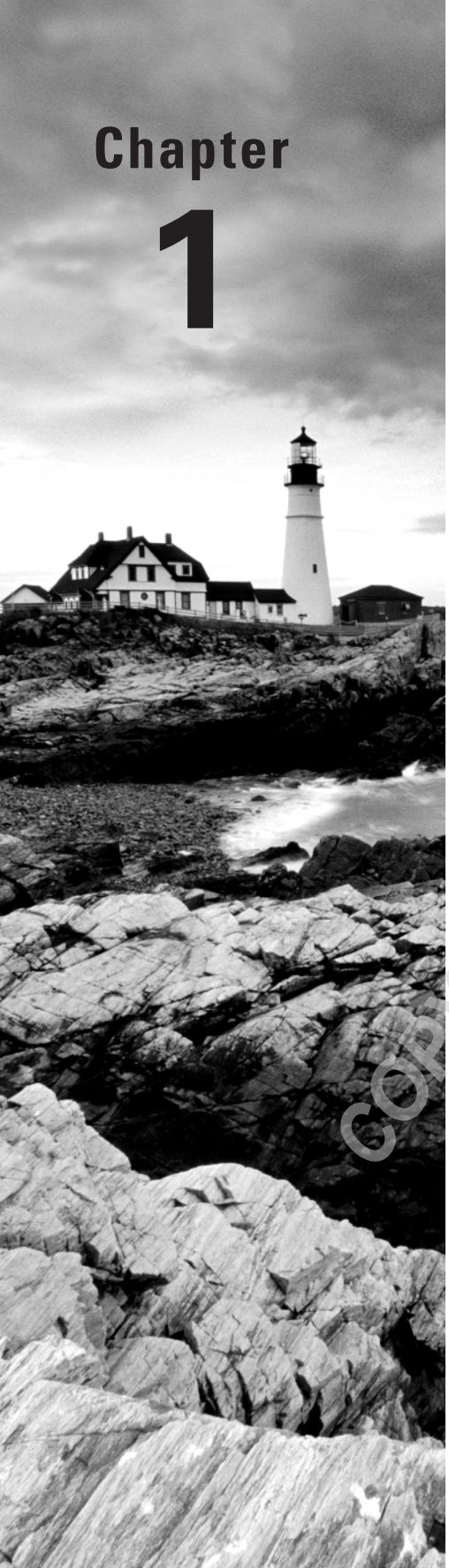


Chapter

1

Introduction to Advanced Networking

THE AWS CERTIFIED ADVANCED NETWORKING – SPECIALTY EXAM OBJECTIVES COVERED IN THIS CHAPTER MAY INCLUDE, BUT ARE NOT LIMITED TO, THE FOLLOWING:

Domain 1.0: Design and implement hybrid IT network architectures at scale

- ✓ **1.4 Evaluate design alternatives that leverage AWS Direct Connect**

Domain 2.0: Design and Implement AWS Networks

- ✓ **2.1 Advanced knowledge of AWS networking concepts**

Domain 4.0: Configure network integration with application services

- ✓ **4.1 Leverage the capabilities of Route 53**
- ✓ **4.4 Given a scenario, determine an appropriate load balancing strategy within the AWS ecosystem**
- ✓ **4.5 Determine a content distribution strategy to optimize for performance**

Domain 5.0: Design and implement for security and compliance

- ✓ **5.3 Evaluate AWS security features for managing network traffic**

Networks are foundational in our connected world. They are simultaneously critical to our everyday lives and frequently overlooked. Although network infrastructures, like the Internet, are likely the most distributed systems on Earth, they are not noticed unless they are operating poorly. This contrast makes networks quite interesting, both to learn about and to work with.

In addition to its distributed characteristics, modern networks are also a combination of new and old. The Internet Protocol (IP) and Transmission Control Protocol (TCP) were created in the 1970s, and though they have been updated over time, they still run the Internet. Meanwhile, new innovations, including advanced encapsulations, automation, and improved security mechanisms, continue to push the capabilities of the network forward. AWS has driven innovation in cloud networking with capabilities like Amazon Virtual Private Cloud (Amazon VPC), which provides customers with their own logical segment of the Amazon network—on demand and in minutes.

