Chapter 5:
Large Enterprise Cloud Design Case Study

Background: This chapter is devoted to the case of a fictitious international network equipment manufacturer called Castlencia, headquartered in California, USA. Castlencia was founded in 2013 by a group of ambitious network and hardware engineers who were trying to break into the market of white-boxes by bringing their own data center top-of-rack switches to the market. Although their initial goal was to remain focused on the design aspect of hardware without entering the market of silicon, firmware, or operating systems, due to the strong demand over time, they acquired multiple other vendors. This exponential growth significantly increased their revenue, but at the same time, they faced very unique challenges to keep their business agile, ironically without having to focus too much on expanding their own data centers which was never the core of their business.

The Politics: Castlencia is no exception. Over the past eight years, although the CEO has never changed, the company witnessed two different CIOs. With the first CIO, he and his circle would meet behind closed doors to decide what to purchase and design next. As of 2019, the IT department underwent a major reorganization. Now, the new CIO, Ilene, has one main mandate from the CEO; to turn the ship around and make IT an enabler for innovation rather than a "bump in the road".

Last year, they also started their architecture board, which alongside the other new ITIL constructs, attempts to bring order and logic to the decisions. According to Ali, one of the network engineering chief architects, they all feel more included in the decision-making process.

The new CIO has also presented the cloud-first idea to the leadership and received a decent annual budget with the promise of improving business agility at the minimum cost. This initiative is particularly important to the CEO because they lost major opportunities on several occasions due to the multi-month IT processes of ordering hardware, renting colocation spaces, ordering circuits, and turning up new data centers. The average time required to enable power, space, and cooling for a new footprint is now close to nine months. This can slip further if they have to hire new people. Last but not least, the leaders have their eyes on a potential IPO which would bring the company under a lot more scrutiny. Hence, in partnership with the CISO, Ilene has the tough mission of securing the environment as well. Castlencia has barely passed the most recent audits and has fallen victim to four major security incidents in the past 18 months.

Castlencia also has recently hired a new CISO, Albert. The interview with him quickly reveals that the initial steps in the cloud-first strategy had raised many concerns, but now he is coming along with the project as he and the team earn each other's trust. For instance, one major concern was the S3 traffic ending up on the Internet, both between the VPCs and S3 and between the offices and S3. Furthermore, he expresses concerns about unencrypted private data and flimsy security solutions such as simple NAT that would give people, especially the engineering teams, a false sense of security. This alone cost the company $2.6M last year.

Description of the Buildings

The idea of keeping the business in Southern California didn't materialize, and today they have a presence in 32 countries and data center footprints in eight different cities. They currently operate 60 buildings.

Today according to their IT department, Castlencia has four types of buildings:

1. *Type 1*: This includes four colocation spaces they rented from data center vendors in different countries and three semi-data centers built in the basements in some of their large buildings. The latter is widely despised among the IT teams as these semi-data centers' physical and power issues

are the root cause of outages over the last three years. They had a major water leak, a few overheating cases, and a handful of power outages in 2018-2019 caused by failed generators or faulty UPSs. Rumor has it, the company is unwilling to move everything to the colocations to keep the costs down.

2. *Type 2*: This covers all the offices with more than 2000 occupants. This includes headquarters or campuses in Los Angeles, Frankfurt, Mumbai, and Hong Kong. The CIO and CEO both sit in the LA office. Castlencia currently has four Type 2 buildings.
3. *Types 3*: This type of building includes all the medium-sized offices usually housing 20-2000 users. Some of the most critical engineering design and customer support offices are housed here. Castlencia currently has 15 Type 3 buildings.
4. *Type 4*: Small or satellite sales offices with less than 50 users. Some of these users are almost always remote. Castlencia currently operates 34 Type 4 buildings.

Description of the Network Environment

Castlencia, due to the constant wave of acquisitions of smaller companies, rapidly turned into a ragtag environment. But over the past 12 months, through multiple major costly projects, the new CIO attempted to retrofit and standardize the network, systems, and storage areas as much as possible.

Currently, every office has its own private AS number, while the company owns a 2-byte public AS number. Castlencia also owns a /16 public IP address range exclusively used in cases where NAT or other techniques cannot be done.

As part of the retrofit project, the IT team also replaced all the random routing protocols with BGP and OSPF. Currently, at a very high level between the buildings, inside their data centers, and the outside world, they run BGP while they exclusively use OSPF within their building. This forms a layer 3 access network in 99% of cases.

The company operates a wide range of connectivity options with various carriers. Currently, they have DIA, MPLS L3VPN, VPLS, LTE, and dark fiber. Except for a few corner cases, all the Type 4 offices have DIA and backup LTE connections, while the Type 2 and Type 3 have a mix of DIA, leased lines and MPLS circuits. VPLS and dark fiber are used almost exclusively for data center connectivity.

From a cost perspective, it's the CIO's absolute preference not to add any major item to the recurring fees, and every new circuit request must be approved by her.

Description of the Tools and Systems

Castlencia has a wide range of software and tools, but at a high level, they fall into two main categories: productivity and core. The productivity tools include email, virtual desktops, internal websites, and databases for a variety of purposes, including payroll and HR. This group includes both on-prem and SaaS applications. The core software includes AutoCAD and other tools used by the engineering departments. Each category has its own stakeholders and challenges. Although not related to the core of their business, the group of productivity applications still act as the lifeline of the business. The core applications have powerful stakeholders inside the company, and the SVP of engineering, Sherwin, has shown no tolerance to service outages.

It goes without saying that many of their cloud-based applications need to constantly communicate with on-premises resources. In some cases, it's temporary, but in the case of engineering tools, this might be permanent as the engineering department would hold on to some of their physical data center footprint.

Description of the Overall Cloud Environment

In the last two years since Castlencia embarked on their cloud journey, they have extensively worked with three public cloud providers: Oracle Cloud Infrastructure (OCI), Microsoft Azure, and Amazon Web Services (AWS). They have yet to develop a solid framework across the board, but each department has taken its own path. It was on the to-do list of the new CIO when she joined the company to optimize and preferably organize the multi-cloud environment. For example, if certain application servers reside in Azure, how do they optimize access to their high throughput databases in Oracle Cloud? Or on the SaaS front, how do they optimize access to their critical SaaS applications? The testing in 2020 revealed that some of them were more than 12 hops away from their key buildings.

Castlencia currently uses SaaS applications from multiple vendors, including Oracle, Microsoft, and ServiceNow. ServiceNow provides their change management, RCA and ticketing system.

Description of IP Addressing Strategy

Although their IP Address Management (IPAM) tool is not the most organized, its data is still very reliable. In addition to the /16 public IPv4 range, the company decided years ago to use the 10.x.y.z/8 range for internal addressing. People drifted from the original policy over time within the engineering environment by using other CIDR blocks from the RFC-1918 range. The blame was assigned to the acquisitions. Currently, the 10.8.0.0/13 subnet is carved out for cloud engineering purposes, and the 10.9.0.0/16 is assigned to AWS networking. The IT department is confident that if needed, they could easily allocate more addresses.

Castlencia currently has no plans to use IPv6 on the cloud yet, although some parts of engineering, mostly for hardware testing and development purposes, use OSPFv3.

Finally, the company has never needed multicast beyond its labs. Hence, the entire Single Source Multicast (SSM) environment is confined within certain engineering buildings with no communications over the WAN. The CIO, based on some bad experiences in the past with multicast-related bugs, rejected any plans to enable multicast on their backbone network.

Description of the VPC addressing strategy

Since their AWS environment was built after 2013, all their designs are based on the modern VPC-based environments without any trace of AWS EC2-Classic. They have also worked with AWS to increase the limitation of 5 VPCs per region.

Since AWS only supports the CIDR blocks with subnet masks between /16 and /28, all of their VPCs have a subnet mask of /24. This decision was made during the early days of their cloud-first project and later led to scaling issues as they started building larger VPCs with more than 250+ addresses in each. One of the notable issues they ran into as they depleted the primary CIDR block ranges for their large VPCs was the concept of block association restrictions. It turned out if their primary CIDR was within the 10/15 range, they could not add a CIDR block from the 10/16 range. In some cases, the IT department ended up adding subnets from the carrier grade NAT range of 100.64.0.0/10 to those large VPCs. Admittedly, before deploying this range, many engineers in the IT department thought the 100.64 range was a public IP range.

The IT department uses the VPC DHCP options sets to propagate the IPs and some other addresses, including those of the NTP servers. In 99% of the cases, this solution works well. There are a small number of cases where, within the same VPC, they had to use different NTP addresses or domain names. This issue surfaced when the software engineering team tested a new code and needed different NTP addresses on different EC2 instances. This didn’t work out as they discovered the DHCP options sets were at the VPC level, and even worse, they could not easily be edited after creation. Their painful solution was to move the instances running the new code to a dedicated VPC with a new DHCP options set.

Prior to the DNS retrofit project, most DHCP options sets pointed to the IP address of the on-premises DNS servers for any sort of name resolution; on-premises, AWS resources, other cloud providers, or just the Internet. Although this introduced a latency of no longer than 15ms, it was still noticeable by some of the engineering teams. As you will see in the following chapters, the issue was addressed in later 2020.

Description of the Offices' Internet Connectivity

All of their offices have access to the Internet, but the office of CISO has banned any sort of local egress. What that means from an architecture perspective is that all Type 2, 3, and 4 offices, regardless of their size, would have to use two of the closest data centers to access the Internet or other resources in an active-passive design managed by BGP. The company has built seven security stacks, including Next Generation firewalls, packet capture tools, and proxy servers in each data center. Depending on their location, importance, and size, the offices use a combination of circuits to connect to the data centers. Each data center has a pair of 100Gbps DIA circuits from two different carriers with fully diverse paths that have been verified through Google map files.

Description of the Challenge of Internet Connections in AWS

Castlencia has come a long way. In the early days of their cloud journey, every VPC had its own dedicated IGW. These random IGWs would be protected by a range of techniques depending on the skill-level of the owners of each VPC. This included the use of AWS NAT gateways, server-based firewalls, EC2-based Next Generation firewalls and any combination of them.

For instance, if VPC_eng045 had to be carefully protected, they would install a firewall image from the AWS Marketplace on a large EC2 instance supported by the firewall vendor. Then they would configure a gateway route table on the IGW using the VPC ingress routing technique to push the Internet traffic through the firewall via its ENI before it made its way to the protected subnets.

Castlencia once missed an RFP deadline just because one of these non-redundant EC2-based firewalls initially slowed down and eventually crashed after several hours and did not recover in time.

The cases protected by the simple NAT solution had their own issues. In some instances, since the AWS NAT solution is effectively a PAT construct, they had no way to perform 1:1 IP address mapping in source or destination NAT cases.

Furthermore, the billing model of the NAT constructs made the case a tough sell for the engineering SVP to offer to the CISO and leadership. The main disaster didn't hit until a malware spread across one of the NAT-only VPCs, which had no Next Generation firewalls or anti-malware installed on its resources. Rumor also has it, at some point in one of those VPCs there was a limited case of ransomware that locked the screens, and the company was asked to pay an unknown amount in bitcoin (BTC).

Given all the threats and the security breaches the company faced, the CISO eventually managed to pull the random Internet and NAT gateways by introducing the concept of central inspection VPCs.

Castlencia now has a VPC designated as an inspection VPC containing a stack of 4 Next Generation firewalls in each region. This central VPC is owned by the cloud engineering team and the office of CISO.

<end of snippet>

Technical Correspondence

At this point, you will review a few email exchanges, including the ones between Albert the CISO and Ilene the CIO:

Dear Ilene,

This morning I was appalled to learn that at least two Network Load Balancers on AWS mask the IP address of the visiting clients by replacing those with their own IPs in the access log files of the Apache servers. This punches a big hole in our security posture as I am not sure about the logging capabilities of the NLBs themselves. Can you please have someone look into this ASAP? We need the right IP address of each visitor on the backend servers. The NLBs in question are nlb-uswest-32 and nlb-useast-11, both sitting in front of very critical servers.

