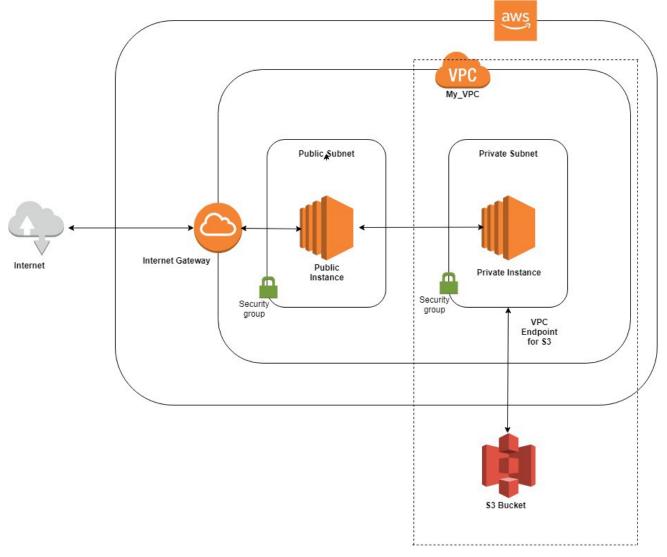
On Setup: Active/Passive HA **IGW** Main Router E1/1: HA2 E1/1: HA2 MGT:HA1 MGT: HA₁ E1/2: External **ACTIVE PASSIVE** E1/4 E1/3 DB

Palo VM series failover



Networking Building Blocks

VPC Endpoints



All AWS SaaS and PaaS services have **public IP addressing**, so your instances, even located in Private subnets should have Internet access to consume them.

From security perspective it is not always allowed.

Solution – **VPC endpoints**

Direct connectivity to AWS without requiring an internet gateway, NAT device, etc.

Kind of "shortcut" to AWS from VPC. Traffic between VPC and the other service does not leave the Amazon network, so it is more like a private "datacentre".

Two types of VPC endpoints – Gateway and Interface, depending on AWS service.

Create Endpoint

A VPC endpoint allows you to securely connect your VPC to another service.

An interface endpoint is powered by PrivateLink, and uses an elastic network interface (ENI) as an entry point for traffic destined to the service. A gateway endpoint serves as a target for a route in your route table for traffic destined for the service.

Service category

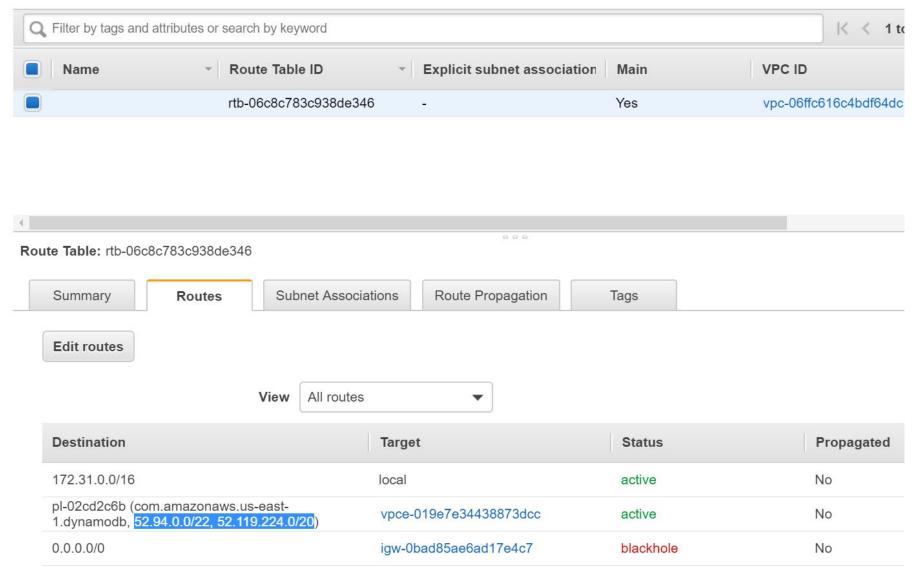
- **AWS** services
- Find service by name
- Your AWS Marketplace services

Service Name Select a service



Filter by attributes			
Service Name		Owner	Туре
com.amazonaws.us-east	-1.config	amazon	Interface
com.amazonaws.us-east	-1.datasync	amazon	Interface
com.amazonaws.us-east	-1.dynamodb	amazon	Gateway
com.amazonaws.us-east	1 000	amazon	Interface

Gateway VPC endpoint



Gateway is a route to a virtual AWS "gateway"

Destination is a prefix list automatically populated by AWS with network ranges specific to the service.

Note blackhole route to igw-xxxx. Route pointing to deleted Internet GW

Interface VPC Endpoint (also referred as PrivateLink endpoint)