This study guide covers the breadth and depth of AWS networking in scope for the AWS Certified Advanced Networking – Specialty exam. The study guide reviews a broad array of topics relevant to the Amazon global infrastructure, various regional AWS networking features, on-premises hybrid networking, and AWS edge networking. The study guide's contents assume that you have a strong understanding of networking concepts and that you have successfully completed the AWS Certified Solutions Architect – Associate exam.

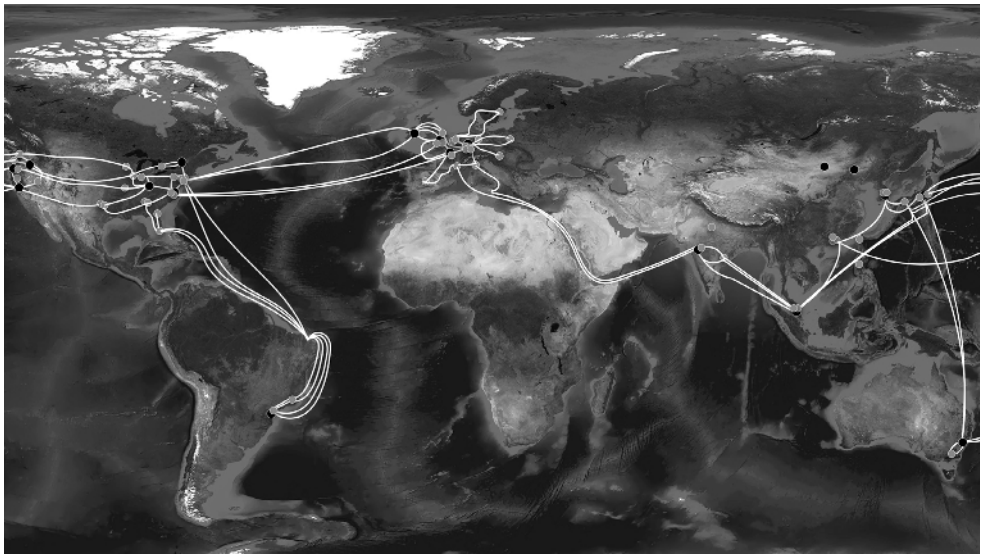
AWS Global Infrastructure

AWS operates a global infrastructure. This network is operated by one company, Amazon, and it spans the continents where AWS has a presence. This infrastructure enables traffic to flow between AWS Regions, Availability Zones, edge locations, and customer cross-connect facilities. Traffic between nodes on this network uses the AWS global infrastructure, with the exception of AWS GovCloud (US) and China. A representation of the global infrastructure is shown in Figure 1.1.

Regions

A *region* is a geographic area in the world where AWS operates cloud services (for example, Amazon Elastic Compute Cloud, also known as Amazon EC2).

FIGURE 1.1 AWS global infrastructure



AWS Regions are designed to be completely independent from other regions. This approach provides fault isolation, fault tolerance, and stability.

Most AWS Cloud services operate within a region. Since these regions are separated, you only see the resources tied to the region that you have specified. This design also means that customer content that you put into a region stays in that region unless you take an explicit action to move it.

Availability Zones

Each region is composed of two or more *Availability Zones*. Each Availability Zone contains one or more data centers. The zones are engineered such that they have different risk profiles. That is, AWS considers factors like power distribution, floodplains, and tectonics when placing Availability Zones within a region. The zones are connected to one another by low-latency, high-bandwidth fiber optics. Availability Zones are typically less than 2 milliseconds apart.

Amazon operates state-of-the-art, highly-available data centers. Although rare, failures can occur that affect the availability of resources that are in the same location. If you host all of your Amazon EC2 instances in a single location that is affected by such a failure, for example, none of your instances would be available. When you launch an Amazon EC2 instance, you can select an Availability Zone or let AWS choose one for you. If you distribute your instances across multiple Availability Zones and then one instance fails, you can design your application so that an instance in another zone can handle requests.

Edge Locations

To deliver content to end users with low latency, AWS provides a global network of edge locations. This content distribution network is called *Amazon CloudFront*. As end users make requests, the AWS Domain Name System (DNS), Amazon Route 53, routes requests to the Amazon CloudFront edge location that can best serve the user's request, typically the nearest edge location in terms of latency.