Regards,

Albert

***

Hi Jose,

Can you please look into this immediately?

Thank you,

Ilene

***

Hi Ilene,

I am OOTO but I just had someone look into that. Unfortunately, Albert was right. The team should have enabled Proxy protocol version 2. Although the root cause of the misconfiguration would need further studies, just for the records nlb-uswest-32 and nlb-useast-11 are the ones with IP addresses and not instance IDs as their targets. In this case, they need the proxy header to carry the source IP address over to the servers. I will cover this in great detail on our 1:1. Also will have the team check all the other ELBs to make sure, regardless of their type we preserve the source IP address of the clients when they're logged on the backend servers.

Thanks,

Jose

Hi Jose,

The SRE team has called out this issue multiple times to the network and cloud architects, but we have yet to come up with a solution for it. As you know, alb-uswest-2, one of the most critical AWS Application Load Balancers, starts acting up. Actually, I should say that it saturates one of the 5 Web servers behind it for some unknown reason. And this goes on until the SRE team intervenes. Aren't they supposed to distribute the load evenly, or at least close to that, using some sort of hashing algorithm? I know the ALB is configured to keep the sessions sticky, but this one appears to be a bit too sticky!

Regards,

Darry

***

Hey Darryl,

Jose forwarded your email to me. I've heard of the issue but can dig further. I have a strong hunch about this that might actually answer your question. Before going too far, can you please confirm if you're observing any of the following? Are all the sessions coming from one IP address? If so, I suppose the IP address does not belong to a known malicious network. Right?

Thank you

Ali

***

Hey Ali,

Spot on, my friend. The ghost got into one of the servers again this morning, and all the sessions were coming from a single IP address. Luckily. I knew the partner who owns the public IP range. Guess what? It's their NAT IP address. For now, the engineers are working on a scaling solution until we come up with a better plan.

Thank you

Darryl

***

Dear Ilene,

To pass a new audit and also for our own benefit, can you please have the AWS architects come up with an idea to track down every single change made to our AWS footprint? The more you can get granular, the better.

Thanks,

Albert

***

Hi Albert,

We're already running AWS CloudTrail. The data sometimes is too much, but we gain visibility into our users and resources activity at the API level. Let me know if you're looking for a specific part of the cloud. I also strongly recommend that you look into AWS Config. From what I understand, it helps your team track and evaluate our config compliance against a set of rules. I suspect you would want both.

Thanks,

Ilene

***

Dear Ilene,

Apparently, I cannot have a quiet weekend. Yesterday I got a call from the SecOps team. The new head of the team let me know that the AWS Flow Logs does not track anything related to the packet payloads. We were counting on it. How do we fix this? Is it a knob we're missing? For now, let's focus on VPC_EU_143.

Thanks,

Albert

***

Hi Albert,

That's a correct statement, but there is no knob in Flow Logs to enable capturing the payloads. In fact, the tool only captures the flow data. Until recently, we had to use Wireshark for that, but recently AWS rolled out traffic mirroring. It basically uses the VXLAN encapsulation to send a copy of packets from the source to the destination. From there, you can use any tool to analyze that. It works with your Suricata, and there are tools to decrypt the traffic. If you don't mind the noise of the VXLAN headers, even a simple tcpdump on your Linux box can read mirrored traffic at the destination. It won't give you all the bells and whistles to plot sessions etc., but if your goal is to capture packets and analyze the payloads, it does the job with a few caveats that I am sure you'll figure out. Please have your team contact Ali.

Thanks,

Ilene

***

Hi Ilene/Ali,

I hate the caveats for a reason. After spending hours, it turned out the low-cost T2 micro instances in our target VPC do not support traffic mirroring. I should've asked you. To my surprise, I learned that BOTH the source and destination of the traffic need to be from the supported list of instances which kind of makes sense when I consider the VXLAN encapsulation. We're falling back to tcpdump and Wireshark until we upgrade the servers. Not sure if it would be worth the cost and effort, though.

Thank you

Albert

<End of snippet>

The Architecture Review Board Meeting

In this part of the scenario, Ali, the chief cloud architect, is attending the weekly architecture board meeting. There are 8 design proposals put forth by various departments for their cloud presence that need to be reviewed by the board members and potentially approved.

Design Proposal 1

As shown in the figure below, the energy and cooling team, an important part of hardware engineering, is proposing an architecture in which they will use an interface endpoint in one of their VPCs to access 45 S3 buckets, not just from within the VPC but also over Direct Connect from their office. They have recently installed their 1G point-to-point Direct Connect.

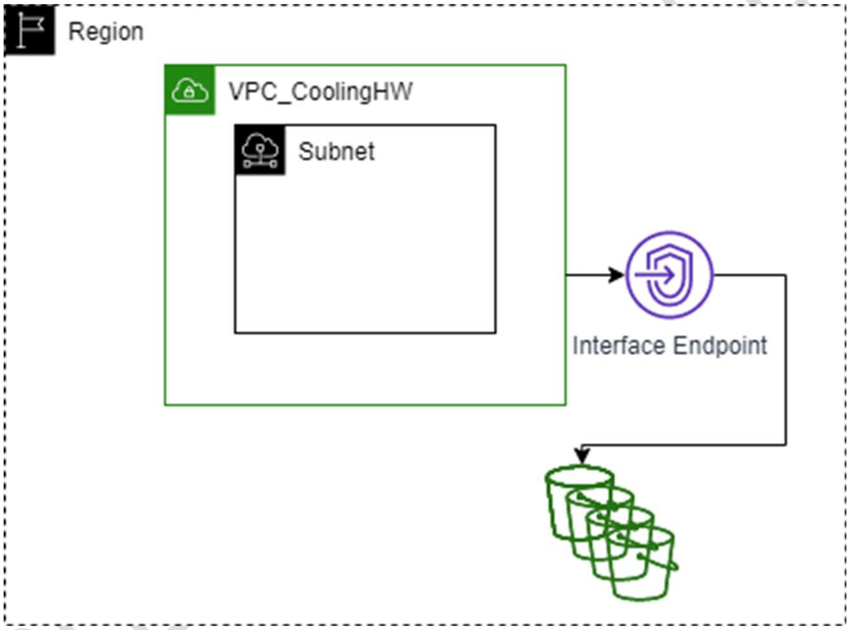


Figure 18

Ruling: Approved.

Comments: Make sure you use the right endpoint type. You need the new interface endpoints for S3 and NOT the gateway endpoint to transit. Also, study the costs carefully and apply for approval. The rates are higher than those of your gateway endpoints.

Design Proposal 2

Related to proposal #1 and the 1G DX link. The energy and cooling team has ordered a point-to-point leased line for their Direct Connect to enable MACsec on it. This will give them extra peace of mind.

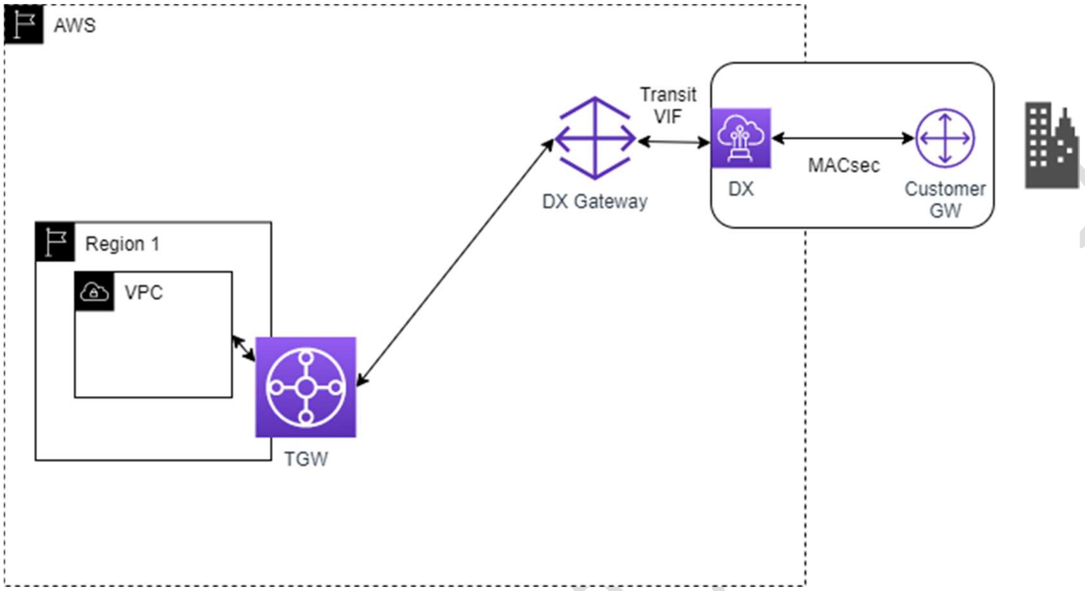


Figure 49

***Ruling*: Denied**

***Comments*: Although your point-to-point circuit might satisfy the basic requirements of MACsec, please keep in mind, AWS only provides MACsec on 10G and 100G links. This limitation is related to the type of switches/routers on the AWS side and cannot be changed by us.**

Design Proposal 3

The public apps team is requesting a change to their ELB architecture for a group of 25 custom-built Web servers to support a high-profile launch event. Currently, they're sitting behind an ALB in us-east-2. In the new architecture, they will be behind a 2-tier ELB model by adding more ALBs as shown below. The goal is to introduce a 3rd party solution of 5 servers for caching and traffic filtration between the two tiers.

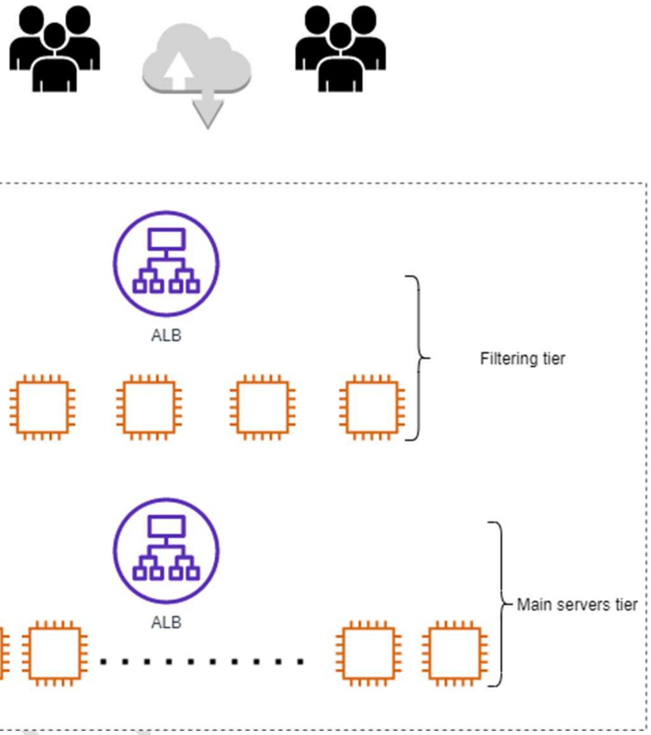


Figure 502

***Ruling*: Rescheduled.**

***Comments*: Not enough information. Does your 3rd party solution only take an IP address as the next hop address? How about FQDN? If having a static IP to point to is a hard requirement, you will have to use NLB in your second tier. Furthermore, do you need stickiness in both tiers? Please submit your clarifications to the next architecture board meeting.**

<End of snippet>

Sample mini-scenarios

52- As part of a large stack configuration on AWS using CloudFormation you also need to configure 5 homegrown business applications running on a fleet of EC2 instances. The development team has provided a set of twelve APIs that can be called using Python to configure the applications. How should you incorporate the applications in your existing templates?

A- As long as the code is written in Python you just need to create a Lambda function for the Python code and call it directly from the template

B- Create a Lambda function for the Python code. In the template use the AWS::ElasticBeanstalk::Application resource type to invoke the Lambda function.

C- Create a Lambda function for the Python code. AWS::CloudFormation::CustomResource resource type to invoke the Lambda function. Manually specify the ARN of the Lambda function under ServiceToken.