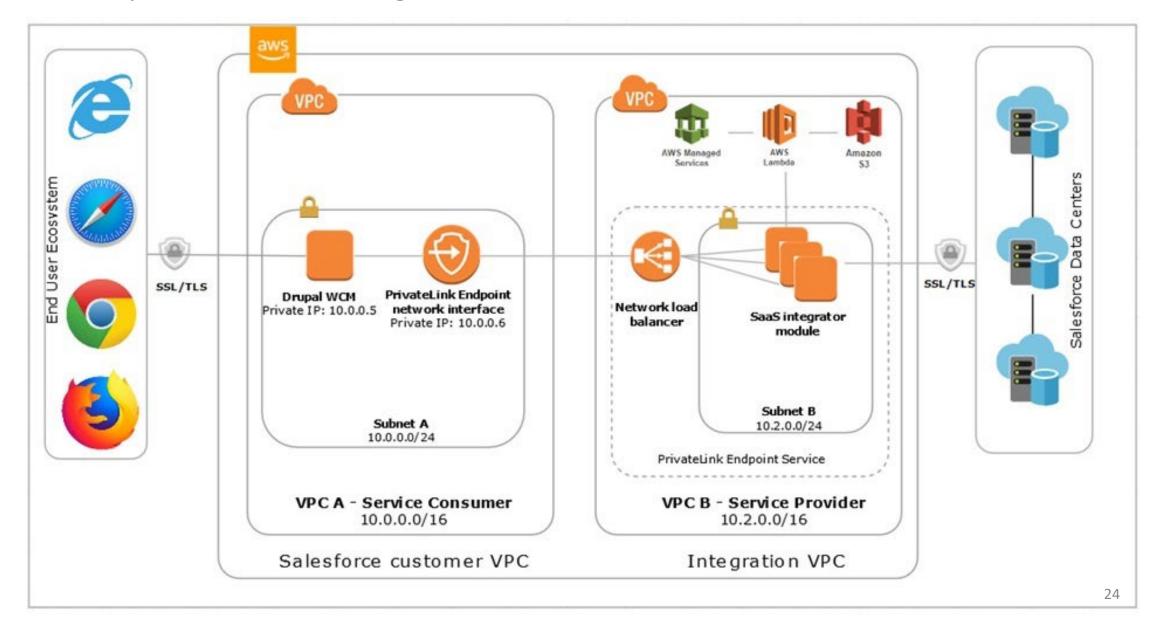
Interface VPC Endpoint looks like a Network interface from other VPC, enabling direct connectivity to the service via private IP addresses. PrivateLink provides private SaaS between VPCs.

Uses concepts as Client VPC and Provider VPC.

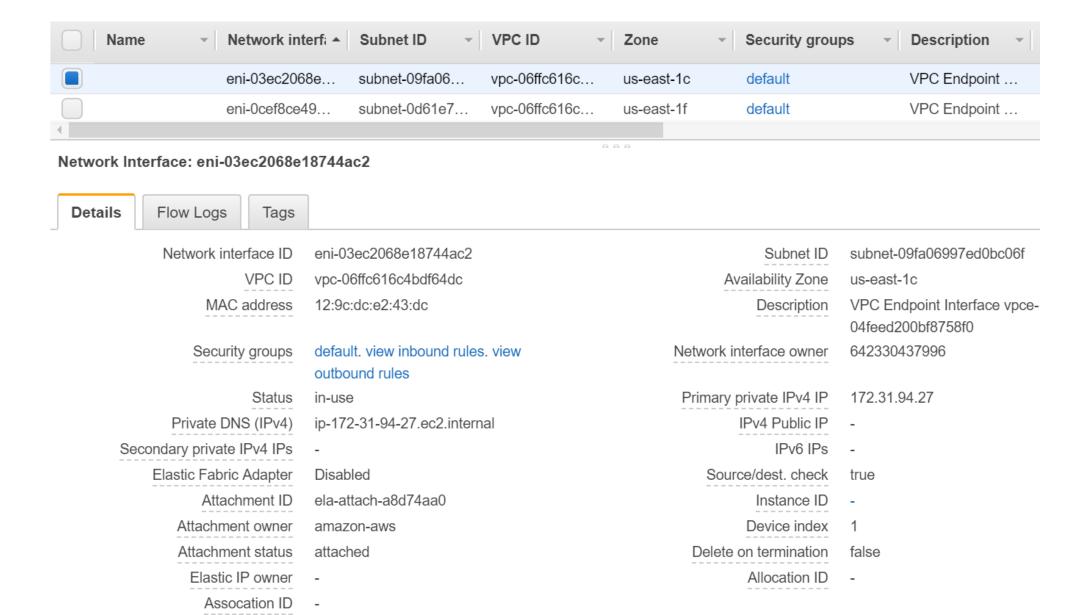
Network Load Balancer's interface from Provider VPC is exposed to Consumer VPC.



Example – connecting to Salesforce via PrivateLink



After creating VPC Interface endpoint it appears as a Network interface directly in your subnet with subnet's IP address:



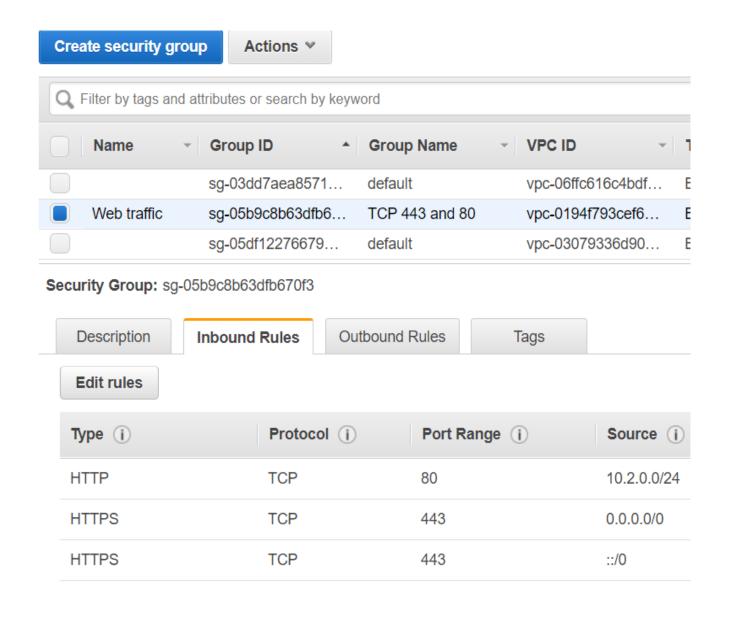
25

Network control:

Security Groups
Network Access Lists

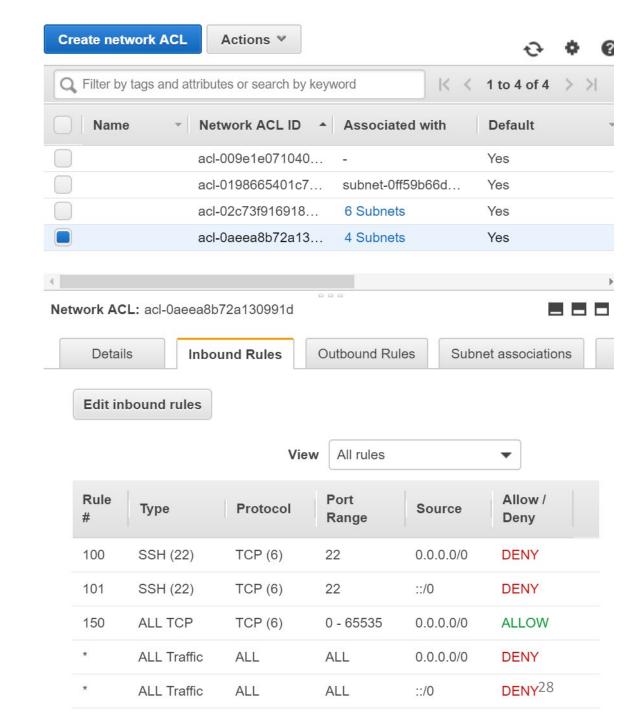
Network Control Security Groups

- Per-Instance Hypervisor-level firewall
- Enforced at the Instance level or Interface (if multiple interfaces)
- Stateful if traffic is allowed by inbound rules, outbound traffic automatically allowed
- Only ALLOW rules, no DENY, all traffic is denied unless explicitly allowed
- Possible to use Security Group names as source
- No order just ALLOW statements

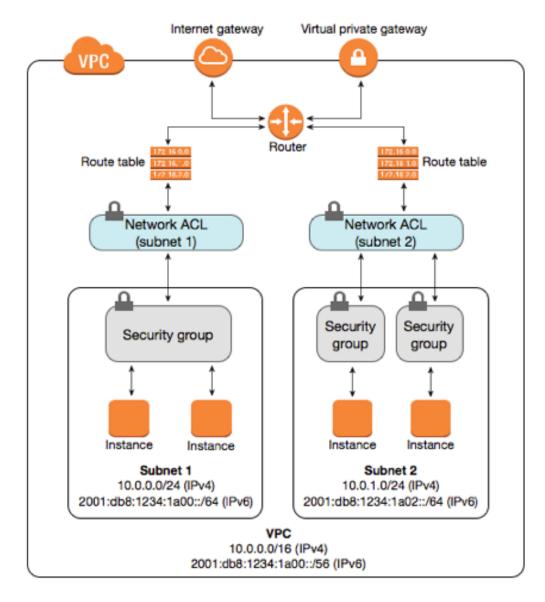


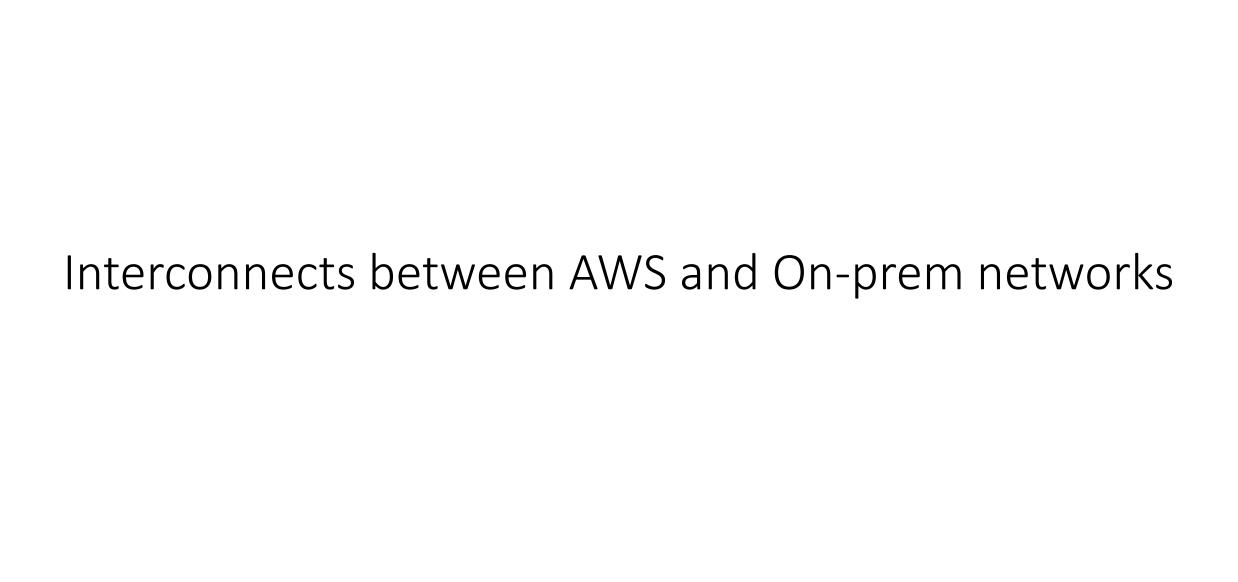
Network Access Lists

- Per-subnet
- Stateless Return traffic must be explicitly allowed by rules
- Process rules in number order when deciding whether to allow traffic
- Very similar to traditional routers' ACLs
- Can be pretty complex to manage if used together with security Groups