In the edge location, Amazon CloudFront checks its cache for the requested content. If the data is locally cached, Amazon CloudFront returns the content to the user. If the data is not in the cache, Amazon CloudFront forwards the request for the files to the applicable origin server for the corresponding file type. The origin servers then send the files back to the Amazon CloudFront edge location.

Amazon Virtual Private Cloud

Amazon Virtual Private Cloud (Amazon VPC) lets you provision a logically-isolated section of the AWS Cloud. You can launch AWS resources like Amazon EC2 instances in a virtual network that you define. You have complete control over your virtual networking environment, including selection of your own IP address range, creation of subnets, and configuration of route tables and network gateways. You can use both IPv4 and IPv6 in your Amazon VPC for secure and easy access to resources and applications.

VPC Mechanics

Amazon VPC enables you to launch resources into a logical network that you define. This network closely resembles the traditional networks that you operate in your own data centers, with the additional scalability and capability benefits of AWS. Amazon VPC uses many traditional concepts, like subnets, IP addresses, and stateful firewalls.

The underlying Amazon VPC mechanics differ, however, from the composition of standard, on-premises networking infrastructures. AWS built a custom network environment that satisfies the scale, performance, flexibility, and security requirements of the millions of active customers who use AWS each day. Consider that each customer has their own isolated network, and many customers are making thousands of changes per day. While the technology underlying Amazon VPC is not within the scope of the exam, understanding how it works will help you reason about its operation and functionality.

The Amazon VPC infrastructure is composed of various support components (such as the Amazon DNS server, instance metadata, and the Dynamic Host Configuration Protocol [DHCP] server) and the underlying physical servers onto which customers launch their Amazon EC2 instances. Each of these physical servers has its own IP address. As customers launch Amazon EC2 instances into their VPCs, AWS determines the physical server on which the instance will run. This decision is based on multiple factors, including the desired Availability Zone, instance type, instance tenancy, and whether the instance is part of a placement group. When different AWS accounts launch instances using Amazon VPC, these instances are not visible to each other.

Tenant isolation is a core function of Amazon VPC. In order to understand which resources are part of a given VPC, Amazon VPC uses a mapping service. The mapping service abstracts your VPC from the underlying AWS infrastructure. For any given VPC, the mapping service maintains information about all of its resources, their VPC IP addresses, and the IP addresses of the underlying physical server on which the resource is running. It is the definitive source of topology information for each VPC.

When an Amazon EC2 instance, say Instance A, in your VPC initiates communication with another Amazon EC2 instance, say Instance B, over IPv4, Instance A will broadcast an Address Resolution Protocol (ARP) packet to obtain the Instance B's Media Access Control (MAC) address. The ARP packet leaving Instance A is intercepted by the server Hypervisor. The Hypervisor queries the mapping service to identify whether Instance B exists in the VPC and, if so, obtains its MAC address. The Hypervisor returns a synthetic ARP response to Instance A containing Instance B's MAC address.

Instance A is now ready to send an IP packet to Instance B. The IP packet has Instance A's source IP and Instance B's destination IP. The IP packet is encapsulated in an Ethernet header with Instance A's MAC as the source address and Instance B's MAC as the destination address. The Ethernet packet is then transmitted from Instance A's network interface.

As Instance A emits the packet, it is intercepted by the server Hypervisor. The Hypervisor queries the mapping service to learn the IPv4 address of the physical server on which Instance B is running. Once the mapping service provides this data, the packet emitted by Instance A is encapsulated in a VPC header that identifies this specific VPC and then encapsulated again in an IP packet with a source IP address of Instance A's physical server and a destination IPv4 address of Instance B's physical server. The packet is then placed on to the AWS network.

When the packet arrives at Instance B's physical server, the outer IPv4 header and VPC header are inspected. The instance Hypervisor queries the mapping service to confirm that Instance A exists on the specific source physical server and in the specific VPC identified in the received packet. When the mapping service confirms that the mapping is correct, the Hypervisor strips off the outer encapsulation and delivers the packet that Instance A emitted to the Instance B network interface.

The details of packet exchange in Amazon VPC should provide you clarity on why, for example, Amazon VPC does not support broadcast and multicast. These same reasons explain why packet sniffing does not work. As you reason about Amazon VPC operation and functionality, consider this example.