D- Create a Lambda function for the Python code. In the template use the AWS::EC2::LaunchTemplate resource type to invoke the Lambda function.

53- As a Cloud Solution Architect working for a Value Added Reseller (VAR) you are assisting a customer to reduce their AWS monthly fees and administrative overhead. The customer is a video game development company with a fleet of 50 test EC2 instances in one VPC for GPU-intensive testing. They are used by the development team and although they need an Internet connection to update the operating systems and packages, they must not be accessed by anyone on the Internet. Which of the following architectures better meets their requirements?

A- Use an Internet Gateway and public IP addresses on the EC2 instances. Use NACLs and Security Groups to secure the environment

B- Use a NAT Gateway and assign private IP addresses to the EC2 instances.

C- Use a NAT Gateway and assign private IP addresses to the EC2 instances. Use NACLs and Security Groups to secure the environment

D- Deploy a firewall to perform source NAT on an EC2 instance. Assign private addresses to the EC2 test instances.

54- You are architecting an AWS solution for a software development company. Currently, they have a physical data center connected with a dedicated 100G Direct Connect link to their AWS presence in the ap-east-1 region in Mumbai. The goal is to enable the EC2 instances in VPC-dev023 to download updates every night at 10PM. You have two options to receive the updates. They can either be downloaded from Amazon RDS in another VPC in the same region and same Availability Zone called VPC-dev101 or directly from the physical data center through the Direct Connect. Which of the following architectures can meet the customer's requirements with minimum cost?

A- Establish VPC peering between VPC-dev023 and VPC-dev101. Download the files from the servers in VPC-dev101

B- Download the files directly from the servers in the datacenter

C- Create an Interface VPC endpoint to download the files from VPC-dev101

D- Place the contents in an S3 bucket in the same region and use a NAT Gateway to download the updates

55- You, as a senior architect, are reviewing a report put together by a cloud engineer to explore different opportunities to help the company reduce its monthly bills and potentially improve availability. The company currently has a wide range of AWS and homegrown services spread across multiple

Availability Zones and North American regions. Which of the following statements is NOT accurate and should be removed from the final report to the senior leadership?

A- There are no charges for data transfer inbound from the Internet to an AWS public IP address

B- Bi-directional traffic between two AWS public IP addresses located in two different regions are charged separately

C- There are charges for data transfers from one EC2 instance to another under the same account in the same region and same Availability Zone using their public IP addresses

D- There are no charges for data transfer from one EC2 instance to another in the same VPC and under the same account but in different Availability Zones

56- The network engineering team of a biomedical company is evaluating public cloud for one of their key applications with very specific needs for low latency network performance. Which of the following statements is correct about AWS EC2 instances and their networking details?

A- To improve resiliency AWS by default attempts to spread out EC2 instances across underlying hardware when EC2 instances are created.

B- To improve performance AWS by default attempts to build EC2 an instance of one account in a VPC as close to each other as possible. This cannot be changed later.

C- To improve resiliency AWS by default never builds two EC2 instances of one account in a VPC on the same underlying hardware.

D- To improve performance AWS by default attempts to build EC2 instances of one account in a VPC as close to each other as possible. This can be changed later in favor of greater resiliency.

57- After implementing a geolocation routing policy in Route 53 for an international Gas and Oil company with AWS resources in 5 regions, you constantly receive complaints from users in Denmark. After reviewing the tickets, it turns out the general theme of the issues is around: (1) Users load pages designed for the Norwegian customers (2) Users cannot successfully resolve the resource names. As the first step, you successfully confirm that the source IP address of the impacted users is indeed registered in Denmark. You also confirm that the DNS resolver used by those corporate customers is based in Norway and is in a healthy condition. Which of the following options show TWO potential root causes of the issue? (Choose TWO)

A- The default (catch-all) record is missing from the zone

B- A firewall along the path is dropping UDP packets larger than 512 bytes

C- The users' local DNS resolver does not support EDNS0 (edns-client-subnet)

D- The users' local clients do not support EDNS0 (edns-client-subnet)

58- The AWS footprint of a global online off-road equipment store has fallen victim to a deliberate security breach. Detailed forensics done by a third-party security company revealed that numerous changes were made to their AWS resources including the Security Groups to open up backdoors to databases related to shocks and snorkel kits, although the attackers only had root access to the environment for minutes. The platforms' team assures the leadership that they have extensively hardened their environment but everyone is still worried about the AWS resources and components. How would you go about tracking such changes to the AWS resources in case of another similar breach in the future?

A- AWS Config

B- AWS CloudTrail

C- AWS Macie

D- AWS GuardDuty

59- You are troubleshooting an AWS Direct Connect traffic engineering issue. The customer's IT department has just commissioned their second 10G Direct Connect link in their local region to access their VPCs in that region. The issue now is that no matter how many times they prepend their own AS number to the AS_PATH of the prefixes they're advertising to AWS, AWS still prefers the other link. Which of the following options could be the potential root cause of the issue? (Choose TWO)

A- They are not advertising their prefixes with the same length

B- They might have set some Local Pref values on some of the routes they are receiving from AWS

C- They might have used AWS predefined BGP communities to set higher Local Pref on the prefixes they are advertising to AWS

D- They might have set a lower IGP metric on the other less preferred link

E- They might be manipulating the MED values on the routes they are advertising AWS

60- You are paged into a Direct Connect troubleshooting incident where the customer is having issues with manipulating AS_PATH to influence traffic from AWS S3 buckets into their network via the two 10G Direct Connect links that have just been put in service. The customer is running eBGP between their AS (AS#64513) and AWS and iBGP between their border routers. Their goal is to have an active/passive setup. Before opening a case with their router manufacturer which of the following options might be true about their architecture?

A- The customer must replace the current AS number with a public AS number that they own.

B- The customer must replace the current AS number with an AS number in the range of 65000-65535.

C- The customer can keep the existing AS number but with the caveat that they cannot use AS_PATH to engineer their traffic. They can instead use AWS BGP communities to set Local Pref on the AWS side into their network.

D- The customer can keep the existing AS number but with the caveat that they cannot use AS_PATH to engineer their traffic. They can instead announce more specific prefixes out of the preferred link.

61- A small business has a star architecture for their 6 VPCs, all connected to the same transit gateway. As a cloud architect, you are helping them implement a private NAT strategy for 5 of their VPCs. Those VPCs will eventually be accessing resources located in the servers' VPC. To get started, they added a NAT subnet to each of the 5 VPCs and placed a private NAT gateway in each. Which of the following options describes the most suitable routing design for this customer? Choose FOUR.

A- Ensure each VPC is attached to the TGW and associated with the main route table. Ensure propagation takes place between each VPC and the TGW.

B- Ensure each VPC is attached to the TGW and associated with a dedicated route table. Ensure propagation takes place between each VPC and the TGW.

C- Add a route pointing to the TGW to each route table of each VPC.

D- Add a route pointing to the private NAT gateway of each VPC to the other route tables of each subnet of each VPC.

E- Add a route pointing to the TGW to the route table of each NAT subnet of each VPC.

F- Add a route pointing to the TGW to the route tables of the servers VPC.

62- As a solutions architect, you are helping a customer with their migration to the cloud project. As part of the migration, they need to deploy private NAT gateways in each one of their 24 VPCs. The cloud engineering team is informed that the office of the CISO for certain EC2 instances would like to have a 1:1 NAT. The rest of the subnets in each VPC can be translated by Port Address Translation (PAT). Which of the following options is accurate about this migration?

A- The AWS private NAT gateway construct does not support 1:1 mapping. The customer would need to use a Linux box with iptables to implement such a design.

B- The AWS private NAT gateway construct does not support 1:1 mapping. The customer would need to use a public NAT gateway to implement such a design.

C- The customer can safely use the private NAT gateway construct to define the 1:1 mapping statement, but the PAT configuration cannot be done on the same gateway.

D- The customer can safely use the private NAT gateway construct to define the 1:1 mapping statement and the PAT configuration as needed.

63- As a chief cloud architect, you are contracted to review the roll-out plans prepared by the cloud engineering team of a startup. In their architecture, they created a dedicated VPC called ONPREM_ATTCH and connected that back to their data center using a single 1Gbps Direct Connect link to be backed up by a VPN connection. The idea is to use this VPC and its NAT gateway to reach the other VPCs as an intermediary platform. Also, there might be resources in ONPREM_ATTCH that would be accessed by the users in the physical data center. All the 12 VPCs owned by the customer are connected to the main route table of a single transit gateway. How would you evaluate this architecture?

A- This is a flawed architecture, and the physical data center will not be able to consume any resources in any VPC.

B- This is a flawed architecture, and the physical data center will not be able to consume any resources in any VPC, except the ones located in ONPREM_ATTCH.

C- This is a valid architecture, and the physical data center will be able to consume any resources in any VPC.

D- As it stands now, this is a valid architecture, and the physical data center will be able to consume any resources in any VPC, but if the customer fails over to the VPN, the physical data center won't have access to any resources except the ones in ONPREM_ATTCH.

64- A small insurance company with only one VPC is trying to architect an inline firewall solution to protect 25 servers that they have deployed up in a subnet called Servers with public IP addresses. They reached an agreement with their firewall vendor to bring their on-premises license to the cloud. The firewall will have two inside and outside interfaces, run on an EC2 instance and protect the servers from the inbound traffic from the Internet. The VPC is connected to the Internet via an Internet Gateway. Which of the following options shows the correct high-level architecture to support this idea?

A- Create two separate subnets for the inside and outside interfaces. Configure a gateway route table for the IGW. Add a route to direct all traffic destined for the servers to outside ENI of the firewall. Add a route to the servers' route table to direct all traffic destined for the Internet to the inside ENI of the firewall.

B- Create two separate subnets for the inside and outside interfaces. Add a route to the route table of the outside interface to direct all traffic destined for the servers to the outside ENI of the firewall. Add a route to the servers route table to direct all traffic destined for the Internet to the inside ENI of the firewall.

C- Create a dedicated firewall subnet and place the inside and outside interfaces in it. Configure a gateway route table for the IGW. Direct all traffic destined for the servers to the outside ENI of the firewall. Add a route to the servers' route table to direct all traffic destined for the Internet to the inside ENI of the firewall.

D- Create a dedicated firewall subnet and place the inside and outside interfaces in it. Configure a gateway route table for the IGW. Direct all traffic destined for the servers to the outside ENI of the firewall. Add a route to the servers' route table to direct all traffic destined for the Internet to the IGW.

65- A high school with a single VPC is looking to insert a dual-interface Next Generation firewall in their existing AWS design to protect their servers' subnet. Currently, they do not have any public addresses on the servers, and their reachability to the outside world is established through a NAT gateway, then the Internet Gateway attached to the VPC. The virtual firewall will be running on an EC2 instance. Which of the following partial architectures should the chief architect approve to maintain connectivity and improve their security posture before the full design is prepared?

A- Create a dedicated subnet for the outside interface of the firewall. Place the inside interface of the firewall in the same subnet as the NAT gateway. Add a route to the route table of the NAT subnet and direct 0.0.0.0/0 to the inside interface of the firewall.

B- Place both interfaces of the firewall in the same subnet as the NAT gateway. Add a route to the route table of the NAT subnet and direct 0.0.0.0/0 to the inside interface of the firewall.

C- Place both interfaces of the firewall in the same subnet as the NAT gateway. Add a route to the route table of the NAT subnet and direct 0.0.0.0/0 to the outside interface of the firewall.

D- Create a dedicated subnet for the outside interface of the firewall. Place the inside interface of the firewall in the same subnet as the NAT gateway. Add a route to the route table of the servers' subnet and direct 0.0.0.0/0 to the inside interface of the firewall.