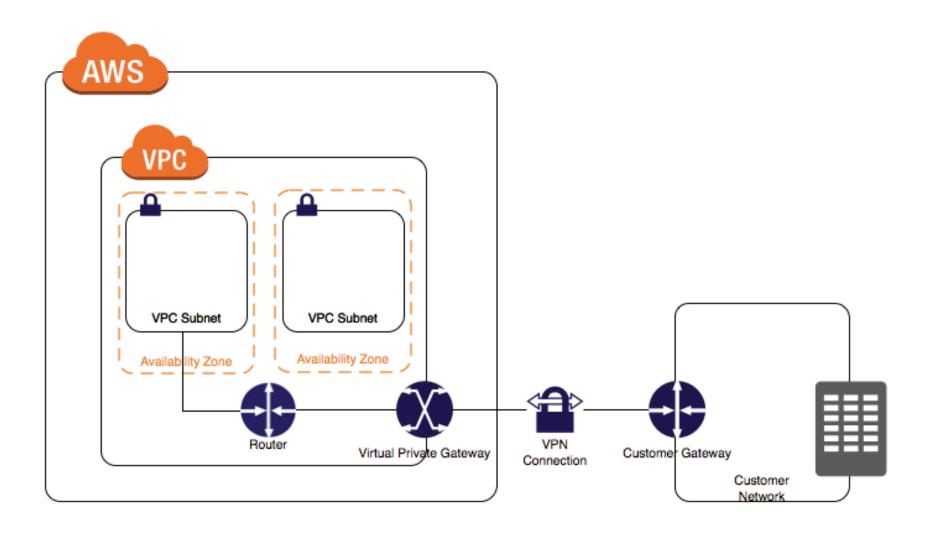
Network Control - Summary





VPN Gateway

Logical construct, create and attach to VPC

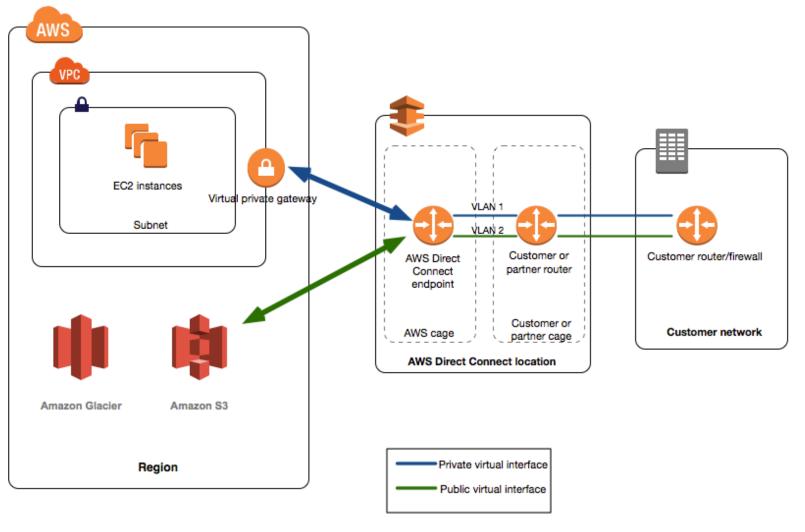


Create VPN Connection

Select the virtual private gateway and customer gateway that you would like to connect via a VPN connection. Yo your customer gateway information already. 0 VPN to DC1 Name tag vgw-0ca92c330aefad621 Virtual Private Gateway **Customer Gateway** Existing New 0 IP Address 84.53.198.10 0 **BGP ASN** 65000 Certificate ARN Select Certificate ARN Dynamic (requires BGP) **Routing Options** Static

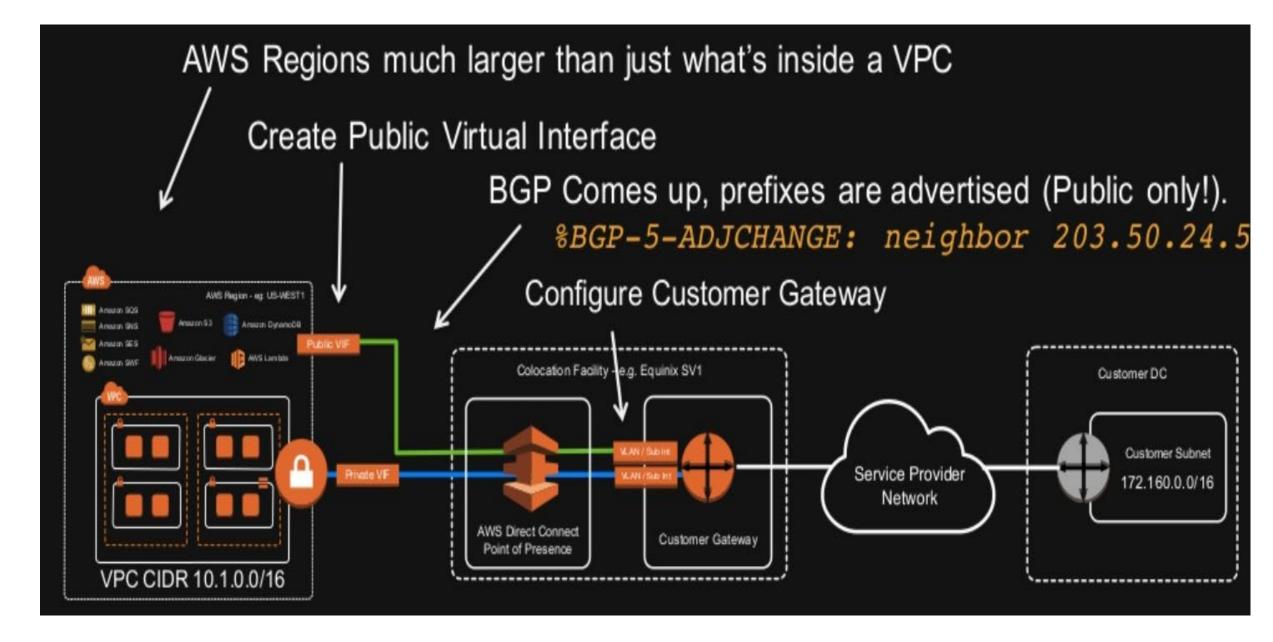
- Receives routes over BGP propagated automatically into VPC Routing table
- Advertises VPC's CIDR subnets via BGP
- Alternatively create static routes in VPC pointing to VPN GW
- Supports PSK or CA-based authentication
 - Builds two tunnels to Customer GW, active/failover, but can be used act/act with BGP
- 1.25 GBps limit per tunnel