Services Outside Your VPC

Many AWS Cloud services are provided from locations outside of your own VPC. These services are delivered from the following:

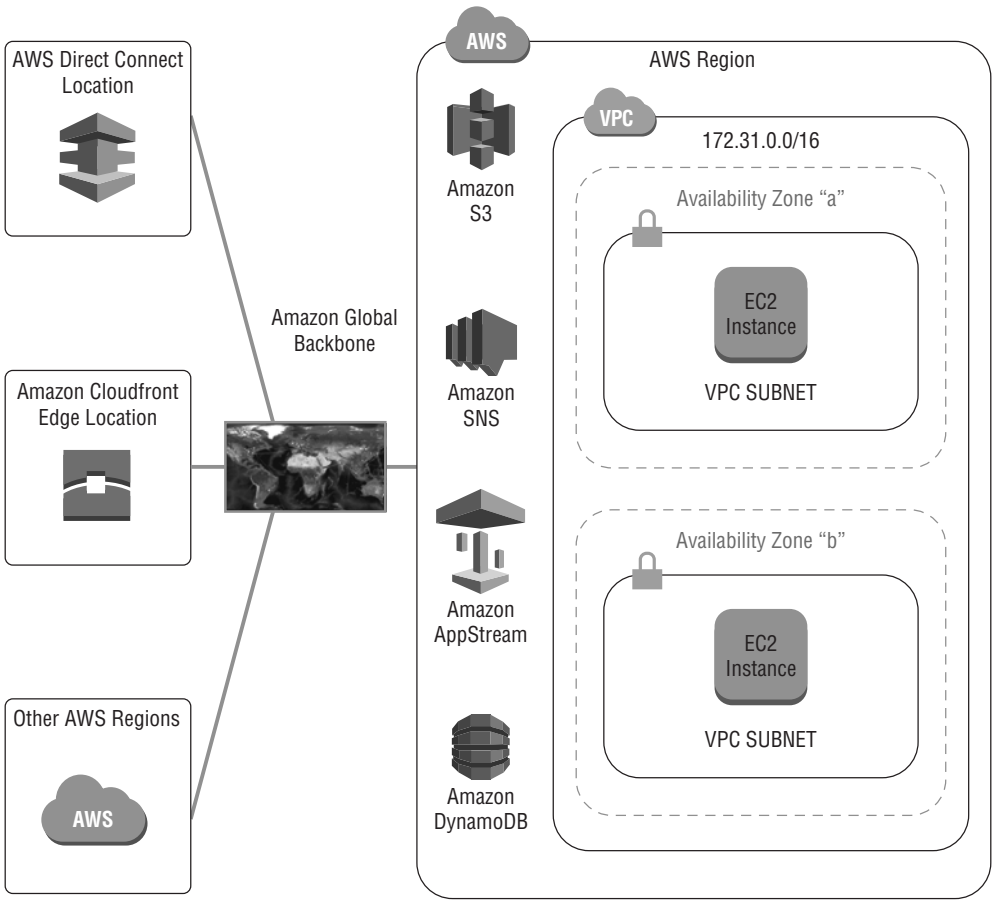
- Edge locations (for example, Amazon Route 53 and Amazon CloudFront)
- Directly inside your VPC (for example, Amazon Relational Database Service [Amazon RDS] and Amazon Workspaces)
- VPC Endpoints in your VPC (for example, Amazon DynamoDB and Amazon Simple Storage Service [Amazon S3])
- Public service endpoints outside your VPC (for example, Amazon S3 and Amazon Simple Queue Service [Amazon SQS])

AWS Cloud services use the same global infrastructure described earlier in this chapter. When you use services that are delivered directly on the Internet, such as edge locations and public service endpoints, you control network behaviors using service-specific mechanisms like policies and whitelists. When you use services that are exposed directly to your VPC, typically through a network interface or a VPC endpoint in your VPC, you may also use Amazon VPC features like security groups, network Access Control Lists (ACLs), and route tables in addition to service-specific mechanisms.

For the exam, you should understand how AWS Cloud services integrate into your overall network architecture and allow you to control network behavior. You do not need to understand the specific mechanisms that AWS uses to deliver services. However, understanding these delivery models will aid you in the development of scalable, performant, and highly-available architectures.