66- A principal architect is being consulted by the cloud engineering team of a legal firm. They only have one VPC in AWS and need to deploy a Next Generation firewall to protect their 16 subnets from the Internet reachable via an Internet Gateway. Currently, all the servers have public IP addresses and run host-based firewalls. The company is looking to use VPC ingress routing on the IGW to eventually remove all the host-based firewalls, centralize security rules at the firewall and maintain connectivity and security. If they are assured in terms of security, their IT leadership has no issues with keeping the public IP addresses on the servers. Which of the following high-level architectures can help the team achieve its goals?

A- Install the virtual firewall on an EC2 instance with 17 elastic network interfaces; 16 on the inside and 1 on the outside. Place each inside interface in one of the server's subnets. Remove the host-based firewalls.

B- Install the virtual firewall on an EC2 instance with 17 elastic network interfaces; 16 on the inside and 1 on the outside. Place each inside interface in one of the server's subnets. Remove the host-based firewalls. Remove the public IP addresses.

C- Install the virtual firewall on an EC2 instance with 2 elastic network interfaces; 1 on the inside and 1 on the outside. Place a NAT gateway between the servers' subnet and the firewall. Create 1:1 NAT map entries for each subnet. Remove the host-based firewalls. Remove the public IP addresses.

D- Install the virtual firewall on an EC2 instance with 2 elastic network interfaces; 1 on the inside and 1 on the outside. Place a NAT gateway between the servers' subnet and the firewall. Remove the host-based firewalls. Remove the public IP addresses.

67- As a solutions architect, you are helping an AWS customer to expand their security stack in the cloud. Currently, in their only VPC, they have a single firewall that processes the traffic between the Internet and a group of 32 servers. The solution uses the concept of VPC ingress routing with a dedicated route table for the IGW to send the traffic to the Elastic Network Interface of the firewall. After experiencing some performance issues, the cloud engineering team is looking to increase the number of firewalls by adding 3 more units. Which of the following two high-level architectures indicate valid and independent options for this customer? Choose TWO.

A- In the same VPC, place all the servers in one subnet and the GWLB, its endpoint, and the group of firewalls in another.

B- In the same VPC, place all the servers, the GWLB, its endpoints, and the group of firewalls in one subnet.

C- In the same VPC, place all the servers and the GWLB endpoint in two different subnets. Create another VPC and place the GWLB and the group of firewalls.

D- In a dedicated VPC, place all the servers in as many subnets as needed. Create another VPC and place the GWLB, its endpoint, and the group of firewalls.

68- A high school with a single VPC is looking to insert a dual-interface Next Generation firewall in their existing AWS design to protect their servers' subnet. Currently, they do not have any public addresses on the servers, and their reachability to the outside world is established through a NAT gateway, then the Internet Gateway attached to the VPC. The virtual firewall will be running on an EC2 instance. Which of the following partial architectures should the chief architect approve to maintain connectivity and improve their security posture before the full design is prepared?

E- Create a dedicated subnet for the outside interface of the firewall. Place the inside interface of the firewall in the same subnet as the NAT gateway. Add a route to the route table of the NAT subnet and direct 0.0.0.0/0 to the inside interface of the firewall.

F- Place both interfaces of the firewall in the same subnet as the NAT gateway. Add a route to the route table of the NAT subnet and direct 0.0.0.0/0 to the inside interface of the firewall.

G- Place both interfaces of the firewall in the same subnet as the NAT gateway. Add a route to the route table of the NAT subnet and direct 0.0.0.0/0 to the outside interface of the firewall.

H- Create a dedicated subnet for the outside interface of the firewall. Place the inside interface of the firewall in the same subnet as the NAT gateway. Add a route to the route table of the servers' subnet and direct 0.0.0.0/0 to the inside interface of the firewall.

69- A principal architect is being consulted by the cloud engineering team of a legal firm. They only have one VPC in AWS and need to deploy a Next Generation firewall to protect their 16 subnets from the Internet reachable via an Internet Gateway. Currently, all the servers have public IP addresses and run host-based firewalls. The company is looking to use VPC ingress routing on the IGW to eventually remove all the host-based firewalls, centralize security rules at the firewall and maintain connectivity and security. If they are assured in terms of security, their IT leadership has no issues with keeping the public IP addresses on the servers. Which of the following high-level architectures can help the team achieve its goals?

E- Install the virtual firewall on an EC2 instance with 17 elastic network interfaces; 16 on the inside and 1 on the outside. Place each inside interface in one of the server's subnets. Remove the host-based firewalls.

F- Install the virtual firewall on an EC2 instance with 17 elastic network interfaces; 16 on the inside and 1 on the outside. Place each inside interface in one of the server's subnets. Remove the host-based firewalls. Remove the public IP addresses.

G- Install the virtual firewall on an EC2 instance with 2 elastic network interfaces; 1 on the inside and 1 on the outside. Place a NAT gateway between the servers' subnet and the firewall. Create 1:1

NAT map entries for each subnet. Remove the host-based firewalls. Remove the public IP addresses.

H- Install the virtual firewall on an EC2 instance with 2 elastic network interfaces; 1 on the inside and 1 on the outside. Place a NAT gateway between the servers' subnet and the firewall. Remove the host-based firewalls. Remove the public IP addresses.

70- As a solutions architect, you are helping an AWS customer to expand their security stack in the cloud. Currently, in their only VPC, they have a single firewall that processes the traffic between the Internet and a group of 32 servers. The solution uses the concept of VPC ingress routing with a dedicated route table for the IGW to send the traffic to the Elastic Network Interface of the firewall. After experiencing some performance issues, the cloud engineering team is looking to increase the number of firewalls by adding 3 more units. Which of the following two high-level architectures indicate valid and independent options for this customer? Choose TWO.

E- In the same VPC, place all the servers in one subnet and the GWLB, its endpoint, and the group of firewalls in another.

F- In the same VPC, place all the servers, the GWLB, its endpoints, and the group of firewalls in one subnet.

G- In the same VPC, place all the servers and the GWLB endpoint in two different subnets. Create another VPC and place the GWLB and the group of firewalls.

H- In a dedicated VPC, place all the servers in as many subnets as needed. Create another VPC and place the GWLB, its endpoint, and the group of firewalls.

71- You are completing a Gateway Load Balancer architecture for an online education company. The company inherited this half-done design after their previous architect abruptly left the firm. The goal is to deploy a group of 4 firewalls in-path of every packet that travels between their AWS environment and the Internet. In his architecture, before his departure, the previous architect provisioned 5 VPCs for the main departments with their own Internet Gateways. The design also had a dedicated VPC for the firewalls. Finally, the documents indicate he had plans to place one GWLB endpoint in each of the 5 VPCs. Choose the right set of routes for this architecture. Choose THREE.

A- 0.0.0.0/0 to gwlbe-xyz for the servers subnet in each VPC

B- 0.0.0.0/0 to igw-xyz for the servers subnet in each VPC

C- 0.0.0.0/0 to igw-xyz for the endpoint subnets in each VPC

D- 0.0.0.0/0 to gwlbe-xyz for the endpoint subnets in each VPC

E- x.y.z.t/p (servers subnet) to gwlbe-xyz for the Internet Gateways of each VPC

F- 0.0.0.0/0 to gwlbe-xyz for the GWLB subnet in its dedicated VPC

72- As a chief architect, you are helping a team of cloud engineers with their security enhancement project. They currently have 10 VPCs, one for each team. They are all connected to the same transit gateway in the same region. Each VPC has a dedicated Internet Gateway, an idea despised by the new CISO. Their goal is to improve their security posture by centralizing the security stack in a separate VPC with a NAT gateway. They are also making a significant investment to purchase 4 Next Generation firewalls. Which of the following statements show a valid component of the central security VPC design? Choose FOUR.

A- Remove the IGWs from all the 10 VPCs. Create a new VPC for the GWLB, its endpoint, the IGW, and NAT gateway.

B- Remove the IGWs from all the 10 VPCs and place a GWLB endpoint in each one of the 10 VPCs. Create a new VPC for the GWLB, the IGW, and NAT gateway.

C- Attach the new VPC to the TGW and propagate the routes of each one of the 10 VPCs to its route table.

D- Add a route on the TGW associated with the VPCs to direct 0.0.0.0/0 to the security VPC.

E- Add a route to the route table of each one of the VPCs and direct 0.0.0.0/0 to the GWLB endpoint (gwlbe-xyz).

F- The subnet of the GWLB and its endpoint will have a route table containing two routes o direct 0.0.0.0/0 to the NAT gateway and the servers' supernet to the TGW.

73- You are architecting a security solution in AWS for a public company. Due to regulatory restrictions, the customer cannot place two groups of servers in the same subnet. In their original design, in the same VPC, those servers were simply placed in different subnets, and access control between them was implemented using network security lists. After three years, now the solution is difficult to manage and is not scalable at all. Your team's suggestion is to deploy a pair of stateful firewalls using Gateway Load Balancers between the two subnets. How would you rate this high-level architecture?

A- They would need to create two GWLB endpoints in two separate subnets. Each server's subnet would need a more specific route for the other servers' subnet via the endpoint.

B- They would need to create one GWLB endpoint in a separate subnet. Each server's subnet would need a more specific route for the other servers' subnet via the endpoint.

C- They would need to create a GWLB endpoint in each one of the server's subnets. Each server's subnet would need a more specific route for the other servers' subnet via the endpoint.

D- This architecture as it stands now is not supported.

74- After a major cloud redesign where you deployed a stack of 4 Next Generation firewalls behind a Gateway Load Balancer to inspect the Internet-bound traffic, the SecOps team is complaining about losing visibility into the IP addresses of their EC2-based servers when they communicate with the Internet. Your initial assessment indicates that the Next Generation firewalls only log the IP address of the NAT gateway and have no trace of the addresses of the EC2-based servers. In the current architecture, you have an Internet Gateway with ingress routing to direct any traffic destined for the servers to the Gateway endpoint. Also, from the servers, a default route sends all the traffic to the NAT gateway and then, from there, to the GWLB endpoint. Which of the following architectural changes can help the SecOps team regain visibility into the servers' IP addresses in the firewalls' logs without compromising any security enhancements?

A- Currently, the NAT gateway is between the GWLB endpoint and the servers. Move and place the NAT gateway between the IGW and GWLB endpoint.

B- Remove the NAT gateway from the design.

C- Replace the AWS NAT gateway with a Linux-based NAT gateway with Iptables.

D- Create a dedicated VPC for the security stack, including the GWLB endpoint, GWLB, firewalls, and NAT gateway.

75- You have architected a dual-Availability Zone security solution for a regional bank. In this design, you used 2 AZs in the us-east-1 region. Each of your 15 department VPCs has subnets in both AZs. You also have a dedicated security VPC with two subnets, one in each AZ. The security VPC has two sets of Gateway Load Balancers, endpoints, and stateful firewalls in AZ1 and AZ2. All the VPCs are connected to a Transit Gateway. The goal of your architecture is to inspect any east-west traffic between the servers in the VPCs. You are asked by a junior engineer, "What would happen if we disabled the appliance mode on the TGW?"

Which of the following options accurately answers this question?

A- Since the firewalls are stateful, the setting can safely be turned off.

B- With the appliance mode turned off, we would not be able to failover immediately if we were to lose one AZ.

C- With the appliance mode turned off, one of the GWLBs would never be used. As a result, we would end up with an active-passive architecture.

D- With the appliance mode turned off, we would have connectivity issues if two servers in the different AZs of different VPCs were to talk to each other.

76- In order to inspect any east-west traffic between your group of 25 VPCs spanned across Availability Zone 1 and 2, you deployed a dedicated VPC. The security VPC has two subnets located in AZ1 and AZ2, each with dedicated a Gateway Load Balancer, its endpoint, and a stack of stateful firewalls. All the VPCs are connected via a single transit gateway. Which one of the following conversations will have issues if the implementation engineer forgets to enable appliance mode for the security VPC attachment?