Alternative to VPN – Direct Connect



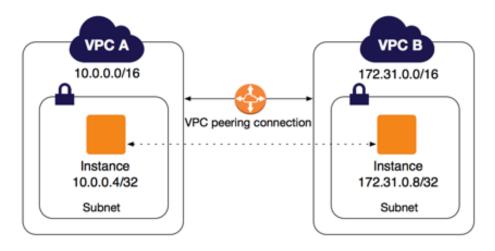
- Dedicated links to AWS
- Direct 1/10Gig optic link if in the same Datacentre
- Or via SP such as Megaport, Telstra, Vodafone, etc.
- In any case it is a VLAN and L3 point-to-point interface
- Uses BGP
- Use at least two DX, and you can leverage BGP standard traffic engineering techniques – AS-Path prepend, communities, etc.
- Traffic is not encrypted. Use
 IPSec VPN over Direct connect if
 there is such a requirement

Two types of Direct Connect Interfaces – Public or Private

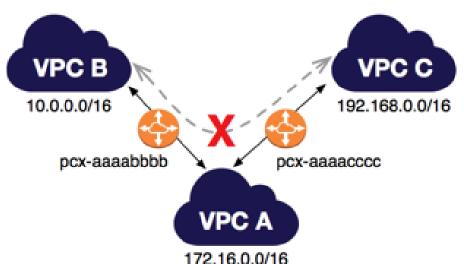


Interconnects between VPCs

Interconnect VPC - VPC Peering



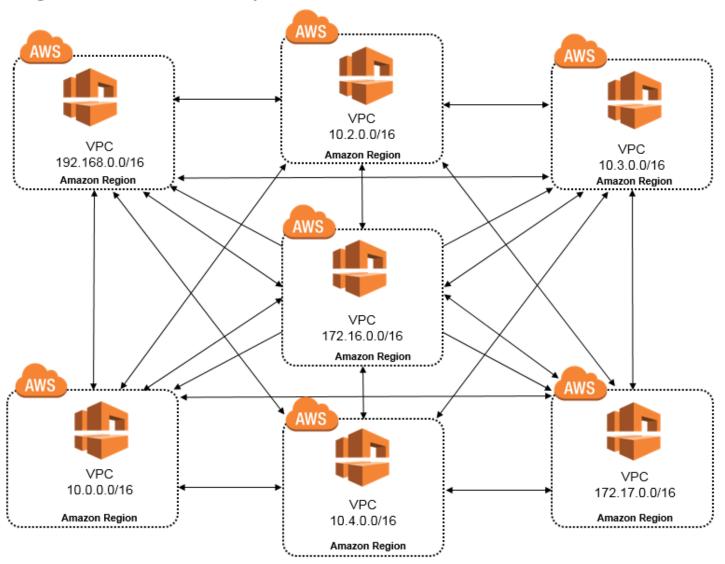
- Add peering connection to another VPC
- Add route to routing table pointing to VPC peering link



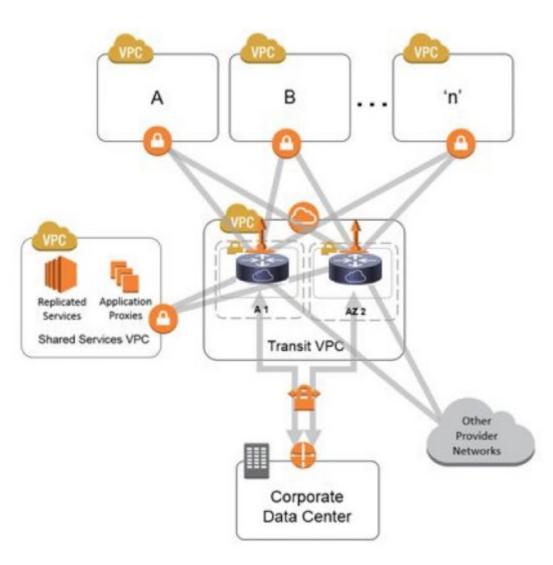
- Transitive peering is not supported, so VPC B can't access VPC C via VPC A.
- Solution Full-mesh peering or Transit Gateway

Full-mesh peering

Max 50 peering connections per VPC. Not scalable, but the cheapest option

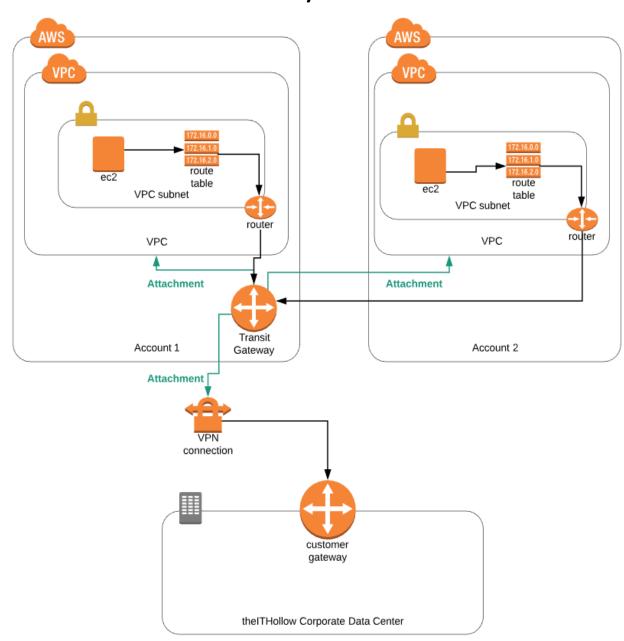


VPN GWs is not only for On-prem connectivity – Transit VPC



- Remember non-transitive routing rule?
- Transit VPC is no overcome this and avoid building full-mesh VPC
- Two Instances are built running Cisco CSR 1000v, Fortigate, Palo, anything which supports IPSec tunnels
- Assign public IP addresses to these Instances
- Build IPSEC VPN with spoke VGWs
- Establish BGP via IPSEC VPN tunnel
- Use AS-path prepend in BGP for one of the firewalls, so all spokes use the same firewall based on BGP
- Expensive, complex, and pretty common, but not relevant anymore after AWS introduced Transit GW