An overview of service locations can be seen in Figure 1.2.

FIGURE 1.2 Overview of the AWS service locations



AWS Networking Services

AWS provides many services that you can combine to meet business or organizational needs. This section introduces the AWS Cloud services specifically related to networking. Later chapters provide a deeper view of the services pertinent to the exam.

Amazon Elastic Compute Cloud

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It allows organizations to obtain and configure virtual servers in Amazon's data centers and to harness those resources to build and host software systems. Organizations can select from a variety of operating systems and resource configurations (for example, memory, CPU, and storage) that are optimal for the application profile of each workload. Amazon EC2 presents a true virtual computing environment, allowing organizations to launch compute resources with a variety of operating systems, load them with custom applications, and manage network access permissions while maintaining complete control.

Amazon Virtual Private Cloud

Amazon Virtual Private Cloud (Amazon VPC) lets organizations provision a logically-isolated section of the AWS Cloud where they can launch AWS resources in a virtual network that they define. Organizations have complete control over the virtual environment, including selection of the IP address range, creation of subnets, and configuration of route tables and network gateways. In addition, organizations can extend their corporate data center networks to AWS by using hardware or software *Virtual Private Network (VPN)* connections or dedicated circuits by using AWS Direct Connect. Amazon VPC is covered in depth in Chapter 2, "Amazon Virtual Private Cloud (Amazon VPC) and Networking Fundamentals," and Chapter 3, "Advanced Amazon Virtual Private Cloud."

AWS Direct Connect

AWS Direct Connect allows organizations to establish a dedicated network connection from their data center to AWS. Using AWS Direct Connect, organizations can establish private connectivity between AWS and their data center, office, or colocation (AWS) environment, which in many cases can reduce network costs, increase bandwidth throughput, and provide a more consistent network experience than Internet-based VPN connections. AWS Direct Connect is covered in depth in Chapter 5, "AWS Direct Connect."

Elastic Load Balancing

Elastic Load Balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances in the cloud. It enables organizations to achieve greater levels of fault tolerance in their applications, seamlessly providing the required amount of load balancing capacity needed to distribute application traffic. Elastic Load Balancing is covered in depth in Chapter 6, "Domain Name System and Load Balancing."

Amazon Route 53

Amazon Route 53 is a highly available and scalable DNS service. It is designed to give developers and businesses an extremely reliable and cost-effective way to route end users to Internet applications by translating human-readable names, such as `www.example.com`, into the numeric IP addresses, such as 192.0.2.1, which computers use to connect to each other. Amazon Route 53 also serves as a domain registrar, allowing customers to purchase and manage domains directly from AWS. Amazon Route 53 is covered in depth in Chapter 6.

Amazon CloudFront

Amazon CloudFront is a global Content Delivery Network (CDN) service that securely delivers data, videos, applications, and Application Programming Interfaces (APIs) to an organization's viewers with low latency and high transfer speeds. Amazon CloudFront is integrated with AWS, both with physical locations that are directly connected to the AWS global infrastructure and software that works seamlessly with other AWS Cloud services. These include AWS Shield for Distributed Denial of Service (DDoS) mitigation, Amazon S3, Elastic Load Balancing, or Amazon EC2 as origins for applications, as well as AWS Lambda to run custom code close to the content viewers. Amazon CloudFront is covered in depth in Chapter 7, "Amazon CloudFront."

GuardDuty

GuardDuty is a continuous security monitoring, threat detection solution that gives customers visibility into malicious or unauthorized activity across their AWS accounts and the applications and services running within them. GuardDuty is capable of detecting threats such as reconnaissance by attackers (for example, port probes, port scans, and attempts to obtain account credentials), Amazon EC2 instances that have been compromised (such as instances serving malware, bitcoin mining, and outbound DDoS attacks), and compromised accounts (for example, unauthorized infrastructure deployments, AWS CloudTrail tampering, and unusual API calls). When a threat is detected, the solution delivers a security finding. Each finding includes a severity level, detailed evidence for the finding, and recommended actions. GuardDuty is covered in depth in Chapter 8, "Network Security."

AWS WAF

AWS WAF helps protect web applications from common attacks and exploits that could affect application availability, compromise security, or consume excessive resources. AWS WAF gives organizations control over which traffic to allow or block to their web applications by defining customizable web security rules. AWS WAF is covered in depth in Chapter 8.