A- Between server 1 in AZ1 VPC1 and server 2 in AZ1 VPC1

B- Between server 1 in AZ1 VPC1 and server 2 in AZ2 VPC1

C- Between server 1 in AZ2 VPC1 and server 2 in AZ2 VPC2

D- Between server 1 in AZ2 VPC1 and server 2 in AZ1 VPC2

77- In a large Gateway Load Balancer deployment, you have a dedicated VPC for the security appliances and 65 spoke VPCs. In the security VPC, the GWLB has registered targets (appliances) in both AZ1 and AZ2 of the us-west-1 region. You have enabled cross-zone load balancing. Which of the following statements is correct about the behavior of the GWLB in this architecture when all the appliances in both AZs have failed?

A- The GWLB will drop all the traffic until at least one appliance recovers in one of the AZs.

B- The GWLB will drop all the traffic until at least two appliances recover in two different AZs.

C- The GWLB will keep sending new sessions to the last healthy target.

D- The GWLB selects a random target and sends the traffic to it.

78- You are configuring routing for a small AWS Network Firewall implementation. The customer only has one VPC with a dedicated Internet Gateway, and by creating a new subnet (10.1.254.0/24) for the firewall endpoint (vpce-123), you need to inspect the traffic between 2 servers subnets (10.1.1.0/24 and 10.1.2.0/24) and the Internet. Which of the following statements is accurate about the proper routing strategy? Choose THREE.

A- Add to the gateway route table of the IGW: route 0.0.0.0/0 to vpce-123

B- Add to the gateway route table of the IGW: route 10.1.1.0/24 to vpce-123 and 10.1.2.0/24 to vpce-123

C- Add to the route table of the endpoint's subnet: route 0.0.0.0/0 to igw-xyz

D- Add to the route table of the endpoint's subnet: route 10.1.1.0/24 to vpce-123 and 10.1.2.0/24 to vpce-123

E- Add to the route table of the servers subnets: route 0.0.0.0/0 to vpce-123

79- As an external consultant, you are evaluating a distributed AWS Network firewall deployment at an online vacation booking company. Currently, they have 12 VPCs in one region, and in a dedicated subnet in each VPC, they have placed an AWS Firewall endpoint. The ultimate goal of the architecture is to inspect any north-south traffic; between the servers in each VPC and the Internet Gateway. Which of the following statements is NOT accurate about this high-level design?

A- Any typo or other misconfigurations will remain contained within the same VPC and will not affect the other VPCs.

B- The customer does not have to deploy transit gateways for connectivity between these VPCs and the IGW.

C- The firewalls will have a common set of rules across all the VPCs which can easily be managed centrally.

D- The engineering team would need to add one route for each one of the server's subnets to the gateway route table of the corresponding IGW of each VPC.

80- You are architecting an AWS Network Firewall solution for a hardware manufacturer. Currently, they have 12 VPCs all connected to the same transit gateway in the us-west-2 region. They also have two 100G Direct Connect links with MACsec attached to the same region via a Direct Connect gateway. The AWS footprint has dedicated connectivity to the public Internet. Your goal is to inspect the traffic between their VPCs and on-premises network. Due to the upcoming release date of their flagship chip, the company is unwilling to take any risks by adding any new subnets to the production VPCs, nor are they willing to update the route tables of any of the production VPCs. Which of the following options provides a valid high-level design?

A- Place a firewall endpoint in one of the existing subnets of each VPC.

B- Create a dedicated VPC, attach it to the transit gateway and add a single route to each VPC to direct the traffic to the new VPC.

C- Create a dedicated VPC with two subnets, one for the TGW attachment and the other one for the firewall endpoint.

D- Create a dedicated VPC with one subnet for the TGW attachment and the firewall endpoint.

<End of Snippet>

Sample Comprehensive Answers

67- (A) This architecture is a typical scenario where we deploy a simple firewall with two interfaces, namely inside and outside. In such scenarios, for the first section, you create a dedicated route table (usually called a gateway route table) for the IGW. This route table will have a route to direct all the inbound traffic from the Internet to the outside interface of the firewall. The second section of this architecture is where you add a route to the route table of the servers' subnet to direct all the Internet-bound traffic to the inside interface of the firewall. This scenario does not need a NAT construct as all the servers have public IP addresses.

68- (A) This architecture is a classic scenario where the customer is not willing to assign public IP addresses to their backend servers, is using a NAT construct for Internet connectivity, but they still need to insert a firewall inline. To accommodate such scenarios, you can logically place the firewall between the IGW and NAT gateway. The outside interface of the firewall with a public address will face the IGW. The inside interface of the firewall with a private address will face the NAT gateway and share the same subnet with it. To route packets in this scenario, you need to add a route to the servers' route table for 0.0.0.0/0, pointing to the NAT gateway. Then you need to add a route to the NAT subnet for 0.0.0.0/0 pointing to the inside interface of the firewall. The subnet containing the outside interface of the firewall will have a route for 0.0.0.0/0 pointing to the IGW to complete the chain. The inbound traffic is handled by adding a route for the NAT gateway to the gateway route table of the IGW. Finally, please note the question is looking for a partial answer which is not uncommon when multiple teams are involved in a design process of a multi-phase effort. All the other choices contain wrong options. From a design perspective, this solution also helps the customer keep their AWS elastic NAT solution without worrying about scaling up the firewall to support the process-intensive NAT.

69- (D)The key to choosing the right architecture here is to pay attention to two facts. Firstly, the AWS NAT gateway does not support 1:1 map; in fact, for the most part, it is a simple PAT (Port Address Translation) device. Secondly, AWS currently supports a maximum of 15 interfaces on one instance. Although this is the maximum for EC2, it does not necessarily mean you can run your firewall on that instance, or even that it financially makes sense. Hence, although you could have one interface per subnet in theory, this solution would not be scalable or even, in this case, doable. As indicated in D, the most reasonable option to achieve all the goals here would be deploying a NAT gateway and removing the public IP addresses from the servers. With that solution in place, you could grow the number of subnets to a much larger number until you start running into issues with the AWS NAT gateway's limitations, such as its total bandwidth (~45Gbps), the maximum number of existing sessions (55000), or the number of new sessions per minute (900). Once that point is reached, you could rearchitect your solution, for example, by deploying more firewalls. On a final note, you may exhaust your firewall resources before those of your NAT gateways. That is another reason why capacity planning and monitoring in such cases are critical. So is the new type of load balancers, the Gateway Load Balancer (GWLB), and its specific architectures.

70- (A)(C) The two options here present two major architecture options available to those customers with simple designs and north-south traffic flows as described here. They usually have a single VPC and would like to accommodate everything in there. Or they have a small number of VPCs and do not mind adding an extra VPC to centralize the GWLB function. The first architecture, as depicted by option A, is to create a dedicated subnet for the GWLB stack, which includes the GWLB itself, the group of firewalls, and the endpoint. The servers' subnet will have a route to direct everything north-bound to the GWLB endpoint. As described in C, the second architecture brings some centralization to the overall picture by allocating a dedicated VPC to the GWLB but leaving the endpoint in the servers' VPC. As you know by this point, the endpoint and the servers are placed in different subnets. The latter gives the customer more flexibility to expand their design in the future by simply adding more VPCs with endpoint subnets instead of dropping another GWLB and a group of firewalls in each future VPC. As you can imagine, the extra flexibility comes at the cost of more overhead to manage two separate VPCs.

71- (A)(C)(E) The logic behind this architecture is to have a dedicated VPC for the GWLB and its group of targets (the firewalls). The endpoints are placed in each one of the servers' VPCs. For such architecture to work, the servers would need to route their north-bound (Internet-bound) traffic to the GWLB endpoint of their VPCs (option E). The GWLB endpoints need to communicate with the IGW as well as the GWLB. You don't need to configure anything for the latter to take place; it's done by design. In fact, the return traffic from the GWLB to the endpoints also flows in the right direction by design. However, in order for the endpoints to talk to the IGW, you would need a route similar to option C. Finally, to properly direct the ingress traffic, you would need to use the VPC ingress routing feature and add a route for the servers subnets with the next-hop of the endpoint to the dedicated route table of each IGW. It goes without saying in large-scale scenarios, this solution can be automated or modified to a more centralized design.

72- (A)(C)(D)(F) The key challenge in this architecture is to enable communication in two major areas. First, we need to set up routing inside the security VPC, and then we need to make sure the security VPC and all the other VPCs can seamlessly communicate with each other. The latter is pretty straightforward. All you need to do is make sure route the two route tables on the TGW have routes to both sides of the equation (the security VPC on one side and the servers VPC on the other). You also need to configure each server's VPC with a route to send its traffic out to the TGW. Next, we need to turn attention to the security VPC. In this VPC, you will have 3 route tables in a line. First, from the TGW, the packet hits the TGW attachment subnet and its route table. You would need a route to direct 0.0.0.0/0 to the GWLB endpoint. The endpoint is placed in the next subnet along with its GWLB. This subnet arguably forms the heart of our design, where you would have two main routes, one directing the egress traffic to the NAT gateway and the other directing the ingress traffic to the TGW. The NAT gateway's subnet will have a default route to the IGW and a route for the servers pointing to the GWLB endpoint.

73- (D) With the exception of gateway route tables defined on Internet Gateways and Virtual Private Gateways for VPC ingress routing, in AWS VPCs, you cannot have any route more specific than the VPC's main CIDR block to direct traffic from one corner of the VPC to another. In this case, an east-west traffic inspection within the same VPC, in theory, any intercepted traffic would have a more specific destination as it belongs to a subnet of the main CIDR block. You cannot have those more specific routes in AWS VPCs. This leads to very interesting design challenges that you will see at least one of in the following questions. In scenarios like this, depending on the customer's requirements and restrictions, your best bet might be to use separate VPCs (major architectural change) or simply automate host-based solutions such as network security groups of firewalls (usually significant overhead).

74- (D) Here, you can easily rule out the two options of B and C. However, the choice between A and D needs to be made carefully. The issue as described in the scenario is that the NAT gateway is sitting between the servers and the GWLB endpoint. So, when the traffic leaves the servers, it hits the NAT gateway first and is PAT'ed to a single IP address. This is the IP address that the SecOps team captures in their firewalls' logs. An absolutely reckless architecture here would be to move the NAT gateway from where it is to a place between the GWLB endpoint and the IGW. Although this idea works in theory, as mentioned in previous questions, AWS will not allow you to create any routes more specific than the main CIDR of the VPC within a VPC. Why is this a requirement if we move the NAT gateway? If you place the NAT gateway between the IGW and the GWLB endpoint (which would make it similar to traditional data centers), you would need to configure a route for the NAT gateway's subnet to direct server-specific traffic to the endpoint. That route will be more specific than the VPC's main CIDR block and cannot be created. Hence, as this architecture demonstrates, you would need to keep the NAT gateway between the servers and the GWLB endpoint. This solution will address the AWS routing restriction but will expose you to other issues, including the one shown here. Your firewall logs will only record the outside interface of the NAT gateway. The AWS NAT gateway is effectively a PAT device so, although it would not scale, you still cannot create 1:1 NAT entries either. There are ways to get around this limitation, including creating a dedicated VPC for the security stack, as shown by option D.

75- (D) This question is not limited to the concept of GWLBs. In fact, many engineers struggle with the whole concept of TGW Appliance Mode. The underlying reason why AWS needed this new mode is simple; the AWS TGW by default tends to keep the traffic within one AZ (i.e., the originating AZ). What this sentence means in our context is that if a server in AZ1 tries to communicate with a server in AZ2, it would leave the server, hit the TGW, and intentionally be kept in AZ1. This effectively means the packets will subsequently hit the GWLB and its firewalls in AZ1. The stateful firewall in AZ1 will register and track this session. So far, we have no issues, and the packet is sent to its final destination server in AZ2. The issue starts coming to the surface when the destination server attempts to respond back. The response packets leave the destination server in AZ2, hit the TGW, and, according to the default behavior of the TGW, remain in AZ2 all the way to the GWLB and its firewalls located in AZ2. But, as you recall, the initial flow landed on the firewall in AZ1, and the unit in AZ2 has no clue about the response. Just imagine that bewildered face "response to what?" And the firewall in AZ2, not having a clue, will drop the packet. This conversation would go well if both servers were located in either AZ1 or AZ2. Because the TGW would keep the traffic in that AZ. But as you noticed, inter-AZ sessions will run into issues with stateful firewalls. This case and other similar scenarios constitute the small percentage of cases where you have to change the TGW's default behavior by enabling the appliance mode. After enabling the mode for the TGW attachment of the security VPC, the TGW will ensure both request and response traffic to the same appliance (or GWLB endpoint in this case).