Transit Gateway

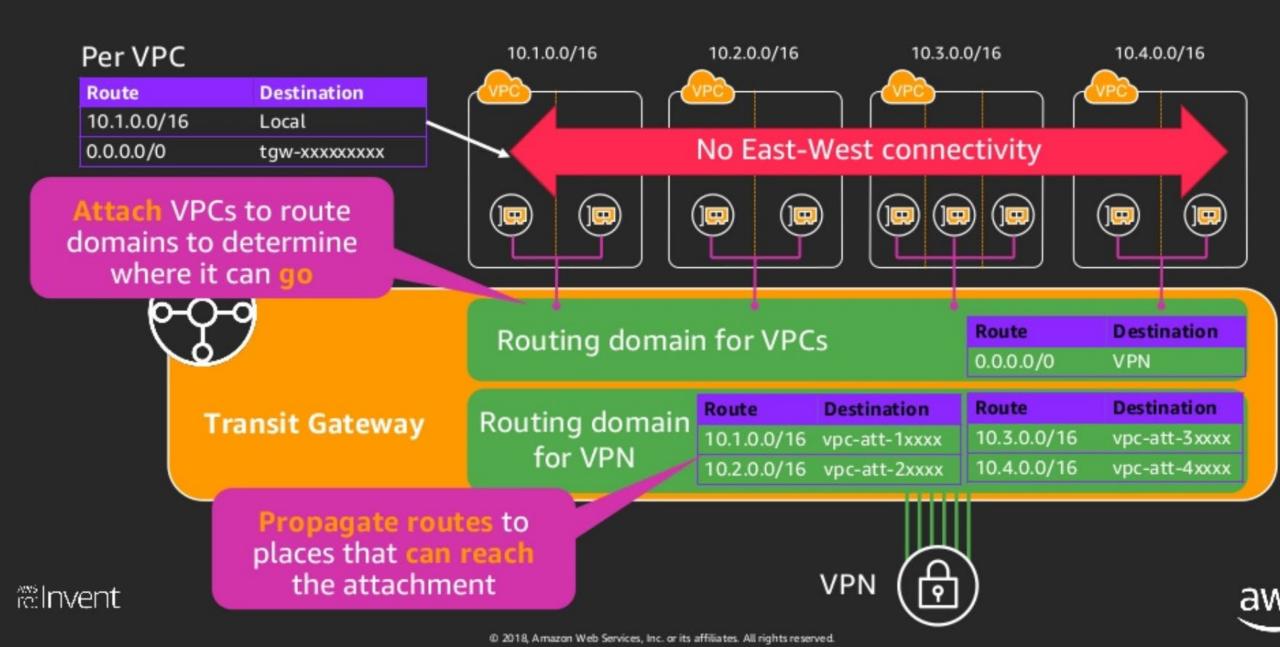


Transit GW is similar to a multi-VRF router which imports/exports routes between VRFs (VPCs in our case) creating Route Domains

TGW concepts:

- TGW Route tables (don't confuse with VPC routing tables) – routing domains with routes imported from VPCs or built manually
- Attachment (routes to import FROM) and propagation (routes imported TO) – Can be VPC, VGW, Direct Connect
- Allows to built complex scenarios, restrict East-West traffic, force traffic to Security VPCs, etc
- Not free Price per GB of data processed and per attachment/hour

Isolated: Transit Gateway route domains



AWS Managed Network Services

Load Balancers

WAF

Shield

AWS Managed Network Services Load balancers

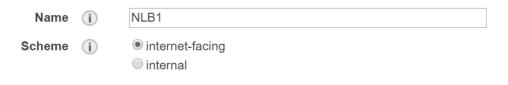
- Two types of managed Load-balancers:
- Network LB
- Application LB
- Classic LB still exists, but considered as Legacy

AWS Managed Network Services

Network Load Balancer

Basic Configuration

To configure your load balancer, provide a name, select a scheme, specify one or more listeners, and select a network. configuration is an Internet-facing load balancer in the selected network with a listener that receives TCP traffic on port



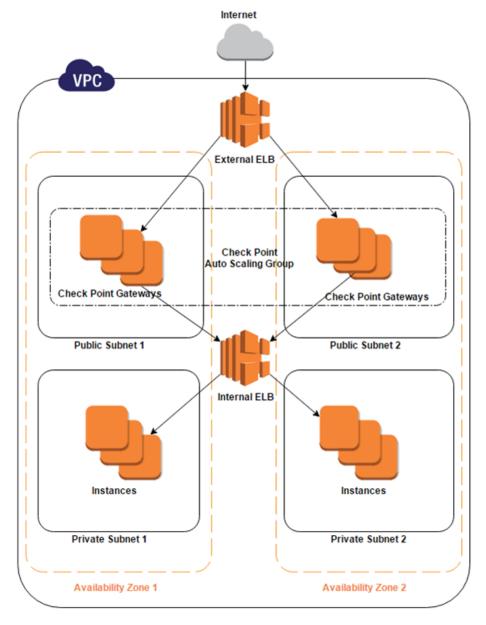
Listeners

A listener is a process that checks for connection requests, using the protocol and port that you configured.