AWS Shield

AWS Shield is a managed DDoS protection service that safeguards web applications running on AWS. AWS Shield provides always-on detection and automatic inline mitigations

that minimize application downtime and latency. There are two tiers of AWS Shield: Standard and Advanced. All AWS customers benefit from the automatic protections of AWS Shield Standard at no additional charge. AWS Shield Standard defends against the most common, frequently occurring network and transport layer DDoS attacks that target websites or applications. AWS Shield is covered in depth in Chapter 8.

Summary

AWS provides highly-available technology infrastructure services with multiple locations worldwide. These locations are composed of regions and Availability Zones. AWS provides networks and network features spanning edge locations, VPCs, and hybrid networks. AWS operates a global network connecting these locations.

Amazon VPC provides complete control over a virtual networking environment, enabling secure and easy access to resources and applications.

This chapter introduced the primary services related to networking on AWS. This chapter also provided the background and context so that you can understand more advanced networking introduced later in this study guide.

Resources to Review

For further review, check out the following URLs:

AWS Global Infrastructure:

`https://aws.amazon.com/about-aws/global-infrastructure/`

Amazon EC2:

`https://aws.amazon.com/ec2/`

Amazon VPC:

`https://aws.amazon.com/vpc/`

AWS Direct Connect:

`https://aws.amazon.com/directconnect/`

Elastic Load Balancing:

`https://aws.amazon.com/ elasticloadbalancing/`

Amazon Route 53:

`https://aws.amazon.com/route53/`

Amazon CloudFront:

`https://aws.amazon.com/cloudfront/`

AWS WAF:

`https://aws.amazon.com/waf/`

AWS Shield:

`https://aws.amazon.com/shield/`

Exam Essentials

Understand the global infrastructure. AWS operates a global infrastructure. This network is operated by one company, Amazon. This infrastructure enables traffic to flow between regions, Availability Zones, edge locations, and customer cross-connect facilities. Traffic between nodes on this network uses the AWS global infrastructure.

Understand regions. A region is a geographic area in the world where AWS operates cloud services such as Amazon EC2. AWS Regions are designed to be completely independent from other regions. Most AWS Cloud services operate within a region. Since these regions are separated, content you put into a region stays in that region, unless you take an explicit action to move it.

Understand Availability Zones. An Availability Zone consists of one or more data centers within a region, which are designed to be isolated from failures in other Availability Zones. Availability Zones provide inexpensive, low-latency, high-bandwidth network connectivity to other zones in the same region. By placing resources in separate Availability Zones, you can protect your website or application from a service disruption affecting a single location.

Understand Amazon VPC. Amazon VPC is an isolated, logical network in the AWS infrastructure. A VPC contains resources, such as Amazon EC2 instances. There is a VPC mapping service that enables the routing capability inside a VPC.

Understand how AWS Cloud service integration works. You should understand how AWS Cloud services integrate into your overall network architecture and how to control network behavior. You do not need to understand the specific mechanisms that AWS uses to deliver services. Understanding these delivery models, however, will aid you in the development of scalable, performant, and highly-available architectures.

Test Taking Tip

Manage your time wisely when taking this exam. Don't waste time on questions where you are stumped. Mark it for later review and move on. Plan on leaving time at the end of the exam for review. Go through each marked question to answer any that you may have skipped or to make sure that you are still happy with previously-marked answers.

Exercise

EXERCISE 1.1

Review Network Service Documentation

Navigate to all of the URLs in the resources to review the section above and review the network service product material.

1. Navigate to the AWS Global Infrastructure website. Review the information provided about AWS Regions and Availability Zones. Become familiar with the AWS Global Infrastructure.
2. Navigate to the Amazon VPC product documentation. Review the product details and FAQs. Become familiar with the additional product documentation in the related links section.
3. Navigate to the AWS Direct Connect product documentation. Review the product details and FAQs. Become familiar with the additional product documentation in the related links section.
4. Navigate to the Elastic Load Balancing product documentation. Review the product details and FAQs. Become familiar with the additional product documentation section.
5. Navigate to the Amazon Route53 product documentation. Review the product details and FAQs. Become familiar with the additional product documentation section.
6. Navigate to the Amazon CloudFront product documentation. Review the product details and FAQs. Become familiar with the additional product documentation section.
7. Navigate to the AWS WAF product documentation. Review the product details and FAQs. Become familiar with the additional product documentation section.
8. Navigate to the AWS Shield product documentation. Review the product details and FAQs. Become familiar with the additional product documentation section.