76- (D) By default and without appliance mode for the security VPC's attachment, the transit gateway will keep the conversation within the originating AZ. This works well until two servers in two different AZs try to talk to each other. The request will go to AZ1 and leave a trace on its firewall, but the response will be sent from AZ2 to an AZ2 firewall with no background of the session and will get dropped.

77- (D) If cross-zone is enabled (or even if it's not) and there are no healthy targets, the GWLB will pick a random target based on the 5-tuples for TCP/UDP sessions and 3-tuples for non-TCP/UDP sessions. This behavior will continue until the session times out. Needless to say, the users will be having issues as the target appliances are not functioning properly.

78- (B)(C)(E) The goal here is to make sure any traffic to and from the servers' subnet is sent to the endpoint. Hence, we add two routes for the two servers' subnets to the gateway route table of the IGW. They both direct the inbound traffic to the vpce-123 (the firewall endpoint). By the same token, we add a default route (0.0.0.0/0) with a next-hop of the endpoint to the route tables of the servers' subnets. Finally, the endpoint's subnet will have a default route for its Internet-bound (north-bound) traffic but uses the "local" route of 10.0.0.0/16 for its south-bound communications down to the servers. Probably the trickiest part of this scenario is where you rule out option D as the process will be done by the local route. As a bit of food for thought, you might also want to figure out why it cannot be done? Remember, this point was previously covered in the book.

79- (C) While all the other options make sense, option C is not correct. In fact, each VPC can have its own set of unique firewall policies without having to do any coordination with other VPCs. Also, option D is absolutely correct, as you would need to take the VPC ingress routing approach by creating specific routes for each of the serves subnets in the VPC. Please note, from a design perspective, a distributed firewall model, like the one presented here, can be implemented without any transit gateways. Given the billing model of transit gateways, to some customers, this is one way to potentially reduce costs, although one needs to carefully consider the endpoint hourly charges as the number of them increases in a distributed deployment model, like the one presented here.

80- (C) This scenario is the prime example of a simple centralized deployment. In this case, the customer does not need to touch their "spoke" (also known as satellite) VPCs at all. However, you would need to leverage a transit gateway to connect the VPCs and Direct Connect (and potentially the site-to-site VPN tunnels). As you have seen before, the endpoint will sit in its own dedicated subnet for the routing logic to work. Please note, option A and B clearly fail to meet the requirements of this customer. We will take a closer look at the centralized deployment model and its routing details in the following scenarios.

<End of Snippet>

Reference Architecture(s)

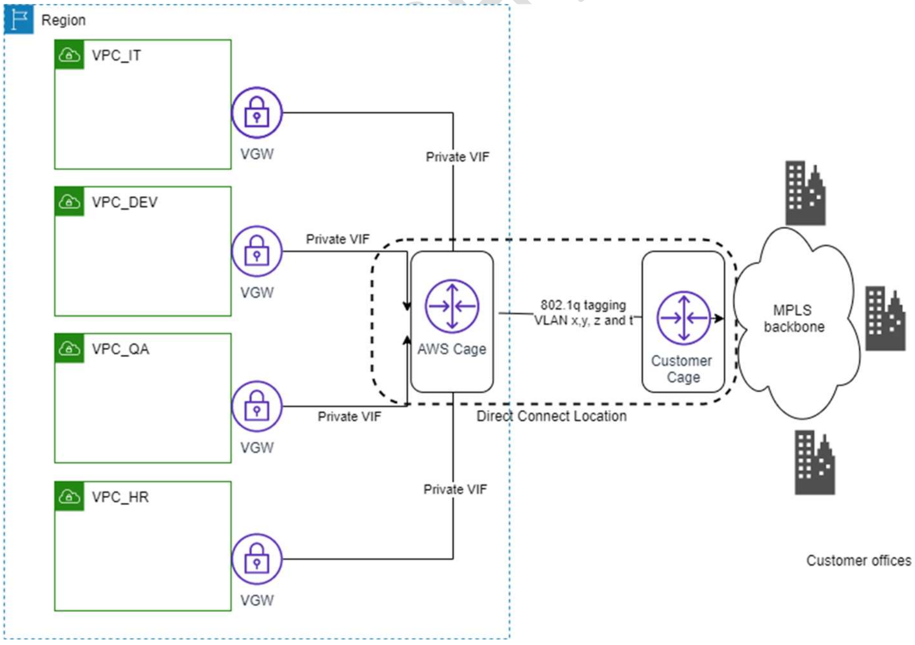


Figure 3

The most basic form of connectivity between a group of VPCs in the same region and the on-premises network via Direct Connect. Notably, the customer is using VGWs instead of transit gateways. They only have presence in one region (hence no need to DX gateway). As always, AWS uses VLAN tagging (802.1q) sub-interfaces on the routers to distinguish between the traffic of each VPC. It's not shown here, but if the Direct Connect link is point-to-point 10Gbps or 100Gbps, you might also be able to encrypt it using MACsec. At the time of this writing, there is no native way to

use IPSec for non-P2P circuits, although, it might come out at some point and as you will see, there are custom architectures that we will discuss extensively. One way or the other, the million-dollar question remains; do we need encryption on a private dedicated fiber between your device and AWS? Some security experts and auditors would require that. That is a battle, rarely won by the cloud and network architects.

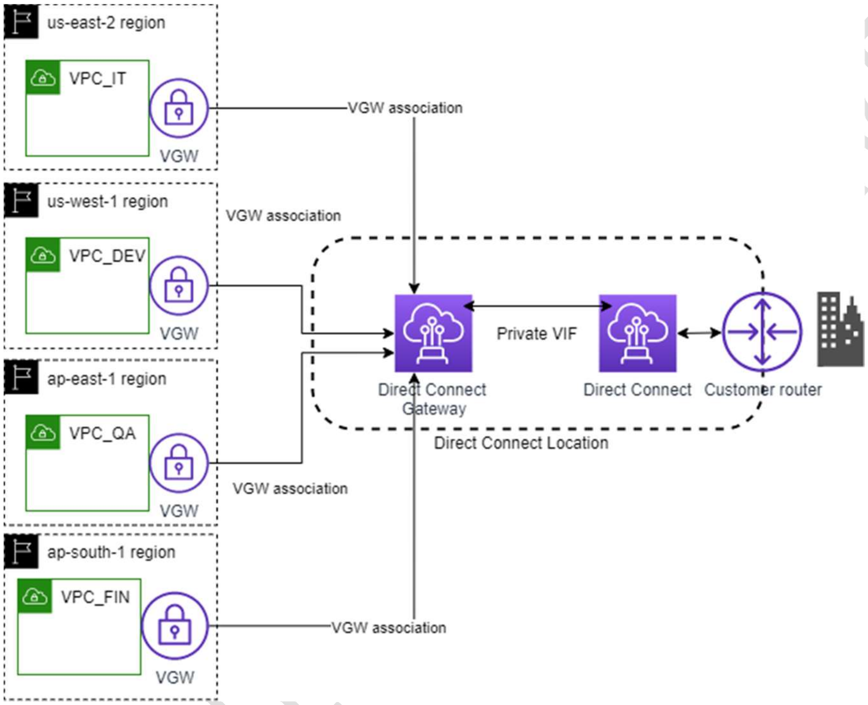


Figure 4

Simple multi-region reference architecture for Direct Connect. Here, although the architect could have used the DX link to reach VPCs in one region, she is using another construct called DX gateway to provide global reach to multiple regions. This design is based on VGWs, but as you will see in the following sections, this model is also extensively used when we have transit gateways (TGWs).

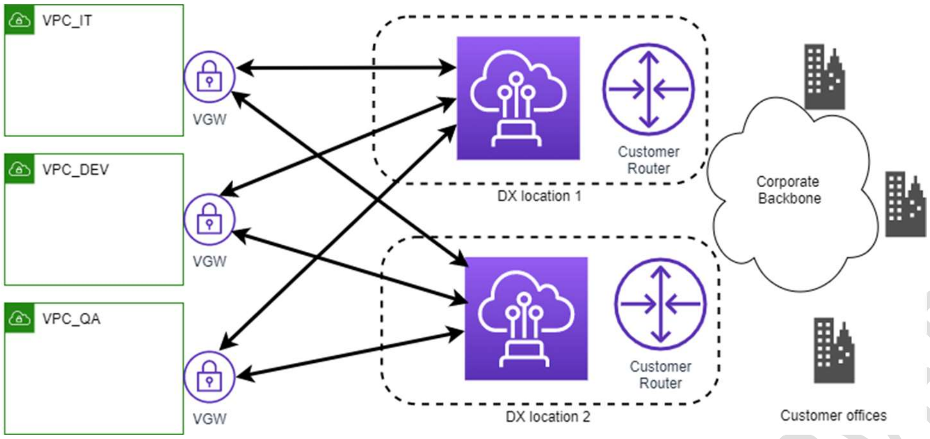


Figure 5

Further enhancements to the previous models by going to multiple Direct Connect locations and using dedicated routers on the customer side. Predictably, the cost would be higher but better numbers in terms of availability can be achieved. The main cost components would be those related to the hourly fees of each one of the four Direct Connect links, the cost of customer equipment, and if any redundancy is needed on the network between the office and those locations, the cost of introducing that also need to be taken into account. Customers with strict availability requirements might require diverse paths from different carriers to get to the AWS locations (if not co-located) and diverse paths within their own backbone (if co-located but have offices in other places).

Detailed cost analysis and budget justification are always among the architect's responsibilities; without them, anyone could design the most available networks in the world with many "9s".

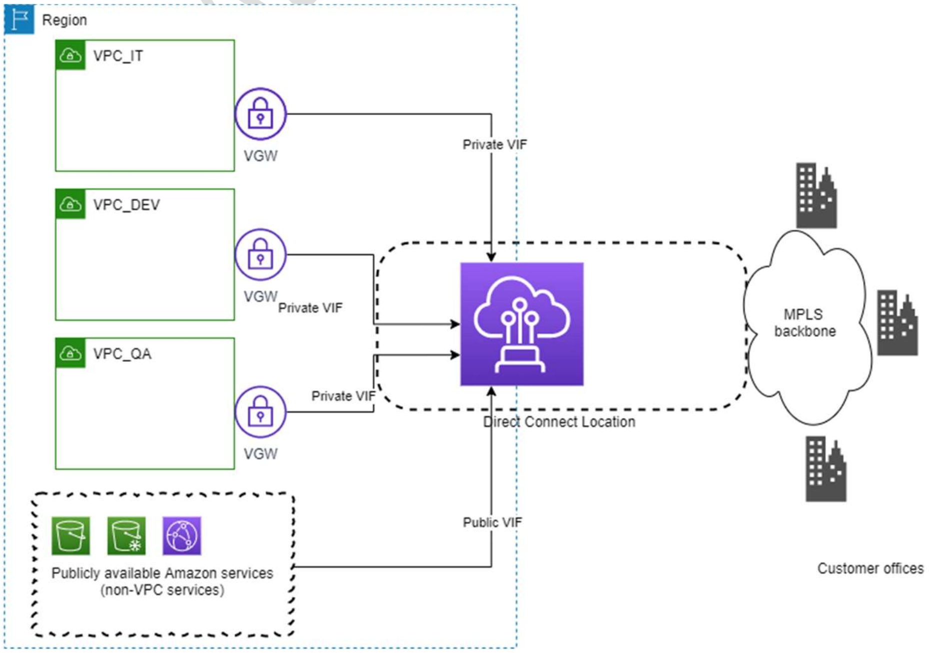
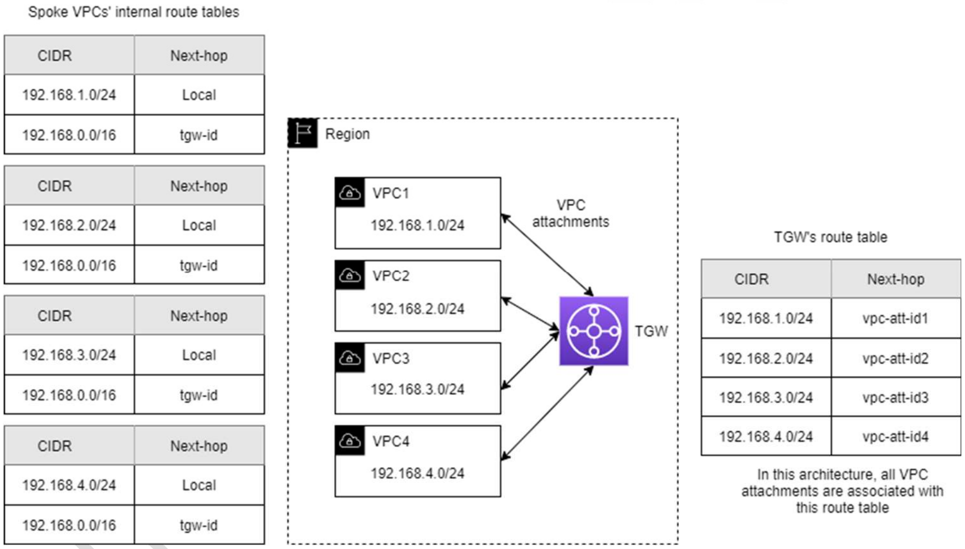