Load Balancer Pro	otocol	Load Balancer Port
TCP	▼	443
TCP_UDP	▼	53
Add listener		

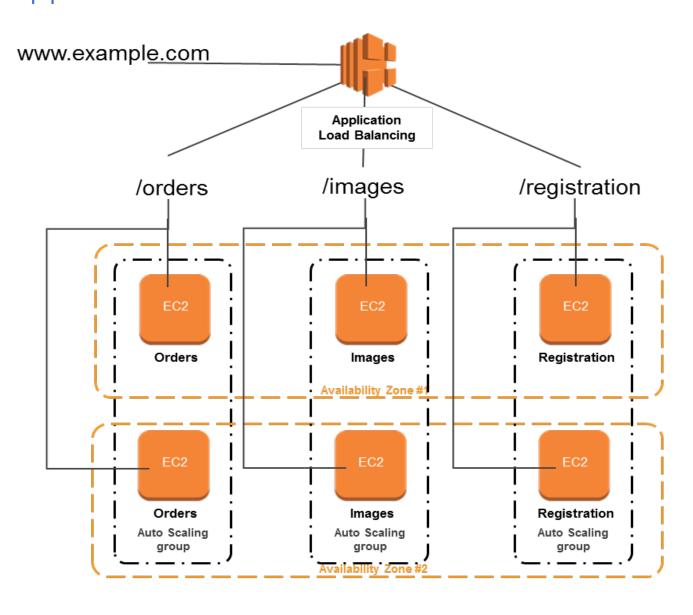
- Network Load Balancers are used to route traffic at Layer 4
- High-throughput, millions request per seconds
- TCP or UDP
- Integrates with EC2 Auto Scaling Elastic to enable you to attach one or more load balancers to an existing Auto Scaling group. After you attach balancer, the load automatically registers the instances in the group and distributes incoming traffic across the instances.
- Add Listeners and then Target Groups

Common design – NLB Sandwich for ingress traffic



- VPC spans 2 availability zones
- Each availability zone has a public subnet and a private subnet.
- An External ELB sends incoming traffic to a Firewall Auto Scaling Group residing on the two public subnets.
- The FWs in the group inspect the traffic and forward the traffic to an Internal ELB.
- The Internal ELB sends incoming traffic to a group of servers residing on the two private subnets.
- The FW Auto Scaling Group is set up to increase or decrease the number of FWs based on AWS Cloud Watch metrics like other instances.

AWS Managed Network Services Application Load Balancer



- Application Load Balancers are used to route HTTP/HTTPS (or Layer 7) traffic
- Use the concept of rules, targets, and target groups.
 Rules determine how to route requests.
- Host-based or path-based routing rules
- Offers SSL offloading.
 Certificates provided by AWS
 Certificate Manager and
 deployed on your load balancer
 can be renewed automatically.

Application Load Balancer rules: if/then



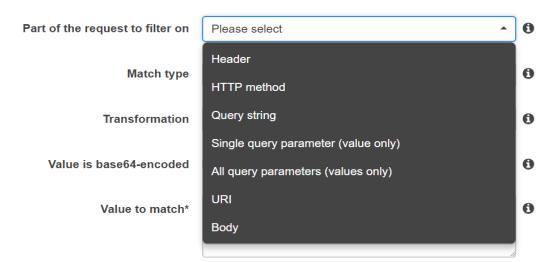
AWS Managed Network Services Web Application Firewall

Create string match condition

A string match condition contains a list of the strings that appear in web requests that you want to allow or block. Learn more Name* Global (CloudFront) Region* Choose Global (CloudFront) to create AWS WAF resources to use with Clo AWS Region to create AWS WAF resources to use with an Application Loa Type* String match To create a standard string match condition, choose String match. To create

Filter settings

Specify the settings that you want to use to allow or block web requests. If you add more than one filter to a string match conditic the request to match the string match condition. (The filters are ORed together.)



- Uses concepts of Web ACLs based on various conditions, such as IP addresses, Headers
- Integrates with **Application Load** Balancer, API Gateway or CloudFront (AWS CDN)
- Rules can be built programmatically

AWS WAF

Web ACLs

Rules

Marketplace

Conditions

Cross-site scripting

Geo match

IP addresses

Size constraints

SQL injection

String and regex matching

AWS Shield

Summary

Protected resources

Incidents

Global threat environment

AWS FMS

Security policies

Rule groups

Settings



AWS WAF is a web application firewall service that helps protect the websites and web apps that you deliver with Amazon CloudFront and ELB Application Load Balancers. Create web access control lists (web ACLS) that define which HTTP and HTTPS requests to allow, block, or count. Learn more

Configure web ACL



Web traffic filtering with custom rules

Create custom rules that can allow, block, or count web requests based on originating IP addresses or strings that appear in web requests.



Block malicious requests

Configure AWS WAF to recognize and block common web application security risks like SQL injection (SQLi) and cross-site scripting (XSS).



Tune your rules and monitor traffic

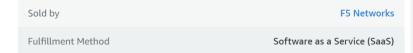
Review details about the web requests that AWS WAF allows, blocks, or counts, and update rules to thwart new attacks.

AWS WAF – Managed rules



Product Overview

Protect against web exploits. F5 Web Exploits OWASP Rules for AWS WAF, provides protection against web attacks that are part of the OWASP Top 10, such as: SQLi, XSS, command injection, No-SQLi injection, path traversal, and predictable resource. Protect your applications and services with F5, the trusted leader in web application security.



Highlights

- Increase protection against web attacks
- Integration with AWS WAF makes it easy to deploy without changes to your infrastructure
- F5 manages your AWS WAF rules, so you don't have to.