After completing this exercise, you will be familiar with AWS network-related products, where to find related documentation, and the different types of additional documentation that AWS provides.

Review Questions

1. Which of the following services provides private connectivity between AWS and your data center, office, or colocation environment?
   - **A.** Amazon Route 53
   - **B.** AWS Direct Connect
   - **C.** AWS WAF
   - **D.** Amazon Virtual Private Cloud (Amazon VPC)
2. Which AWS Cloud service uses edge locations to deliver content to end users?
   - **A.** Amazon Virtual Private Cloud (Amazon VPC)
   - **B.** AWS Shield
   - **C.** Amazon CloudFront
   - **D.** Amazon Elastic Compute Cloud (Amazon EC2)
3. Which of the following statements is true?
   - **A.** AWS Regions consist of multiple edge locations.
   - **B.** Edge locations consist of multiple Availability Zones.
   - **C.** Availability Zones consist of multiple AWS Regions.
   - **D.** AWS Regions consist of multiple Availability Zones.
4. Which of the following describes a physical location around the world where AWS clusters data centers?
   - **A.** Endpoint
   - **B.** Collection
   - **C.** Fleet
   - **D.** Region
5. What feature of AWS Regions allows you to operate production systems that are more highly available, fault-tolerant, and scalable than is possible using a single data center?
   - **A.** Availability Zones
   - **B.** Replication areas
   - **C.** Geographic districts
   - **D.** Compute centers
6. What AWS Cloud service provides a logically-isolated section of the AWS Cloud where you can launch AWS resources in a logical network that you define?
   - **A.** Amazon Simple Workflow Service (Amazon SWF)
   - **B.** Amazon Route 53
   - **C.** Amazon Virtual Private Cloud (Amazon VPC)
   - **D.** AWS CloudFormation

7. Which AWS Cloud service provides Distributed Denial of Service (DDoS) mitigation?
   - **A.** AWS Shield
   - **B.** Amazon Route 53
   - **C.** AWS Direct Connect
   - **D.** Amazon Elastic Compute Cloud (Amazon EC2)

8. How many companies operate the AWS global infrastructure?
   - **A.** 1
   - **B.** 2
   - **C.** 3
   - **D.** 4

9. Amazon Virtual Private Cloud (Amazon VPC) enables which one of the following?
   - **A.** Connectivity from your on-premises network
   - **B.** Creation of a logical network defined by you
   - **C.** Edge caching of user content
   - **D.** Network threshold alarms

10. Which Amazon Virtual Private Cloud (Amazon VPC) component maintains a current topology map of the customer environment?
    - **A.** Route table
    - **B.** Mapping service
    - **C.** Border Gateway Protocol (BGP)
    - **D.** Interior Gateway Protocol (IGP)

11. You may specify which of the following when creating a Virtual Private Cloud (VPC)?
    - **A.** AWS data centers to use
    - **B.** 802.1x authentication methods
    - **C.** Virtual Local Area Network (VLAN) tags
    - **D.** IPv4 address range

12. Amazon Route 53 allows you to perform which one of the following actions?
    - **A.** Create subnets
    - **B.** Register domains
    - **C.** Define route tables
    - **D.** Modify stateful firewalls

13. Which service provides a more consistent network experience when connecting to AWS from your corporate network?
    - **A.** AWS Direct Connect
    - **B.** Amazon CloudFront
    - **C.** Internet-based Virtual Private Network (VPN)
    - **D.** Amazon Route 53

14. Which AWS Cloud service enables you to define customizable web security rules?
    - **A.** Amazon Route 53
    - **B.** AWS Shield
    - **C.** AWS WAF
    - **D.** GuardDuty

15. Which service increases the fault tolerance of your Amazon Elastic Compute Cloud (Amazon EC2) applications on AWS?
    - **A.** AWS Direct Connect
    - **B.** Elastic Load Balancing
    - **C.** AWS Shield
    - **D.** AWS WAF