Figure 6

A common reference architecture used by Direct Connect customers with requirements to connect to both a group of VPCs and AWS public services. Notably, the VLAN tagging process would allow you to carve out a dedicated channel for the public VIF and treat it similar to another VPC. The architect needs to keep in mind that a public VIF could potentially give them access to public resources in more than one region, while the private VIFs would be dedicated to one region if no DX gateway is used. At this point, you need to ask yourself; what would happen to (1) access to the VPCs (2) access to S3/Glacier, if we were to lose the DX location? With or without the backup VPN, their fate would not necessarily be the same.

Reference Architecture(s)



Spoke VPCs' internal route tables

CIDR	Next-hop
192.168.1.0/24	Local
192.168.0.0/16	tgw-id

CIDR	Next-hop
192.168.2.0/24	Local
192.168.0.0/16	tgw-id

CIDR	Next-hop
192.168.3.0/24	Local
192.168.0.0/16	tgw-id

CIDR	Next-hop
192.168.4.0/24	Local
192.168.0.0/16	tgw-id

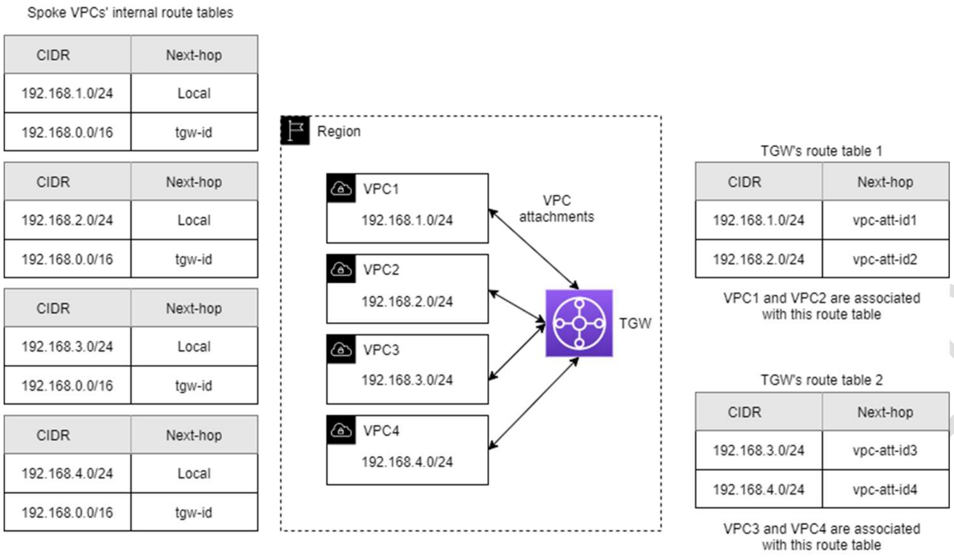
TGW's route table

CIDR	Next-hop
192.168.1.0/24	vpc-att-id1
192.168.2.0/24	vpc-att-id2
192.168.3.0/24	vpc-att-id3
192.168.4.0/24	vpc-att-id4

In this architecture, all VPC attachments are associated with this route table

Figure 7

The most basic reference architecture to deploy a TGW on AWS. This figure shows a totally flat network where all the VPCs can use the "hub" to communicate with each other. Currently, there is only one route table on the TGW that each VPC is associated to. This route table is used to find the destination attachment (VPC) if the traffic does not match the local entry of each spoke route table.



Spoke VPCs' internal route tables

CIDR	Next-hop
192.168.1.0/24	Local
192.168.0.0/16	tgw-id

CIDR	Next-hop
192.168.2.0/24	Local
192.168.0.0/16	tgw-id

CIDR	Next-hop
192.168.3.0/24	Local
192.168.0.0/16	tgw-id

CIDR	Next-hop
192.168.4.0/24	Local
192.168.0.0/16	tgw-id

TGW's route table 1

CIDR	Next-hop
192.168.1.0/24	vpc-att-id1
192.168.2.0/24	vpc-att-id2

VPC1 and VPC2 are associated with this route table

TGW's route table 2

CIDR	Next-hop
192.168.3.0/24	vpc-att-id3
192.168.4.0/24	vpc-att-id4

VPC3 and VPC4 are associated with this route table

Figure 8

A common reference architecture for simple network segmentation on AWS using TGWs. Each group (also known as segment) of VPCs is attached to the same route table of the same TGW and routes whatever is not local (meaning in the same VPC) to its "hub router," known as the TGW. In this case, VPC 1 and 2 are in the same group and VPC 3 and 4 are in another segment. Each segment uses the route in the corresponding route table of the TGW to find the other VPC members of the same group.

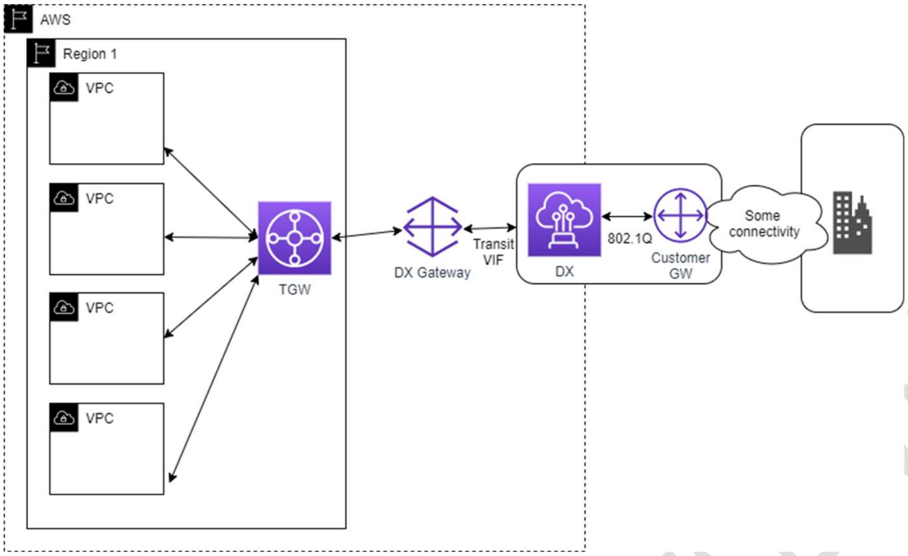


Figure 9

The architecture shows a simple single-region network with Direct Connect and TGW. In contrast to the traditional VGW-based architectures, here, we use the DX gateway between the DX and TGW. The VIF between the two is of a special type called transit VIF. As you will see in the following designs, this architecture can be easily expanded to multiple regions.

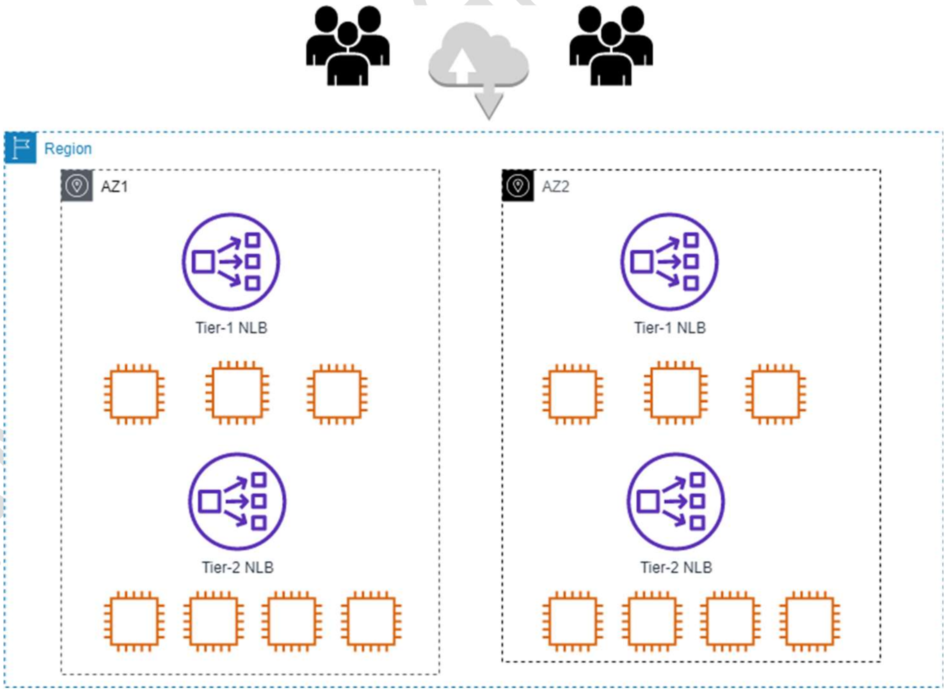


Figure 10

Reference architecture for a simple high-throughput design. While there will be no Layer 7 processing, but the architect can handle millions of sessions. the architect needs to pay close attention

to the capacity planning process for the servers between the two tiers of ELBs. They can easily turn into a bottleneck.

Although it's beyond the scope of this book, but for such servers you need to take some extra steps. This topic is extensively discussed in Linux related training courses where we show how to optimize high-throughput cache servers.

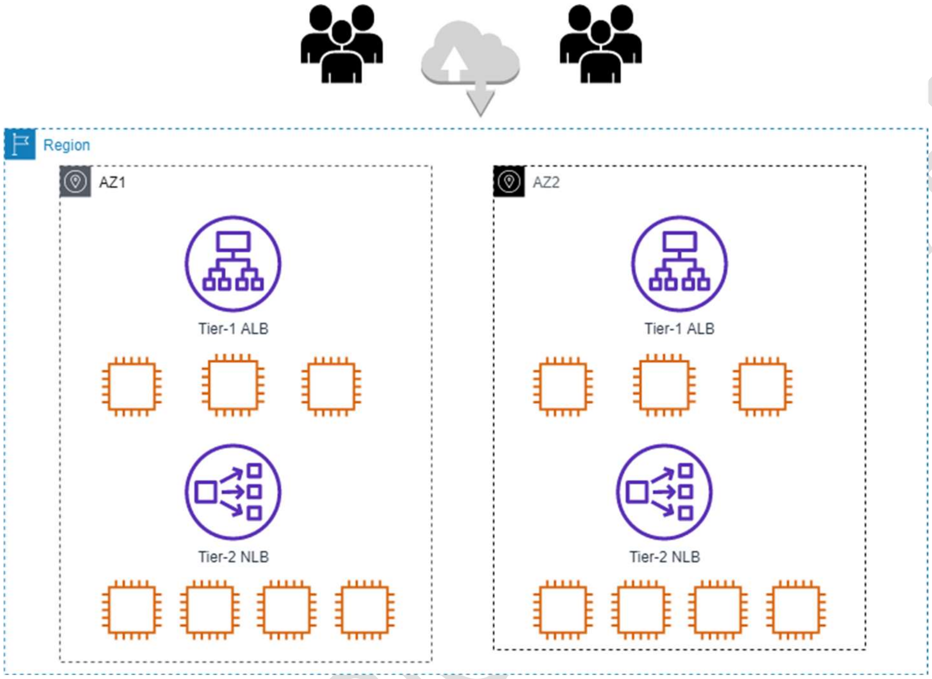


Figure 11

This reference architecture shows a dual-AZ sandwich model for load balancing. Here, the architect is inserting some logic such as caching or firewalling between the two tiers. Once the processing by the caching engines or firewalls is complete, the traffic is sent to the second-tier load balancers. Two key points stand out in this architecture: (1) deployment of an NLB at the second layer because currently, it's the only type of ELB that can be addressed by the cache servers by its static IP address, if needed (2) if stickiness is required in both tiers, the architect cannot deploy two lines of ALBs or two lines of CLBs. It's strongly recommended to deploy a mix where the cookie names would be different. This topic is covered in the mini-scenarios.