- An option to use thirdparty managed rules – see AWS Marketplace
- See rating and reviews

Pricing Information

This software is priced along a consumption dimension.
Your bill will be determined by the number of units you use. Additional taxes or fees may apply.

F5 Rules for AWS WAF - Web exploits OWASP Rules

Units

Cost

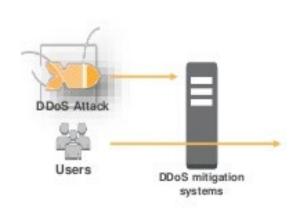
Charge per month in each available region (pro-rated by the hour) \$20 / unit

Charge per million requests in each available region \$1.2 / unit

AWS Shield

DDoS protections built into AWS

- Protection against most common infrastructure attacks
- SYN/ACK Floods, UDP Floods, Refection attacks etc.
- No additional cost



amazon

- AWS Shield is a managed DDoS protection
- Two flavours –
 Standard and
 Advanced

- Standard Free, already turned on, notconfigurable
- Advanced \$3K
 USD per month

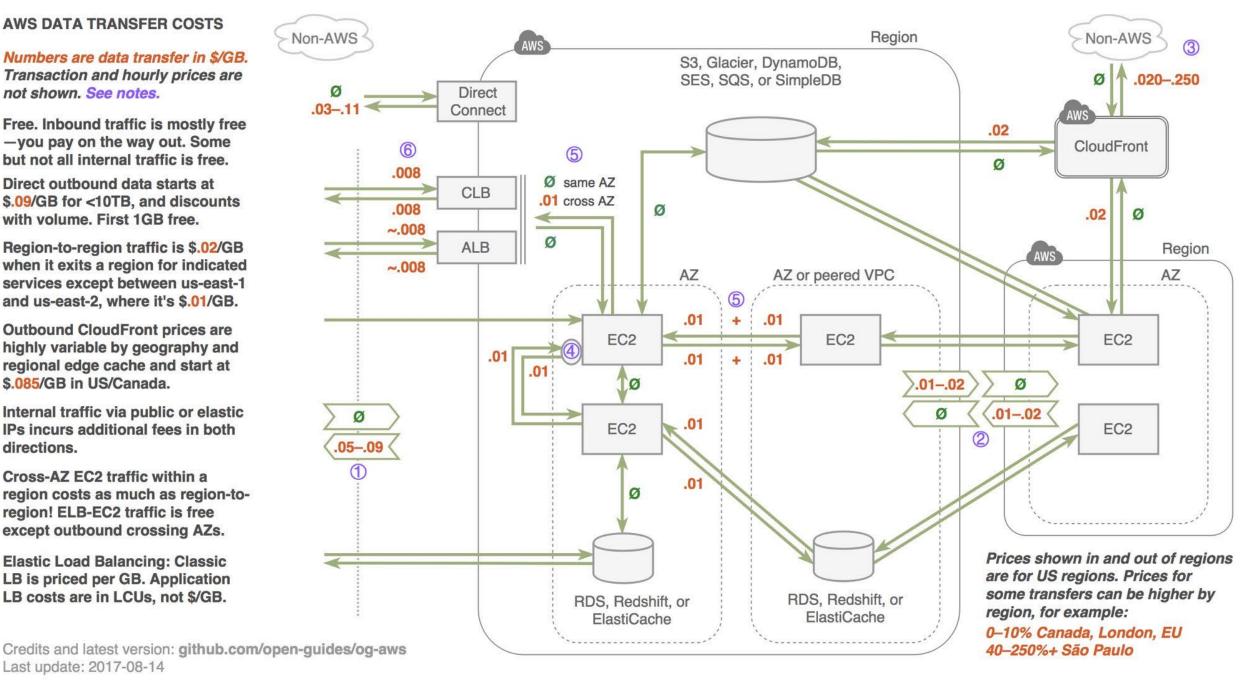
And finally....

AWS DATA TRANSFER COSTS

Numbers are data transfer in \$/GB. Transaction and hourly prices are not shown. See notes.

- Ø Free. Inbound traffic is mostly free -you pay on the way out. Some but not all internal traffic is free.
- Direct outbound data starts at \$.09/GB for <10TB, and discounts with volume. First 1GB free.
- Region-to-region traffic is \$.02/GB when it exits a region for indicated services except between us-east-1 and us-east-2, where it's \$.01/GB.
- **Outbound CloudFront prices are** highly variable by geography and regional edge cache and start at \$.085/GB in US/Canada.
- Internal traffic via public or elastic IPs incurs additional fees in both directions.
- Cross-AZ EC2 traffic within a region costs as much as region-toregion! ELB-EC2 traffic is free except outbound crossing AZs.
- **Elastic Load Balancing: Classic** LB is priced per GB. Application LB costs are in LCUs, not \$/GB.

Last update: 2017-08-14



Replicating on-prem networks in the Cloud?



- Lift-and-shift is not the best approach in most cases
- Know native tools offered by AWS: CloudFormation, CloudWatch, GuardDuty, etc.
- See what can be done with event-driven automation
- Public Cloud
 Networks are still the same IP networks, yet different

Good read

AWS Certified Advanced Networking Official Study Guide (a bit outdated, but still OK) https://www.amazon.com/Certified-Advanced-Networking-Official-Study/dp/1119439833

Amazon VPC for On-Premises Network Engineers from Nick Matthews

https://aws.amazon.com/blogs/apn/amazon-vpc-for-on-premises-network-engineers-part-one/

https://aws.amazon.com/blogs/apn/amazon-vpc-for-on-premises-network-engineers-part-two/

https://packetpushers.net/podcast/show-387-aws-networking-view-inside/ - useful links in this podcast!

Re:Invent videos related to networking, for example, from Matt Lehwess (he's an Australian;-))

Slack:

https://aws-programming-tools.slack.com/

Labs:

https://github.com/awslabs/aws-well-architected-labs