The choice of ALBs is usually based on the customer's potential layer 7 requirements. Do they need to inspect the HTTP header to make any decisions? Do we need separate routing strategies for different URLs? Etc.

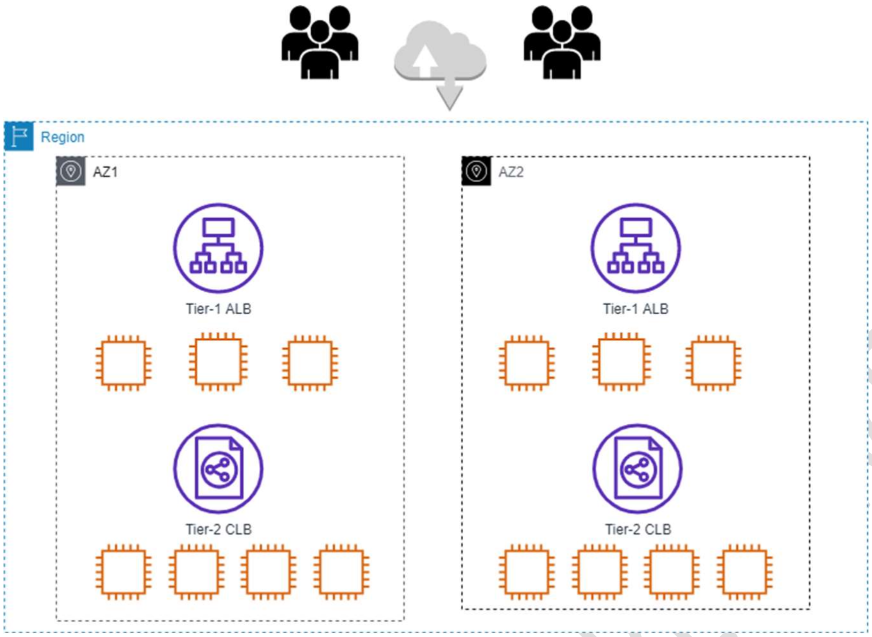
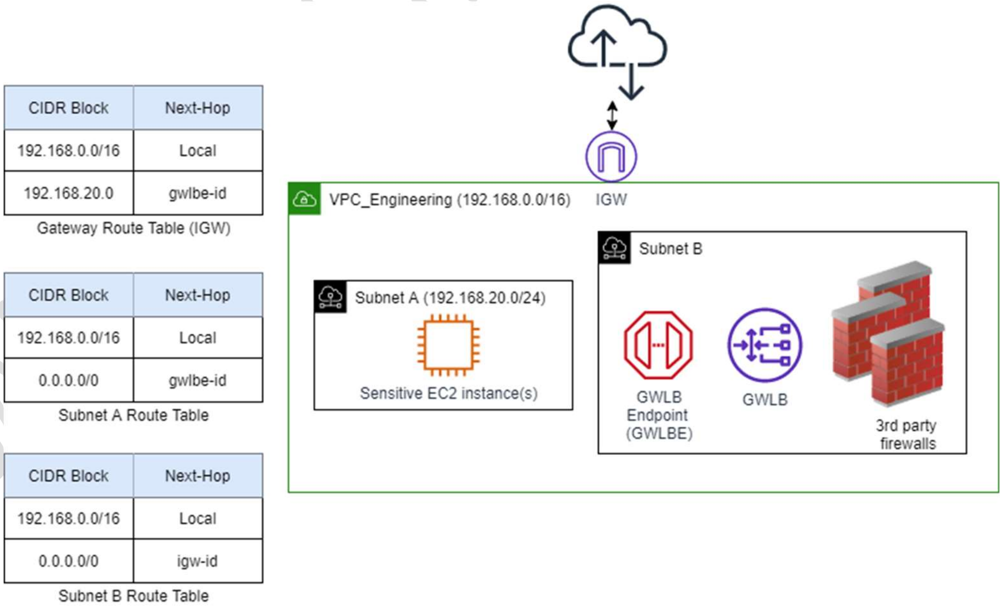


Figure 12

Reference architecture for a sandwich ELB design. Here the architect deploys a number of firewalls or cache engines between the two tiers of load balancers, but since they absolutely need sticky sessions everywhere, they would need different cookie names. The use of two different types of load balancers with two different sets of cookies satisfies this requirement. There are multiple mini-scenarios in the book to get to the bottom of this issue.



CIDR Block	Next-Hop
192.168.0.0/16	Local
192.168.20.0	gwlbe-id

Gateway Route Table (IGW)

CIDR Block	Next-Hop
192.168.0.0/16	Local
0.0.0.0/0	gwlbe-id

Subnet A Route Table

CIDR Block	Next-Hop
192.168.0.0/16	Local
0.0.0.0/0	igw-id

Subnet B Route Table

VPC_Engineering (192.168.0.0/16)
IGW
Subnet A (192.168.20.0/24)
Sensitive EC2 instance(s)
Subnet B
GWLB Endpoint (GWLBE)
GWLB
3rd party firewalls

Figure 13

Reference architecture for a basic GWLB deployment. Two obvious components: (1) GWLB endpoint (2) GWLB. This solution was an enhancement to the simple VPC ingress routing that came out about a year before that. The ingress routing feature, as seen in the gateway table, is still extensively used. As you will see later, this architecture also provided a bridge from simple deployments to AWS Network Firewall.

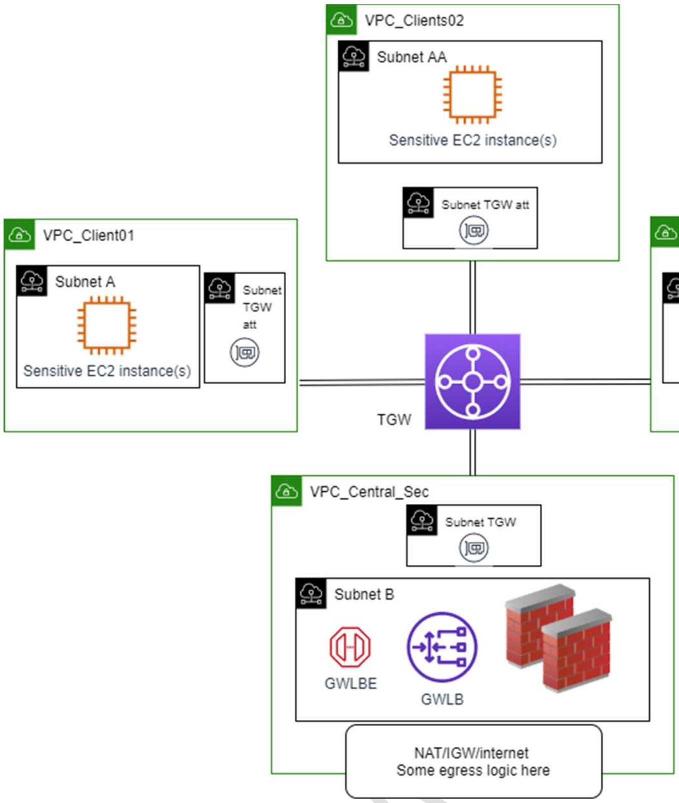


Figure 14B

Reference architecture to show further enhancements using transit gateways. Also, as shown here you can add a NAT construct to the design for more secure north-south traffic flow. A similar architecture with minimum changes can be used to inspect the east-west traffic between the VPCs. The architect needs to keep in mind that currently the maximum bandwidth per VPC connected to the transit gateway is around 50Gbps. Although this might not be an issue for VPC_clientX but needs to be carefully studied for VPC_Central_Sec.

Now, let's talk about something more important.

What if we cared more about availability and built more than one security pod in VPC_Central_Sec in two different AZs? We could also have different AZs in our client VPCs. Wouldn't be nice?

It would be a great idea but let's follow the flows and see where it might fail. As a network or cloud architect, you are expected to predict and avoid situations similar to this:

Host1 in AZ1 part of VPC_client01 is trying to establish a session with Host2 located in AZ2 part of VPC_client02. This is absolutely normal. The packets leave the host and make it to the TGW. Then, since we're inspecting the east-west traffic, the packets are sent to the central VPC, but you have two security pods in AZ1 and AZ2 and need to decide which path to take. The TGW, by design, favors the pod in AZ1 to keep the traffic local within one AZ; from AZ1 to AZ1. All good, we're smart and minimizing our data transfer charges. Subsequently, the packets pass the inspection process and eventually make it to Host2 in AZ2.

One leg of the mission is accomplished.

Host2, replies by sending its response back to the TGW. The TGW, again dutifully, favors the same AZ and sends the traffic to the pod in AZ2 for inspection. The stateful firewall has no trace or precedence of this session and drops the packets. You didn't get lucky in this particular case, but you could have.

We need a deterministic and reliable behavior.

The TGW Appliance Mode addresses this particular issue with asymmetric routing by sending the traffic to the firewall that has been working on the session and not to a random firewall just because it's located in the same AZ as the sending host.

There are quite a few scenarios in this book to discuss the details of this issue.

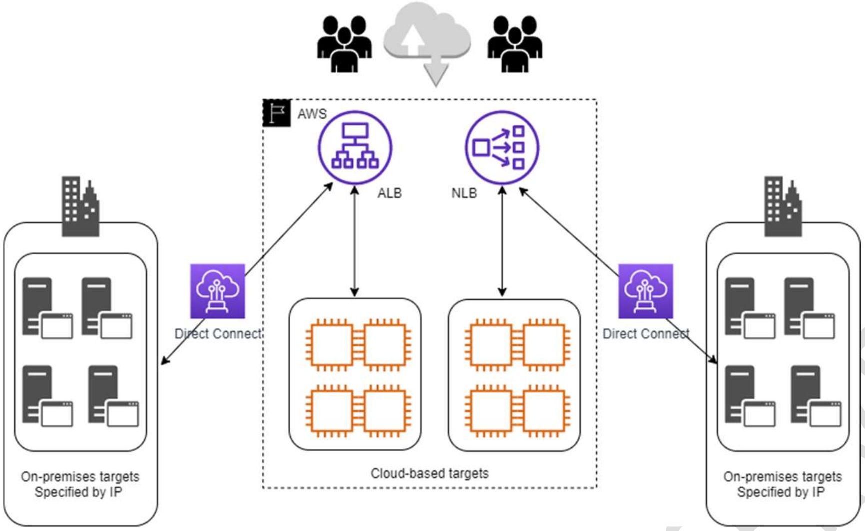


Figure 15

The reference architecture presented here demonstrates the fact that both NLB and ALB elastic load balancers can have targets registered by their private IP addresses. In addition to architecting for microservices, this feature comes in handy when the targets are split between AWS and on-premises. This might be for a range of reasons, including an ongoing migration project.